

CONNECTICUT INSURANCE LAW JOURNAL

Volume 16, Number 1
Fall 2009



University of Connecticut School of Law
Hartford, Connecticut

*Connecticut Insurance Law Journal (ISSN 1081-9436) is published at least twice a year by the Connecticut Insurance Law Journal Association at the University of Connecticut School of Law. Periodicals postage paid at Hartford, Connecticut. Known office of publication: 55 Elizabeth Street, Hartford, Connecticut 06105-2209. **Printing location: Western Newspaper Publishing Company, 537 East Ohio Street, Indianapolis, Indiana 46204.***

Please visit our website at <http://www.insurancejournal.org> or see the final page of this issue for subscription and back issue ordering information.

Postmaster: Send address changes to Connecticut Insurance Law Journal, 55 Elizabeth Street, Hartford, Connecticut 06105-2209.

The Journal welcomes the submission of articles and book reviews. Both text and notes should be double or triple-spaced. Submissions in electronic form are encouraged, and should be in Microsoft™ Word™ version 97 format or higher. Citations should conform to the most recent edition of A UNIFORM SYSTEM OF CITATION, published by the Harvard Law Review Association.

It is the policy of the University of Connecticut to prohibit discrimination in education, employment, and in the provision of services on the basis of race, religion, sex, age, marital status, national origin, ancestry, sexual preference, status as a disabled veteran or veteran of the Vietnam Era, physical or mental disability, or record of such impairments, or mental retardation. University policy also prohibits discrimination in employment on the basis of a criminal record that is not related to the position being sought; and supports all state and federal civil rights statutes whether or not specifically cited within this statement.

Copyright © 2009 by the Connecticut Insurance Law Journal Association.

Cite as CONN. INS. L.J.

CONNECTICUT INSURANCE LAW JOURNAL

VOLUME 16

2009-2010

NUMBER 1

EDITORIAL BOARD 2009-2010

Editor-in-Chief
MARK KOCHANOWICZ

Managing Editor
AGRON ETEMI

Administrative Editor
GWAINA WAULDON

Assistant Managing Editor
LINDSAY CORLISS

Lead Articles Editors
ALANA BARTLEY
SARA NADIM

Notes & Comments Editors
SCOTT SCHAEFFER
CAROLINE VASQUEZ
JESSA YANNI-LAZARUS

Executive Editors
KELLY BARRETT
STEVEN COOPER
RACHEL GINSBURG
SUSAN KIRKEBY
MELISSA WONG

Legal Abstracts Editor
ANNA HEHENBERGER

Research Editor
LOUIS CRUZ

Associate Editors
MARGARET GOODELL
DOUGLAS SAUVE
DOMINIC SPINELLI

Symposium Editor
ERIN O'LEARY

Technology Editor
TIM NAST

MEMBERS

ADAM J. ALLEGRO
KENSLEY BARRETT
OLIVER BOWERS
KENNEX CHAN
DANIEL CSUKA
RAPHAEL DEUTSH
SHANIQUE FENLATOR
CRYSTAL FRASER

NATHANIEL GENTILE
JONATHAN HARDING
K. RUSSELL HASAN
RYAN LORUSSO
WILLIAM MAUKE
T. RYAN MCNEIL
JAMES MOHER
JOHN NEUMON

SHETERA PATTERSON
JONATHAN PORZUC
GREGORY RADA
SANA SHEIKH
MICHAEL YEDINAK
STEVEN ZAKRZEWSKI
ZHIJUN ZHANG

FACULTY ADVISOR
PATRICIA MCCOY

**UNIVERSITY OF CONNECTICUT
SCHOOL OF LAW**

**FACULTY AND OFFICERS OF ADMINISTRATION
FOR THE ACADEMIC YEAR 2009-2010**

OFFICERS OF ADMINISTRATION

Michael J. Hogan, M.A., Ph.D., *President, University of Connecticut*
Peter J. Nicholls, Ph.D., *Provost and Executive Vice President for Academic Affairs*
Jeremy Paul, J.D., *Dean, School of Law*
Anne Dailey, J.D., *Associate Dean for Academic Affairs*
Michael Fischl, J.D., *Associate Dean for Research and Faculty Development*
Darcy Kirk, J.D., *Associate Dean for Library and Technology*
Ann Crawford, M.Ed., *Acting Associate Dean for Finance and Administration*
Ellen K. Rutt, J.D., *Associate Dean for Admissions, Student Finance and Career Services*
Karen L. Demeola, J.D., *Assistant Dean for Admissions and Student Finance*

FACULTY EMERITI

Robert L. Bard, B.A., M.A., LL.B., J.S.D., *Professor of Law, Emeritus*
Phillip I. Blumberg, A.B., J.D., LL.D. (Hon.), *Dean and Professor of Law and Business, Emeritus*
John C. Brittain, B.A., J.D., *Professor of Law, Emeritus*
Clifford Davis, S.B., LL.B., *Professor of Law, Emeritus*
Nell J. Newton, B.A., J.D., *Dean and Professor of Law, Emeritus*
Howard Sacks, A.B., LL.B., *Dean and Professor of Law, Emeritus*
George Schatzki, A.B., LL.B., LL.M., *Dean and Professor of Law, Emeritus*
Craig Shea, A.B., LL.B., LL.M., *Professor of Law, Emeritus*
Phillip Shuchman, B.A., M.A., LL.B., *Professor of Law, Emeritus*
Eileen Silverstein, A.D., J.D., *Professor of Law Emeritus*
Lester B. Snyder, B.S., LL.B., LL.M., *Professor of Law, Emeritus*
Terry Tondro, A.B., LL.B., M.Phil., *Professor of Law, Emeritus*
Nicholas Wolfson, A.B., J.D., *George and Helen England Professor of Law, Emeritus*

FACULTY OF LAW

Jill Anderson, B.A., University of Washington; J.D., Columbia University; *Visiting Assistant Professor of Law*
Paul Bader, B.A., Duke University; J.D., Mercer University Walter F. George School of Law; *Assistant Clinical Professor of Law*
Robin D. Barnes, B.A., J.D., State University of New York at Buffalo; LL.M., University of Wisconsin; *Professor of Law*
Jon Bauer, A.B., Cornell University; J.D., Yale University; *Clinical Professor of Law*
Loftus E. Becker, Jr., A.B., Harvard College; LL.B., University of Pennsylvania; *Professor of Law*
Bethany Berger, B.A., Wesleyan University; J.D., Yale University; *Professor of Law*
Robert Birmingham, A.B., J.D., Ph.D. (Econ.), Ph.D. (Phil.), University of Pittsburgh; LL.M., Harvard University; *Professor of Law*
William Breetz, B.A., Dartmouth College; LL.B., University of Virginia; *Executive Director, Connecticut Urban Legal Initiative, Inc.*

Sara Bronin, B.A., University of Texas; M.Sc., University of Oxford (Magdalen College); J.D., Yale University; *Associate Professor of Law*

Deborah A. Calloway, B.A., Middlebury College; J.D., Georgetown University; *Professor of Law*

Marcia Canavan, B.A., University of California, Los Angeles; MPH., University of California, Los Angeles; J.D., Golden Gate University School of Law and University of Colorado School of Law; *Assistant Clinical Professor of Law*

Paul Chill, B.A., Wesleyan University; J.D., University of Connecticut; *Clinical Professor of Law*

Anne C. Dailey, B.A., Yale University; J.D., Harvard University; *Evangeline Starr Professor of Law and Associate Dean for Academic Affairs*

Steven M. Davidoff, B.A., University of Pennsylvania; M.A., London Business School; J.D., Columbia University; *Associate Professor of Law*

Geoffrey Dellenbaugh, A.B., Princeton University; M.A., Stanford University; Ph.D., University of Pittsburgh; J.D., University of Chicago; *Assistant Clinical Professor of Law*

Timothy H. Everett, B.A., M.A., Clark University; J.D., University of Connecticut; *Clinical Professor of Law*

Todd D. Fernow, B.A., Cornell University; J.D., University of Connecticut; *Professor of Law and Director, Clinical Programs*

Richard Michael Fischl, B.A., University of Illinois; J.D., Harvard University; *Professor of Law and Associate Dean for Research and Faculty Development*

Hillary Greene, B.A., J.D., Yale University; *Associate Professor of Law and Director, Intellectual Property Clinic*

Kaaryn Gustafson, A.B., Harvard University; Ph.D., J.D., University of California, Berkeley; *Associate Professor of Law*

Mark W. Janis, A.B., Princeton University; B.A., M.A., Oxford University; J.D., Harvard University; *William F. Starr Professor of Law*

Richard S. Kay, A.B., Brandeis University; M.A., Yale University; J.D., Harvard University; *Wallace Stevens Professor of Law*

Darcy Kirk, A.B., Vassar College; M.S., M.B.A., Simmons College; J.D., Boston College; *Professor of Law and Associate Dean for Library & Technology*

Peter R. Kochenburger, A.B., Yale University; J.D., Harvard University; *Park Associate Clinical Professor and Director of Graduate Programs*

Lewis S. Kurlantzick, B.A., Wesleyan University; LL.B., Harvard University; *Zephaniah Swift Professor of Law*

Alexandra Lahav, B.A., Brown University; J.D., Harvard University; *Professor of Law*

Leslie C. Levin, B.S.J., Northwestern University; J.D., Columbia University; *Professor of Law*

Diana L. Leyden, B.A., Union College; J.D., University of Connecticut; LL.M., Georgetown University; *Associate Clinical Professor of Law and Director, Tax Clinic*

Peter L. Lindseth, B.A., J.D., Cornell University; M.A., M. Phil, Ph.D, Columbia University; *Professor of Law*

Justin Long, A.B., Harvard University; J.D., University of Pennsylvania; *Visiting Assistant Professor of Law*

Hugh C. Macgill, B.A., Yale University; LL.B., University of Virginia; *Oliver Ellsworth Research Professor of Law and Dean Emeritus*

Neysun Mahboubi, A.B., Princeton University; J.D., Columbia University; *Visiting Assistant Professor of Law*

Jennifer Brown Mailly, B.A., Brown University; J.D., Ohio State University; *Assistant Clinical Professor of Law*

Margaret Martin, B.A., Boston University; J.D., Columbia University; *William R. Davis Clinical Teaching Fellow*

Ruth Mason, B.A., Columbia University; J.D., Harvard University; *Associate Professor of Law*

Patricia A. McCoy, B.A., Oberlin College; J.D., University of California, Berkeley; *Director of Insurance Law Center and George J. and Helen M. England Professor of Law*

Barbara McGrath, B.A., Yale University; J.D., University of Connecticut; *Assistant Director, Connecticut Urban Legal Initiative, Inc.*

Willajeanne F. McLean, B.A., Wellesley College; B.S., University of Massachusetts; J.D., Fordham University; LL.M., Free University of Brussels; *Professor of Law*

Thomas H. Morawetz, A.B., Harvard College; J.D., M.Phil., Ph.D., Yale University; *Tapping Reeve Professor of Law and Ethics*

Lily Neff, B.S.E.E., City College of New York; J.D., Pace University; *Assistant Clinical Professor of Law*

R. Kent Newmyer, B.A., Doane College; Ph.D., University of Nebraska; *Professor of Law and History*

Angel R. Oquendo, A.B., M.A., Ph.D., Harvard University; J.D., Yale University; *Olimpiad S. Ioffe Professor of Law*

Leonard Orland, B.A., Rutgers University; LL.B., University of Pennsylvania; *Oliver Ellsworth Research Professor of Law*

Sachin Pandya, B.A., University of California, Berkeley; M.A., Columbia University; J.D., Yale University; *Associate Professor of Law*

Richard W. Parker, A.B., Princeton University; J.D., Yale University; D.Phil., Oxford University; *Professor of Law*

Jeremy Paul, A.B., Princeton University; J.D., Harvard University; *Dean and Thomas F. Gullivan, Jr. Professor of Real Property Law*

Ellen Ash Peters, B.A., Swarthmore College; LL.B., Yale University; LL.D., Yale University; University of Connecticut; et al.; *Visiting Professor of Law*

Richard D. Pomp, B.S., University of Michigan; J.D., Harvard University; *Alva P. Loisselle Professor of Law*

Jessica S. Rubin, B.S., J.D., Cornell University; *Assistant Clinical Professor of Law*

Susan R. Schmeiser, A.B., Princeton University; J.D., Yale University; Ph.D., Brown University; *Professor of Law*

Peter Siegelman, B.A., Swarthmore College; M.S.L., Ph.D., Yale University; *Roger Sherman Professor of Law*

James H. Stark, A.B., Cornell University; J.D., Columbia University; *Professor of Law and Director, Mediation Clinic*

Martha Stone, B.A., Wheaton College; J.D., LL.M., Georgetown University; *Director, Center for Children's Advocacy*

Kurt A. Strasser, B.A., J.D., Vanderbilt University; LL.M., J.S.D., Columbia University; *Phillip I. Blumberg Professor of Law*

Colin C. Tait, A.B., Cornell University; LL.B., Yale University; *Oliver Ellsworth Research Professor of Law*

Stephen G. Utz, B.A., Louisiana State University; J.D., University of Texas; Ph.D., Cambridge University; *Professor of Law*

Carol Ann Weisbrod, J.D., Columbia University; *Ellen Ash Peters Professor of Law*

Robert Whitman, B.B.A., City College of New York; J.D., Columbia University; LL.M., New York University; *Professor of Law*

Steven Wilf, B.S., Arizona State University; Ph.D., J.D., Yale University; *Professor of Law*

Richard A. Wilson, BSc., Ph.D., London School of Economics and Political Science;
Professor of Law

ADJUNCT FACULTY OF LAW

Peter G. Austin, B.S., University of Connecticut; M.S., University of Hartford; *Adjunct Professor of Law*

Morris W. Banks, A.B., Dartmouth College; LL.B., Columbia University; LL.M., New York University; *Adjunct Professor of Law*

Anne Davis Barry, B.S., University of Connecticut; M.S., Union College; J.D., University of Connecticut; *Lecturer in Law*

Richard Baxter, B.A., Bucknell University; M.A., Yale University; J.D., Duke University; *Lecturer in Law*

James W. Bergenn, B.A., Catholic University; J.D., Columbia University; *Adjunct Professor of Law*

Hon. David M. Borden, B.A., Amherst College; LL.B., Harvard University; *Adjunct Professor of Law*

John Buchanan, A.B., Princeton, B.A., Oxford; J.D., Harvard; *Lecturer in Law*

Michael A. Cantor, B.S., J.D., University of Connecticut; *Adjunct Professor of Law*

Christine Chabot, B.A., Northwestern University; J.D., Notre Dame Law School; *Lecturer in Law*

James N. Cohen, B.A., J.D., Harvard; *Lecturer in Law*

Gary Collins, B.A., State University of New York, Buffalo; J.D., Vanderbilt University; *Lecturer in Law*

David W. Cooney, B.A., Rutgers University; M.A., Wesleyan University; J.D., University of Connecticut; *Lecturer in Law*

Dean M. Cordiano, B.A., State University of New York at Binghamton; J.D., Duke University; *Lecturer in Law*

Michelle Cruz, B.A., M.A., Mount Holyoke College; J.D., University of Connecticut; *Lecturer in Law*

Renee M. Dailey, B.A., J.D., University of Connecticut; *Lecturer in Law*

John G. Day, A.B., Oberlin College; J.D., Case Western Reserve University; *Professor in Residence*

Raymond DeMeo, B.A., Williams College; J.D., University of Connecticut; *Lecturer in Law*

Mark Dubois, B.A., College of the Holy Cross; J.D., University of Connecticut; *Lecturer in Law*

Gerald P. Dwyer, A.B., Colgate University; J.D., Gonzaga University; LL.M., Georgetown University; *Lecturer in Law*

Paul Eddy, B.A., Ohio Wesleyan University; J.D., University of Connecticut; *Lecturer in Law*

Thomas Farrish, B.A., J.D., University of Connecticut; *Lecturer in Law*

Rebecca Flanagan, B.A., M.A., University of Connecticut; J.D., University of North Carolina; *Lecturer in Law*

Evan D. Flaschen, B.A., Wesleyan University; J.D., University of Connecticut; *Adjunct Professor of Law*

Charles W. Fortune, B.A., University of Rochester; J.D., University of Connecticut; *Lecturer in Law*

Michael A. Gailor, B.A., Cornell University; J.D., University of Connecticut; *Lecturer in Law*

William Goddard, B.A., M.B.A., Dartmouth; J.D., University of Connecticut; *Lecturer in Law*

Ira H. Goldman, B.A., Cornell University; J.D., Yale University; *Adjunct Professor of Law*

Daniel L. Gottfried, B.A., J.D., University of Connecticut, *Lecturer in Law*
Andrew S. Groher, B.A., University of Virginia; J.D., University of Connecticut; *Adjunct Professor of Law*
Albert B. Harper, B.A., University of Texas; J.D., Ph.D., University of Connecticut; *Adjunct Professor of Law*
John Harris, B.S.E., Princeton; J.D., University of Wisconsin; *Adjunct Professor of Law*
Wesley Horton, B.A., Harvard University; J.D., University of Connecticut; *Adjunct Professor of Law*
John Houlihan, B.A., Providence College; J.D., St. John's University; *Lecturer in Law*
Daniel Klau, B.A., University of California, San Diego; J.D., Boston University; *Lecturer in Law*
Charles H. Klippel, B.A., M.Ed., M.P.H., J.D., Harvard University; *Lecturer in Law*
Barry Kramer, B.Ch.E., Rensselaer Polytechnic Institute; J.D., New York University; *Lecturer in Law*
John H. Lawrence, Jr., B.S., Washington and Lee University; J.D., University of Virginia; *Lecturer in Law*
Henry C. Lee, B.S., John Jay College of Criminal Justice; M.S., Ph.D., New York University; Dr.Sci. (Hon.), University of New Haven; Dr.Hum. (Hon.), St. Joseph College; *Adjunct Professor of Law*
Erik T. Lohr, B.S., Thomas A. Edison State College; J.D., University of Connecticut; *Lecturer in Law*
James Lotstein, B.S., Northwestern University; J.D., University of Connecticut, *Lecturer in Law*
Joseph A. MacDougald, B.A., Brown University; M.B.A., New York University; J.D., University of Connecticut; M.E.M., Yale University; *Lecturer in Law*
Robert G. Madden, B.S.W., Providence College; M.S.S.W., Columbia University School of Social Work; J.D., University of Connecticut, *Lecturer in Law*
Thomas S. Marrion, A.B., College of the Holy Cross; J.D., University of Connecticut; *Adjunct Professor of Law*
James F. Meehan, B.A., University of Arizona; J.D., University of Connecticut; *Lecturer in Law*
Dwight Merriam, B.A., University of Massachusetts, M.R.P., University of North Carolina, J.D., Yale University; *Lecturer in Law*
Jamie L. Mills, B.A., Eastern Connecticut State University; J.D., University of Connecticut; *Lecturer in Law*
Joseph Mirrione, B.A., Marist College; J.D., Vermont Law School; *Lecturer in Law*
Derek L. Mogck, B.A., Gordon College; M.S., University of Massachusetts; J.D., University of Connecticut; *Lecturer in Law*
Umar F. Moghul, B.A., M.A., University of Pennsylvania; J.D., Temple University; *Lecturer in Law*
Thomas B. Mooney, B.A., Yale University; J.D., Harvard University; *Adjunct Professor of Law*
Ronald T. Murphy, B.S., Springfield College; J.D., University of Connecticut; *Lecturer in Law*
Henry F. Murray, B.A., Yale University; J.D. University of Connecticut; *Lecturer in Law*
Kevin O'Connor, B.A., University of Notre Dame; J.D., University of Connecticut; *Lecturer in Law*
Andrew J. O'Keefe, B.S., College of the Holy Cross; J.D., University of Connecticut; *Adjunct Professor of Law*
Cornelius O'Leary, B.A., Williams College; M.A., Trinity College; J.D., University of

Connecticut; *Adjunct Professor of Law and Mark A. Weinstein Clinical Teaching Fellow*

John Pagini, B.A., Central Connecticut State University, M.P.A., University of Hartford; *Lecturer in Law*

Rosemarie Paine, B.S., Southern Connecticut State University; J.D., University of Connecticut; *Lecturer in Law*

Humbert J. Polito, Jr., A.B., College of the Holy Cross; J.D., University of Connecticut; *Adjunct Professor of Law*

Elliott B. Pollock, A.B., Columbia College; LL.B., Columbia Law School; *Lecturer in Law*

Hon. Elliot D. Prescott, B.A., University of Massachusetts; J.D., University of Connecticut; *Lecturer in Law*

Renee C. Redman, B.A., Michigan State University; J.D., Brooklyn Law School; *Lecturer in Law*

Leah Reimer, B.S., J.D., Ph.D., University of Connecticut; *Lecturer in Law*

Roger Reynolds, B.A., Macalester College; J.D., New York University; *Lecturer in Law*

Barbara Reznar, B.A., Gettysburg College; J.D., Washington and Lee University; *Lecturer in Law*

Louis Ricciuti, B.A., Boston College; J.D., University of Connecticut; *Lecturer in Law*

Tracy Rich, B.A., Union College; J.D., New York University School of Law; L.L.M., Boston University School of Law; *Lecturer in Law*

James K. Robertson, Jr., B.A., Yale University; M.A., Hartford Seminary; J.D., University of Connecticut; *Lecturer in Law*

Morgan Paul Rueckert, B.A., J.D., University of Connecticut; *Lecturer in Law and Mark A. Weinstein Clinical Teaching Fellow*

Patrick J. Salve, B.S., J.D., University of Pennsylvania; *Adjunct Professor of Law*

Hon. Michael R. Sheldon, A.B., Princeton University; J.D., Yale University; *Adjunct Professor of Law*

Sandra L. Sherlock-White, B.A., Central Connecticut State College; J.D., Western New England College; *Lecturer in Law*

Jay. E. Sicklick, B.A., Colgate University; J.D., Boston College; *Adjunct Professor of Law*

Douglas Simpson, A.B., Dartmouth College; J.D., University of Connecticut; *Lecturer in Law*

Mark H. Taylor, B.A., Drew University; J.D., University of Connecticut; *Lecturer in Law*

James R. Turcotte, B.A., Boston University; J.D., Western New England College; *Lecturer in Law*

Stefan Underhill, B.A., University of Virginia; B.A., Oxford University; J.D., Yale University; *Lecturer in Law*

Arthur Webster, B.A., Bradley University; J.D., Georgetown University; *Adjunct Professor of Law*

William Wilcox, B.A., University of Connecticut; J.D., Suffolk University; *Lecturer in Law*

Michael Wilder, B.A., Yale University; LL.B., Harvard University; *Lecturer in Law*

Paul B. Zolan, B.A., Trinity College; J.D., University of Connecticut; *Lecturer in Law*

CONNECTICUT INSURANCE LAW JOURNAL

VOLUME 16 2009-2010 NUMBER 1

CONTENTS

ESSAYS

CREDIT DERIVATIVES ARE NOT “INSURANCE” *M. Todd Henderson* 1

THE ROAD FROM “TWIN PEAKS” –
AND THE WAY BACK *Michael W. Taylor* 61

ARTICLES

WHITHER THE DUTY OF GOOD FAITH
IN UK INSURANCE CONTRACTS? *John Lowry* 97

RISK DATA IN INSURANCE INTERPRETATION *Michelle Boardman* 157

THE LAW AND ECONOMICS OF FIRST-PARTY
INSURANCE BAD FAITH LIABILITY *Sharon Tennyson*
William J. Warfel 203

REGULATION OF LARGE FINANCIAL
INSTITUTIONS: LESSONS FROM
CORPORATE FINANCE THEORY *John P. Harding*
Stephen L. Ross 243

NOTES AND COMMENTARIES

PREDATORY LENDING AND ITS
INSURANCE CONSEQUENCES *Erin O’Leary* 261

THE 2008 MENTAL HEALTH PARITY AND ADDICTION EQUITY ACT: AN OVERVIEW OF THE NEW LEGISLATION AND WHY AN AMENDMENT SHOULD BE PASSED TO SPECIFICALLY DEFINE MENTAL ILLNESS AND SUBSTANCE USE DISORDERS	<i>Sara Nadim</i> 297
EXAMINING CURRENT PROPOSALS FOR INCREASING THE FEDERAL ROLE IN DEALING WITH COASTAL HURRICANE RISK	<i>Louis Cruz</i> 323

CREDIT DERIVATIVES ARE NOT “INSURANCE”

M. Todd Henderson*

This article explores whether credit derivatives should be regulated as insurance and offers an alternative form of regulation for these financial instruments. The largely unregulated credit derivatives market has been cited as a cause of the recent collapse of the housing market and resulting credit crunch. The article explores the possibility that the credit derivatives market should be regulated as insurance. It shows that the argument that some credit derivatives help banks and other providers of debt share risk with other investors is not sufficient for credit derivative contracts in general to be deemed “insurance.” It concludes that insurance regulation is not suitable for the credit derivatives market, while conceding that some sort of regulation may be necessary. The first section provides an overview of the basics of credit derivatives. The second section presents the argument for regulating credit derivatives as insurance. Section III describes why, although credit derivatives contracts can result in risk sharing or transfer, they are not within insurance law. The final section describes what one form of regulation of credit derivatives could look like and contrasts this with insurance regulation.

I. INTRODUCTION

The collapse of the housing bubble and the resulting credit crunch has caused untold harm to the economy and the lives of millions by destroying trillions of dollars in global wealth. The search for causes and remedies has begun in earnest, and chief among these is the largely unregulated credit derivatives market. Regulation of one form or another is the proposed solution in many quarters, and one of the prominent proposals is insurance regulation. At the very least, the analogy between credit derivatives and insurance is often made, and this faulty comparison may lead regulators astray, regardless of the mode of regulation ultimately

* University of Chicago Law School. Thanks to the George J. Phocas Fund for research support.

chosen. This Essay explores the suitability of insurance regulation to the credit derivatives market, concluding that it is a bad fit along many dimensions. Regulation of some sort may indeed be needed to remedy some fairly obvious market failures,¹ but insurance regulation and regulators have little if any role to play in any new regulatory regime.

The most basic form of credit derivative, known as a “credit default swap” (CDS), is simply a contract through which a lender can protect against the risk of default by paying premiums to a third party who agrees to make the lender whole in the event of default by the underlying borrower. The superficial similarity to typical insurance products, like property or life insurance, has caused some politicians and pundits to argue that credit derivatives are a form of insurance and should be regulated as such.² The former director of the Commodities Futures Trading Commission (CFTC), which regulates most derivative products, declared: “A credit default swap . . . is an insurance contract, but [the industry has] been very careful not to call it that because if it were insurance, it would be

¹ One market failure was the lack of a centralized clearinghouse to manage and reduce counterparty risk in credit derivative transactions. The Federal Reserve and Treasury are encouraging exchanges, like the Chicago Mercantile Exchange, to handle these transactions. See USA Exchanges: Geithner Pushes for Derivatives Shake-Up, FINREG21, July 11, 2009, <http://www.finreg21.com/news/usa-exchanges-geithner-pushes-derivatives-shake-up>.

² As shown in *supra* note 1, numerous politicians and observers have made the linkage. It has also crept casually into numerous media accounts. For example, in an account of the AIG catastrophe, an author for The New Republic calls credit derivatives insurance: “Between March, when Greenberg left AIG, and the end of 2005, Cassano's division issued more than \$40 billion in credit-default swaps (essentially insurance) for portfolios of securities backed by subprime mortgages. This was more than half of all the insurance of this type the company had on its books.” Noam Scheiber, *A New Theory of the AIG Catastrophe*, THE NEW REPUBLIC, Apr. 15, 2009, at 10, 11. Legal scholars believe this too. See Robert F. Schwartz, *Risk Distribution in the Capital Markets: Credit Default Swaps, Insurance and a Theory of Demarcation*, 12 FORDAM J. CORP. & FIN. L. 167, 181 (2007) (arguing that certain credit derivative contracts have “general form and function reflect[ing] many basic insurance arrangements.”); William K. Sjostrom, Jr., *The AIG Bailout*, 66 WASH. & LEE L. REV. (forthcoming 2009) (“A CDS certainly appears to fall within this definition given that the protection seller contractually agrees to compensate the protection buyer following the occurrence of a credit event.”), available at http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1346552.

regulated.”³ New York State went even further. On September 22, 2008, Governor David Patterson announced plans to regulate credit derivatives as insurance under the auspices of the state’s insurance department.⁴ New York State Insurance Commissioner Eric Dinallo then testified before a House Committee investigating credit derivatives: “the insurance regulator for New York is a relevant authority on credit default swaps,” because “[w]e believe . . . [they are] insurance.”⁵ Although New York has delayed its regulatory plans pending a federal review of credit derivative regulation,⁶ the question of whether credit derivatives are insurance remains an open and much bandied about one that needs to be analyzed.

³ *60 Minutes: Wall Street’s Shadow Market; Credit default swaps* (CBS television broadcast Oct. 5, 2008), available at http://www.cbsnews.com/stories/2008/10/05/60minutes/main4502454_page1.shtml. Dr. Greenberger argued that the industry was able to avoid regulation by simply using the word “swap” instead of “insurance” to describe the transaction. *See id.* (“So they use a magic substitute word called a ‘swap,’ which by federal law is deregulated.”). Swaps were specifically excluded from regulation by the CFTC by the Commodities Futures Modernization Act of 2000.

It is true that a typical CDS transaction does not involve a “swap” in the same way that an interest rate swap does. In the latter case, two parties agree to swap risks: the holder of a fixed-rate note agrees to swap that income stream with the holder of a variable-rate note. But while the term swap is not technically accurate, it is difficult to imagine that insurance or other regulators would be fooled by the label.

⁴ Press Release, Governor Patterson Announces Plan to Limit Harm to Markets from Credit Derivatives (Sept. 22, 2008) (on file with Errol Cockfield) available at <http://www.ins.state.ny.us/press/2008/p0809224.pdf>. *See also* Danny Hakim, *New York to Regulate Financial Tool Behind the Credit Crisis*, N.Y., NY TIMES, Sept. 23, 2008, at C10, available at <http://www.nytimes.com/2008/09/23/business/23swap.html?ref=business> (“The governor said the state’s insurance department would begin regulating credit-default swaps as insurance products in cases where the buyer of the swap also owns the underlying bond it is meant to back.”).

⁵ *Hearing to Review the Role of Credit Derivatives in the U.S. Economy: Hearing before the H. Comm. on Agriculture Committee*, 110th Cong. (2008) [hereinafter *Hearing*] (testimony of Eric Dinallo, Ins. Comm’r, N.Y. State), available at <http://agriculture.house.gov/testimony/110/h91120/Dinallo.pdf>.

⁶ *See id.* (“Based on the developments reported on by the President’s Working Group, it is clear they are committed to comprehensive and effective federal oversight of credit default swaps. . . . Accordingly, New York will delay indefinitely our plan to regulate part of this market.”). It is clear from Dinallo’s testimony that New York is using the threat of insurance regulation as a weapon to

This Essay argues that it makes little or no sense to regulate credit derivatives as or like “insurance,” regardless of whether they are used as insurance, that is, to reduce risk for one party. The instinct to call credit derivatives “insurance” is sensible enough – the lender buying credit protection looks much like an insured and the party selling credit protection looks much like an insurer, at least where the protection seller is in privity with holders of notes of indebtedness.⁷ The analogy is obvious: in a plain-vanilla CDS, the bank making an original loan pays a premium to a third party that in turn agrees to make the bank whole in the event of a future liability, that is, a default on the underlying loan or bond. This transaction resembles a typical insurance contract, where the insured pays a premium to a third party (an insurance company) in return for a promise to make the insured whole in the event of a loss.

But observing that something resembles or provides insurance against loss is not enough to warrant regulating it as “insurance.” Many contracts that are not called insurance or regulated as insurance imbed some component of insurance or risk sharing. For instance, when a farmer enters into a contract that allows the farmer to sell wheat at a fixed price in the future – a forward contract called a put option – the farmer is in effect insuring against a drop in the price of wheat. On the other side of this transaction, there may be a baker who enters into a forward contract that allows the baker to insure against an increase in the price of wheat. Both parties are buying price insurance from each other, likely with a middleman, known as a market maker, standing between and reducing the counterparty risk in the transaction. But these contracts, and all similar hedging contracts entered into by regular consumers and sophisticated financial entities, are not regulated as insurance contracts.⁸ The point can

encourage more comprehensive federal regulation. (“Based on the developments reported on by the President’s Working Group, it is clear they are committed to comprehensive and effective federal oversight of credit default swaps. . . . Accordingly, New York will delay indefinitely our plan to regulate part of this market.”).

⁷ New York State has proposed regulating only these credit derivatives – about one-fifth of the total market – because the argument that the parties are engaged in an insurance transaction is more difficult in cases where they are simply wagering on the default without an actual interest in the underlying debt instrument. *See id.*

⁸ Option contracts generally trade on exchanges, like the Chicago Mercantile Exchange, and as such are regulated by the Commodities Futures Trading Commission (CFTC).

be made more bluntly: it would be fanciful to argue that every contract in which a party could be said to be reducing its risk and another party was willing to take on some of that risk is or should be called insurance. If this were the case, state insurance regulators would be involved in regulating hedge funds, commodities, options, swaps, and countless other contracts entered into by consumers and firms. In fact, every contract assigns, shares, and apportions some sort of risk. No one seriously advocates this scope for insurance regulation. Simply providing some risk sharing is not enough to be regulated by state insurance commissioners.

The reason insurance regulation does not extend to every contract that involves some element of insuring risk has to do with the purpose of insurance regulation, as opposed to other types of regulation. There are broadly two justifications for a special law of insurance: first, the peculiar governance problems associated with insurance firms; and second, worries about unsophisticated consumers being duped by complicated and essential products. This Essay will show that neither of these justifications obtains or makes sense for the regulation of credit derivatives.

Governance problems arise because insurance companies have an inverted production cycle and do not generally have concentrated creditors like non-insurance firms. This means that two crucial constraints on the potential misuse of investment resources are missing: the feedback to the firm provided by product and other markets is missing given the fact that the insurance company produces its product (that is, payment of claims) many years after the consumers pay for it; and when things go badly for the insurance company, there is no concentrated interest to keep the firm from adopting an excessively risky strategy (from the perspective of creditors [that is, policy holders]).

Insurance law is designed to prevent the risk that insurers competing for policyholders, but unconstrained by normal forces, will charge too little for their products. This happens because of the continuous nature of insurance company inflows and outflows, coupled with a delinkage between the time of pricing a risk and the time of paying out the loss from the risk. In other words, insurance can look a bit like a Ponzi scheme, where new creditors of the firm are paying off the liabilities to old creditors. And, just as in a Ponzi scheme, when things go badly for the firm (that is, when actuarial estimates of liability turn out to be wrong), there is a natural tendency to offer new investors an attractive return to increase cash flows to pay for higher-than-estimated outflows.

The second part of the governance problem – the lack of concentrated creditors – exacerbates this problem, since there is no sophisticated entity with bargaining power that can keep the firm from

adopting a shareholder-friendly, go-to-Vegas strategy in the event liability estimates were erroneous. Without these governance constraints, initial misestimates and mistakes can fester and lead to large losses. This Essay shows how the counterparties in credit derivative contracts do not have this continuous investment problem or these governance problems, unless, of course, they are insurance companies, and how insurance regulation would be futile in any event.

Consumer problems arise because the consumers of insurance company products (and as such creditors of the insurance company) are average individuals without the expertise or sophisticated judgment to assess what they are buying in insurance products. The consumer-centric element of insurance regulation consists of three commonly recited justifications: to make sure insurers don't charge too much; to regulate the substance and terms of policies; and to regulate service and coverage issues. This basis for regulation is, to be sure, driven by a rather dim view of the philosophy of *caveat emptor*, the wisdom and skill of the average consumer, and the power of a small number of informed individuals to set market prices for others. This Essay does not take on the soundness of these consumer issues for insurance products, but simply compares their rationale with what is known about the participants in credit derivative contracts. Unlike the average consumer of insurance, the average participant in credit derivative markets is large, sophisticated, and capable of bearing losses. There is simply no basis for transferring the paternalistic impulses of insurance to this market.

This Essay shows that neither the governance problem nor the consumer abuse problem obtain in significant ways in the context of the credit derivatives market. Section I introduces the basics of credit derivatives. Section II presents the argument for regulating credit derivatives as insurance based on the rough analogy describe above. Section III then shows why the simple fact that credit derivatives sometimes result in risk sharing or transfer does not justify bringing these contracts and the parties to them within the ambit of insurance law. Section IV concludes by briefly sketching out what a sensible regulation of credit derivatives might look like, contrasting this with the approach of insurance regulation.

II. A PRIMER ON CREDIT DERIVATIVES

Credit derivatives exist in many forms and flavors, but the essence is simple: it may be more efficient for different entities to handle the various aspects of lending. A typical loan has many parts, including:

origination, servicing, monitoring, and funding or risk bearing. In a world without risk-sharing mechanisms, all of these are contained within one entity, that is, a bank. The bank has the relationships (origination), scale in the back office (servicing), experience (monitoring), and cash from depositors (funding). But the bank might not want to do all these things. It might want to become an arranger rather than a lender. One reason is because federal regulations designed to protect depositors require the bank to hold cash on hand to offset risk in loans.⁹

Another reason is that other potential lenders may be shut out of the corporate lending market, say because they do not have relationships with borrowers, but would provide a cheaper source of funding or be more efficient holders of particular aspects of corporate borrowing risk. Smaller regional banks and insurance companies come to mind here.

A final reason is that the bank may not be the most efficient monitor of firm conduct. The bank has experience with monitoring in general and (likely) with the specific borrower, but these advantages come with costs too. The relationships that led to the loan may corrupt the monitoring function—a sort of monitor capture by the borrower. Fees earned by banks for workouts and new loans may also distort incentives. So too might the fact that the workout group for a loan may be comprised of only a few individuals, who are subject to biases and shortcomings that a larger, market-based monitoring mechanism might be able to overcome. In short, there are many reasons why banks might prefer to decouple the bundle of loan features, but cannot without financial contracts that allow default risk to be shared. Credit derivatives, although much maligned as a result of current events, can help investors of various sorts allocate the lending market to its most efficient participants.

There are many variations of credit derivatives, but to answer the threshold question of whether credit derivatives in general can be considered “insurance”, it makes sense to consider the two most generic versions: the credit default swap (CDS) and the collateralized debt obligation (CDO).

A. CREDIT DEFAULT SWAPS

A CDS is a contract in which credit risk (that is, the expected losses arising from defaults) is transferred from one party to another. A

⁹ Capital reserves required by the Basle Accords are non-productive, and therefore reduce a bank’s return.

bank makes a loan to a borrower. There is some risk that the borrower will default on the loan, causing losses to the bank.¹⁰ Naturally, the bank wants to minimize these losses. The bank can do this through ex ante selection (that is, due diligence during the underwriting process), through ongoing monitoring of the borrower, and through effective ex post workout procedures. The bank can also contract with a third party to make the bank whole in the event the borrower defaults.

Consider a simple example: Bank holds on its balance sheet a \$100 note for a loan made to Borrower. Bank may want to shift some of the risk that Borrower will not repay the loan, say because of costly federal capital adequacy requirements that require Bank to hold some percentage of the loan's outstanding balance in cash reserves.¹¹ For a period of (normally) five years, Investor, who wants to hold risk of Borrower, agrees to make Bank whole in the event of default, thus assuming the risk of *default*,¹² in return for a stream of periodic payments from Bank. Voilà, the risk of the loan to Borrower has been swapped from Bank to Investor. The premium paid by the Bank is expressed as a risk spread in basis points, say 100 basis points or 1 percent. For a \$100 loan, this would mean the bank would make quarterly payments of \$0.25 to buy protection on the note. (The spread, which expresses the risk of default during the five-year term of protection, varies over time, allowing information about the quality of the debtor to be revealed and allowing investors unrelated to the loan contract to speculate on changing credit quality for profit.) In the event of "default,"¹³ Bank delivers the underlying credit instrument, in this case the loan, to Investor, and Investor makes a payment to Bank that puts Bank in the position it would have been in if Borrower had not defaulted.

¹⁰ The default risk is only one of many risks embedded in a loan. Lenders (and borrowers) face interest rate risks, volatility risk, currency risks, and so on. The significance of credit derivatives (CDSs) is the ability to unpack and isolate credit risk, and allowed it to be transferred to others who may be more efficient holders of it.

¹¹ In general, US banks are subjected to the capital adequacy requirements of the so-called Basle Accords, implemented by the Bank of International Settlements. See Basle Committee on Banking Supervision, International Convergence of Capital Measurement and Capital Standards, July 1998, available at <http://www.bis.org/publ/bcbs04a.htm>.

¹² Default risk is only one of many types of risk. Others include: interest rate risk, counterparty risk, currency risk, and so on.

¹³ As described below, see *infra* p. 7, the issue of when a credit derivative contract triggers payment is a complicated and tricky issue.

These contracts constitute the primary market for credit derivatives since the parties to the transaction interact with the underlying debt instrument: the lender writes and initially holds the instrument, while the counterparty ends up holding the instrument in the event of default, and, in any event, one of the two parties to the contract will hold the underlying debt at the termination of the CDS contract. The participants in this market are large commercial banks, as risk sellers, and insurance companies, hedge funds, pension funds, mutual funds, and a mix of investment banks, commercial banks, and smaller regional banks, as risk buyers.¹⁴ In this way, CDS contracts resemble other risk-sharing arrangements, like the syndication of credit or the sale of loans by banks. Most large loans are shared between a lead lender and other banks with which it contracts to share the risk of default, and there is a large and robust market for the sale of all or parts of loans to other banks. (Although risk sharing contracts, these are not considered or regulated as insurance.)

Returning to the example above, Bank will want to reduce the \$100 risk by getting other investors to participate, both to reduce its own risk and also to comply with capital adequacy rules. Bank could sell the loan, but this might mean giving up its relationship with Borrower, something neither party might want.¹⁵ Shifting the risk using a CDS preserves this relationship – in fact, Borrower may not even know the risk has been shifted – while also allowing conservative investors, like insurance companies and pension funds, to participate in the credit market.¹⁶ A small regional bank in Germany, an insurance company in Indonesia, and a pension fund in California are thus able to achieve desired risk-return investments in new ways.

¹⁴ See U.S. GEN. ACCOUNTING OFFICE, CREDIT DERIVATIVES: CONFIRMATION BACKLOGS INCREASED DEALERS’ OPERATIONAL RISK, BUT WERE SUCCESSFULLY ADDRESSED AFTER JOINT REGULATORY ACTION, GAO-07-716, p.6 n.8 (2007), available at <http://www.gao.gov/new.items/d07716.pdf> (“The top five end-users of credit derivatives are banks and broker-dealers (44 percent), hedge funds (32 percent), insurers (17 percent), pension funds (4 percent), and mutual funds (3 percent).”).

¹⁵ Borrower might not want Bank to sell the loan, since this may signal something bad about Borrower. The positive signal derived from having Bank be a creditor and monitor of Borrower may be quite valuable and for this reason, loan agreements often include no-sale clauses.

¹⁶ Conservatism here may derive internally, that is, from managers and shareholders, or from regulation.

The risk is not just swapped between two parties. In a typical CDS contract, the risk can be swapped many times, so that there are numerous parties in the risk-sharing chain. For instance, A, a manufacturing firm may borrow from B, a commercial bank. B, not wanting to hold the risk, may pay a premium to C, a dealer in CDS contracts, in return from a promise to be made whole if A defaults. (C, like all protection sellers, will be required to post some collateral with B to ensure payment on this obligation.) Since C is a dealer, it will look for a buyer willing to provide the ultimate risk-bearing function. D, an insurance company, agrees to make C whole in the event of a default by A in return for premium payments by C. Then E, a different commercial bank, wants exposure to A's credit risk, so it may agree to make D whole in the event of a default by A, in return for the payment of premiums by D. And on and on and on. There is no limit on the number of links in the risk-sharing chain, and, in practice, credit risk is often transferred dozens of times after its original creation. A typical credit derivative contract has hundreds of investors selling protection for hundreds of lenders and even more underlying borrowers. In the mortgage securitization market, for example, one of the problems in the collapse of US house prices was figuring out who actually held the risk of mortgage default so that workouts or foreclosures could happen efficiently.¹⁷

As discussed below, although C, D, E, and parties on down the chain could be said to be providing risk-sharing contracts in this example, it would be a dramatic expansion of the concept of insurance regulation to call them insurance companies. These entities might be individuals, banks, hedge funds, university endowments, or any other pool of investment money looking for return. In addition, the kind of insurance they are providing is not dissimilar from the insurance provided by nearly every contract that involves risk sharing (that is, every contract), and therefore raises irresolvable line drawing problems. The closest entity to an insurance company is, C the original CDS dealer. But, as discussed below, these are brokers who are regulated by numerous securities and banking laws, and subject to the oversight of numerous federal regulators.

CDS contracts do have characteristics similar to typical insurance contracts. Specifically, risk sharing and information asymmetries inevitably give rise to problems of moral hazard and adverse selection. Bank knows more about default risk of Borrower than the counterparties, and therefore

¹⁷ See, e.g., Mike McIntire, "Tracking Loans Through a Firm That Holds Millions," N.Y. TIMES, Apr. 23, 2009, available at <http://www.nytimes.com/2009/04/24/business/24mers.html>.

the latter may be unwilling to sell protection on the notes Bank brings to the market, since the counterparties may believe these are the debtors most likely to default. Another problem is that the existence of credit protection may make Bank less diligent in its monitoring role, thereby increasing the risk of default as a result of the risk-sharing contract. In theory and practice, there are steps that can be taken to mitigate these risks. Bank can hold back a portion of the risk of default, perhaps the first-loss position, thereby giving it incentives to monitor. This is analogous to a deductible in insurance contracts, and it can address both the moral hazard and adverse selection problem. (As it turns out, however, the nature of the securitization process made these first-loss tranches more valuable, on a risk-adjusted basis, than their price, while more senior tranches were less valuable.) These problems and the steps taken to mitigate them are discussed below.

There is also a rich secondary market in which the risk of default of a particular borrower (known as a “reference entity”) is traded among parties that have no contact with or affiliation with either the borrower or the lender. For instance, auto parts maker Delphi had \$2 billion in bonds outstanding at the time it declared bankruptcy, but there were over \$25 billion in credit derivative bets outstanding on whether or not Delphi would default on those bonds.¹⁸ The term “bet” is chosen deliberately, since these contracts are nothing more than wagers on whether Delphi would default. (As a side note, we do not regulate these bets as gambling for the same reason that the secondary market in stocks, that is, the New York Stock Exchange, is not regulated as gambling, even though it is. The reason is that the gambling is socially useful.) This large ratio of secondary to primary market is common across companies used as reference entities. After all, there is nothing (other than perhaps gambling law) that prevents two parties from writing a contract that replicates the payoffs from the payment or default of any debt instrument entered into anywhere. These contracts are called “synthetic,” since they do not involve any physical obligations to deliver on the underlying debt instrument.

The proposals to date to regulate credit derivatives have focused entirely on the primary market, specifically disclaiming any authority over the secondary market. As discussed below, this has something to do with what insurance experts call “insurable interest,” which is a requirement that the party allegedly doing the insuring has to pay only when the party that is

¹⁸ *The Ballooning Credit Derivatives Market: Easing Risk or Making It Worse?*, KNOWLEDGE@WHARTON, Nov. 2, 2005, <http://knowledge.wharton.upenn.edu/article.cfm?articleid=1303>.

allegedly insured actually suffers a harm unrelated to the insurance contract. But once regulators limit control over one part of the market, the fungibility of financial products will allow investors to move to other unregulated products that give them the same mix of risk and return. This is discussed further below.

B. COLLATERALIZED DEBT OBLIGATIONS

The other type of basic credit derivative is a collateralized debt obligation (CDO). A CDO is, at its core, the same as a CDS contract.¹⁹ As in a CDS contract, the parties to a CDO contract are shifting the risk of an underlying debt instrument from lender to investor, but instead of doing so for a corporate loan or bond issuance from a single borrower they do so for a series of loans or bonds from many borrowers. In this way, some portfolio theory-based diversification is achieved, since the risk for any investor of any one buyer defaulting is absorbed by gains on other debtors that do not default.

Here is a cartoon of how the basic, plain-vanilla CDO is formed. A CDO manager, usually an investor specializing in these products from a large investment house like Goldman Sachs, forms a special purpose vehicle (SPV), basically a stand-alone, bankruptcy-remote firm, and then chooses loans or bonds or mortgages from many borrowers to put into the SPV. The SPV then sells interests in the cash flows it will generate from these debt contracts to numerous investors. The SPV generates cash from the instruments it holds as the borrowers pay back the debts. This cash is then distributed to the investors according to the terms of their investment. So far, the SPV creating the CDO looks like any firm selling a service or product. The SPV raises money from investors, uses this money to invest in assets (in this case, debt instruments), manages the assets, and then distributes the profits it earns to the investors.

SPVs investing in credit derivatives have two somewhat unique features that enable them to be attractive risk-sharing mechanisms: tranching and securitization. The concepts are quite simple.

In a normal debt investment, a group of investors share a vertical slice of the expected payouts from a debtor. Three investors funding a \$100 loan to a firm each bear exactly the same risk if the borrower defaults – as the recovery on the loan falls from \$100 to, say, \$80, each investor suffers

¹⁹ In the nomenclature, if the underlying is a bond, the instrument is called a CDO, while if it is a loan, it is called a CLO.

a 20 percent loss. In a securitized debt investment, by contrast, the same three investors can slice the risk horizontally, instead of vertically, allowing them to assign different payouts, called “tranches,” based on the amount of recovery. For example, the investments could be structured such that the first investor bears the first \$10 in losses, the second the next \$10, and the third investor the remaining \$70. In that case, if the \$100 loan falls in value to \$80, the first two investors would suffer complete losses, while the third investor would suffer none (although its risk would increase, since any additional diminution in value would impair its position). This approach can dramatically reduce the probability of default for a particular tranche, and thus make even risky debt investment attractive for conservative investors. For example, the third investor reduces the riskiness of its investment by investing in the second type of vehicle. Insurance companies, which traditionally invest in only relatively safe instruments, used this approach to expand the types of investments they made.

The sharing and recategorization of risk can be enhanced by pooling together many risks through a process called securitization. This can be seen through a simple example. Consider two banks loaning to two companies in different and uncorrelated industries. The loans both pay \$100 in the good state of the world and \$0 in the bad state of the world, with a probability of default of 10 percent. If the banks take vertical positions, the expected value of the loans for each bank is \$90. Each faces the identical risks.

If instead, the cash flows from the two loans are pooled and tranced in a CDO, the 10 percent risk of default can be reduced for one of the investors. If one of the banks bears the first \$100 in losses and the second loses only if both borrowers default (and assuming the defaults are not correlated), the risk of default for the senior bank falls from 10 percent to 1 percent.²⁰ The expected value for the two banks is thus \$90 for the junior bank and \$99 for the senior bank. The process can be extended indefinitely, with each additional risk added to the pool further reducing the risk up the tranching scheme. For instance, adding a third investor and a third uncorrelated loan to the pool reduces the risk for the most senior bank, which suffers only if all three firms fail, to 0.1 percent.²¹

²⁰ The loss for the senior bank is the probability of firm one defaulting (10 percent) times the probability of firm two defaulting (10 percent).

²¹ Correlation of risks is obviously the key assumption in the creation of a CDO. If the risks in the three-bond case are perfectly correlated (that is, the failure of one firm means all three firms will fail), then the probability of loss for all three

Even when considering the role of tranching and securitization, the structure of a CDO is roughly the same as any firm in any business – they are nothing new under the sun. Tranching is just a fancy way of saying that the SPV replicates the priority of liquidation claims created by bankruptcy law and contracts for other firms. When a firm liquidates, equity interests lose first, and because of the absolute priority rule in bankruptcy, more junior interests, like unsecured creditors, lose only after equity interests are wiped out. This is true whether the firm is a donut maker or SPV holding debt instruments in a CDO structure. In addition, as discussed below, the shareholders investing in a traditional firm are selling insurance to the firm's debt holders, managers, and other stakeholders in the same way that the protection sellers are for the original bank in a credit derivative contract. Equity provides a downside cushion, since no payments are legally due equity holders, and thus provides risk sharing on favorable terms for holders of fixed claims on a firm's balance sheet.

Unlike regular firms, however, SPVs holding credit derivatives generally have only a single investment period. Whereas an insurance company is constantly adding new policy holders, SPVs are typically formed, buy debt instruments, raise money to fund the risk of these instruments, and then make payouts according to the terms of the credit derivative contracts. As mentioned above and discussed more fully below, this distinction is a crucial factor in the appropriateness of insurance regulation.

Before moving to the merits of the arguments for and against regulating credit derivatives as insurance, it is important to point out a few other features of credit derivative markets. First, the CDO market is at least two times larger than the CDS market. At the height of these markets in 2007, the single-name CDS market (that is, an entity selling protection to a bank for a loan to a single company) had a notional value of about \$20 trillion, while the total credit derivatives market was about \$60 trillion in notional value.²² This means that the CDO market, which makes up the rest of the market, was about \$40 trillion, or twice as big as the single-name CDS market.

Second, there are numerous index products and more complicated CDO products (such as the CDO²) that allow individual investors to buy

investors is the same 10 percent. Thus, there are no credit reduction benefits from securitization.

²² See Posting of Tyler Durden to, *Some More Facts about How CDS Market*, <http://zerohedge.blogspot.com/2009/02/> (Feb. 1, 2009, 10:31 EST).

exposure to a portfolio of credit derivative investments. For instance, in the single-name CDS world, there are several indexes, like the Dow Jones “Investment Grade CDX” and the “High Yield CDX,” that consist of over 100 borrowing firms of different credit quality. Investors can buy securities that track these indices in the same way they can invest in the S&P 500 or Wilshire 5000 equity indices. (As discussed below, investments in these indices are no more providing insurance to the underlying participants in the borrowing transaction than a regular firm selling equity, since both provide mechanisms of risk sharing.) Moreover, firms or investors seeking exposure to these credit default risks often hold a portfolio of risks or an index product for a few months or less, rolling the investments on a fairly constant basis to meet the investors’ or firms’ balance sheet needs.²³ As such, the investors in credit derivative indexes are not generally exposed to the possibility of actually having to pay for any losses on the original debt, but rather are susceptible to the change in the price of the indexed securities depending on the changing nature of the credit quality of the underlying borrower.

A CDO² (and more exotic credit derivative products) basically achieves the same result for portfolio products. A CDO² is simply a two-pool portfolio of tranching and securitized loans in which investors face exposure to both pools. If the credit risks are not perfectly correlated, this structure allows investors to lower their overall risk to something less than they would have from investing in both pools separately. These products also allow investors to invest in funds of CDO products, in which the exposure becomes more and more attenuated from any individual underlying borrower and starts to look more like generic risk exposure to the debt markets or even the market in general. As the case of CDS indexes, the investors in these products can now be thought of as merely identifying a unique risk-return investment as opposed to making bets about the credit quality of individual borrowers or pools of borrowers.

III. THE ARGUMENT FOR REGULATING CREDIT DERIVATIVES AS INSURANCE

At first blush, the similarity between property (or other) insurance and credit derivative contracts makes the call for insurance regulation of the latter seem reasonable. But a better case than that must be made,

²³ Interview with executive at insurance company responsible for credit derivative transactions, on Mar. 21, 2009.

because deeming something “insurance” brings with it a special set of laws and rules administered by state insurance departments. This would mean significant increases in the cost of using credit derivatives, and, on the margin, less use of them. This might be fine, but we need to believe that insurance regulation brings us something worth the price. The benefits of the regulation, therefore, must be well calibrated to the particular risks involved, lest otherwise socially beneficial transactions be deterred. This Part fleshes out the analogy between insurance and credit derivatives, while the next section shows how this analysis is highly misleading by looking behind the analogy to the purposes and justifications for calling something “insurance.”

The standard definition of “insurance” is an “agreement in which one party (the insurer), in exchange for consideration provide by the other party (the insured), assumes the other party’s risk and distributes it across a group of similarly situated persons, each of whose risk has been assumed in a similar transaction.”²⁴ There are two parts of this definition – (1) risk transfer; and (2) risk pooling. The insurer assumes not only the risk of loss, but distributes the risk across many other similarly situated individuals or entities, so as to reduce unpredictable events into a more predictable cash flow stream. In technical jargon, insurance companies try to pool risk by attracting a sufficiently large number of diverse policyholders such that the law of large numbers will reduce the aggregate variance of claims.

The credit derivative contracts discussed above have many characteristics that seem to fit well within the scope of at least the first part of this definition. As in a typical insurance contract, a CDS contract involves a party with an asset (the loan)²⁵ with a risk of loss (default by the borrower), paying a reoccurring premium to a counterparty, which in turn agrees to make the first party whole in the event there is a loss. To analogize, just as a homeowner that pays a monthly premium to an insurance carrier in return for a promise to make the homeowner whole in the event of a loss related to the home, so too does the lender pay a monthly

²⁴ ROBERT H. JERRY, II & DOUGLAS R. RICHMOND, UNDERSTANDING INSURANCE LAW § 10(d) (4th ed. 2007). *But see* THE NEW OXFORD AMERICAN DICTIONARY 881 (Elizabeth J. Jewell & Frank Abate eds., 2001) (defines “insurance” is as “a practice or arrangement by which a company . . . provides a guarantee of compensation for specified loss . . . in return for payment of a premium.”). This definition misses a key component of insurance – the pooling of risk.

²⁵ The underlying credit instrument need not be a loan, but could be any debt instrument, such as a mortgage, bond, note or any other form of indebtedness.

premium to a third party in return for a promise to make the bank whole in the event of a loss related to the loan. If this analogy holds, the lender is the “insured” and the counterparty is the “insurer.” And, the insurance law of the fifty states would then regulate the insurer and the content of contracts it enters into with insureds.

Building on the apparent similarity between typical insurance contracts (regulated by state insurance agencies) and credit derivatives, New York State recently proposed deeming credit derivatives “insurance.” The chief state regulator, Eric Dinallo, offered the rationale during testimony before a congressional committee: “With [plain-vanilla CDS contracts], if the issuer of a bond defaults, then the owner of the bond has suffered a loss and the [CDS] provides some recovery for that loss.”²⁶ Dinallo limited the reach of the proposed regulations of credit derivatives, however, by asserting the state’s jurisdiction covers only cases where the credit derivative contract is between an original lender and a third-party investor, that is, single-name CDS contracts in which an individual or entity sells protection to an originating bank.²⁷ These are so-called “covered” transactions (as opposed to “naked” ones), since there is privity between the insured and the underlying debt instrument. The reason for this limited scope for insurance regulation is based on a generally accepted argument that the party being insured has an “insurable interest” in the underlying amount at risk under the contract. In other words, a contract is “insurance” only if the insuring party pays when the insured party actually suffers a harm unrelated to the insurance contract.

This concept can be illustrated by comparing the primary and secondary credit derivative markets. Where a bank issues a loan and then buys credit protection on that loan that pays off if the loan defaults, the argument is that the buyer of credit protection has an insurable interest in the loan, and that the protection acts as insurance against this loss. In the secondary market, by contrast, two parties unrelated to the issuance of the loan (and without the knowledge of the bank making the loan, the borrower

²⁶ *Hearing, supra* note 5, at 3 (testimony of Eric Dinallo, Ins. Comm’r, N.Y. State).

²⁷ *Id.* (“We believe that the first type of swap, let’s call it the covered swap, is insurance. The essence of an insurance contract is that the buyer has to have a material interest in the asset or obligation that is the subject of the contract. That means the buyer owns property or a security and can suffer a loss from damage to or the loss of value of that property. With insurance, the buyer only has a claim after actually suffering a loss.”).

taking out the loan, and any parties contracting with either of them), bet on whether or not the borrower will repay the loan. According to the conventional wisdom of insurance regulation, this bet is not insurance. The reason for this concession is that otherwise “insurance” would include far too much and things far beyond the ken, expertise, or conceivable reach of state insurance regulators.

The concession is manifest in the findings of insurance regulators. For instance, in 2000 after Congress exempted “swaps” and other derivatives from certain regulation,²⁸ the New York Insurance Department was asked whether credit derivatives were in fact insurance, which would be subject to state regulation. The question it was asked by federal regulators was: “Does a credit default swap transaction, wherein the seller will make payment to the buyer upon the happening of a negative credit event and such payment is not dependent upon the buyer having suffered a loss, constitute a contract of insurance under the insurance law?”²⁹ This question is aimed at the secondary market, and was answered in the negative for reasons of a lack of privity with the loss on the part of the entities engaging in the derivative transaction. In his testimony before the House, Commissioner Dinallo distinguished this prior finding of the New York State Insurance Commission that credit derivatives were not insurance, by pointing out that the question asked was focused only on non-privity cases or “naked” credit derivatives.³⁰ From this, Dinallo concluded that a different result could obtain in the privity case (that is, CDS contracts), since the protection seller was insuring a real loss outside of the context of the contract. The analogy described above was thus sufficient for him to conclude that, with privity and a real potential loss, credit derivatives of the plain-vanilla CDS variety are insurance products.

The argument is not preposterous on its face. Insurance is about risk sharing, and in that sense credit derivatives, which are fundamentally risk-sharing contracts, are akin to insurance. But, as shown below, the fact that credit derivative contracts are providing an insuring or risk-hedging

²⁸ See Commodity Futures Modernization Act of 2000 or [CFMA], H.R. 5660, 106th Cong. §§ 103, 105, 407 (2000); CFMA of 2000, S. 3283, 106th Cong. §§ 103, 105, 407 (2000); see also Gramm-Leach-Bliley Act [GLBA] (Financial Services Modernization Act of 1999), Pub. L. No. 106-102, 113 Stat. 1338, 1393-94 (1999).

²⁹ *Hearing, supra* note 5, at 4-5 (testimony of Eric Dinallo, Ins. Comm’r, N.Y. State).

³⁰ *Id.* (“So at the same time, in 2000, the New York Insurance Department was asked a very carefully crafted question.”).

function does not mean that it is sensible or efficient to regulate them as insurance. There are lots of contracts in which one party is effectively offering insurance as part of the deal, and yet none of these are regulated by state insurance commissioners. The reason is the underlying policy justifications for having a separate body of insurance law do not reach these contracts, and thus applying that law would raise the costs of the contracts without any likely benefit. Another reason is that the second part of the definition of insurance – risk pooling – is absent in many of these transactions, as it is in most CDS contracts. In fact, if there is any risk pooling by protection sellers (the alleged insurers), it occurs in CDO contracts or secondary-market CDSs, exactly the place where New York claims its regulatory reach does not extend. These arguments are made in the next Part.

A few other features shared by insurance and credit derivatives provide some support for the analogy to insurance. The first similarity between insurance and credit derivatives is the incomplete nature of the risk transfer. The insured (either the homeowner or the lender) swaps the risk of loss with respect to the underlying asset (either the home or the loan) for the risk that the insurer will not be able to make the insured whole. This latter risk is called “counterparty risk,” and it is a central justification for insurance regulation. An individual who takes out an automobile insurance policy is swapping the risk of loss from an auto accident for the risk of loss that the insurance company will not be around to pay the claim. Capital adequacy rules, investment restrictions, and other aspects of insurance regulation exist to decrease this counterparty risk. Although CDSs and other credit derivatives share this similar feature, as discussed below, this alone does not justify regulating them as insurance as there are many other ways of reducing the counterparty risk problem that do not involve the full panoply of insurance regulations.

The second similarity between insurance and credit derivatives is the presence of moral hazard. Whenever risk is transferred, there is the possibility of misbehavior on the part of the transferor or the transferee. If the transferor (that is, the bank) has an obligation to prevent the loss from occurring, say by monitoring the conduct of the borrower, the transfer of risk reduces the incentive to do this on the margin. In addition to shirking, protection buyers may act deliberately to force the debtor into bankruptcy, say by withholding lending that would otherwise be efficient or by invoking covenants outside the normal usage in the industry. These examples of destroying value to simply collect on a CDS contract can obtain in both the primary and secondary markets – nothing prevents the holder of synthetic protection, say a hedge fund, from taking steps to harm

the borrower in order to collect on its bet. Although plausible, this facial similarity does not justify treating credit derivatives as insurance. As discussed below, this argument proves too much. Numerous opportunities exist for similar moral hazard problems outside of the insurance context, there already exist mechanisms (both market and from industry trade groups) to ameliorate any moral hazard, and there is nothing about insurance law that makes it a good fit for further reducing these potential harms, if they are substantial.

An additional argument for regulating credit derivatives as insurance is the absence of any existing regulation by other federal or state agencies, especially of certain players in the market, like hedge funds and other private pools of money. Many experts and pundits blame the lack of regulation of the credit derivatives market as contributing to the credit crisis. The argument goes like this: credit derivatives are not traded on an exchange, but rather through individualized contracts, known as the over-the-counter market, and the lack of regulation, either directly or indirectly through regulation of the exchanges on which securities trade, allowed private parties to externalize systemic risk costs onto society. The lack of regulation thus generated an inefficient number or type of these transactions from a social welfare standpoint.

There may be something to the premise of this argument, that is, that the lack of regulation exacerbated the risk that private parties would act in ways that would be privately optimal but increase the risk of a global financial meltdown. The premise is debatable, but even if it is true, this Essay shows that insurance regulation is not the only way in which these systemic costs can be internalized by firms. Most obviously, direct regulation of the credit derivatives market by existing federal departments responsible for derivatives and markets, such as requiring derivatives to be traded on an exchange, is possible under current law.

In fact, it seems from the public statements of New York officials that the purpose of the characterization of credit derivatives as insurance is intended to stoke federal regulators to act, more than a firm belief that credit derivatives are insurance. After all, if they are insurance, then there should in fact be no need for or call for federal regulation. In testimony before Congress and other public comments, New York State's insurance officials "stopped short of endorsing comprehensive state-level regulation of this privately negotiated market" and agreed to delay its plan to regulate credit derivatives based on the indication that federal regulators are "committed to comprehensive and effective federal oversight of credit

default swaps.”³¹ The fact that Commissioner Dinallo’s testimony outlined a proposed regulatory agenda for federal agencies also supports the threat-of-regulation-as-leverage claim.³²

A final argument for state-based insurance regulation is the fact that numerous insurance companies were involved in the credit derivatives markets as buyers and sellers of protection, as well as acting as brokers and speculators in secondary markets. According to one estimate, insurance companies represented about 20 percent of end users of credit derivatives.³³ For instance, insurance giant AIG, invested heavily in credit derivatives of various kinds – its portfolio of CDSs reached \$526 billion at its height.³⁴ And it is widely viewed that the losses on these credit derivatives – over \$30 billion in 2007 and 2008 alone – were the cause of the failure of AIG and the need for the massive government bailout.³⁵ The logic of regulation would thus be that these products were misused by insurance companies, among others, and this justifies regulating them as insurance. President Obama seemed to endorse this view when he described the situation as follows: “You’ve got a company, AIG, which used to be just a regular old insurance company. ... Then they decided--some smart person decided--

³¹ *Id.* at 7.

³² Dinallo testified that :

Effective regulation of credit default swaps should include the following provisions: All sellers must maintain adequate capital and post sufficient trading margins to minimize counterparty risk; A guaranty fund should be created that ensures that a failure of one seller will not create a cascade of failures in the market; There must be clear and inclusive dispute resolution mechanisms; To ensure transparency and permit monitoring, comprehensive market data should be collected and made available to regulatory authorities; The market must have comprehensive regulatory oversight, and regulation cannot be voluntary. *Id.*

³³ U.S. GEN. ACCOUNTING OFFICE, *supra* note 14, at 6 n.8 (“The top five end-users of credit derivatives are banks and broker-dealers (44 percent), hedge funds (32 percent), insurers (17 percent), pension funds (4 percent), and mutual funds (3 percent).”).

³⁴ Sjostrom, *supra* note 2, at 40 (“A CDS certainly appears to fall within this definition given that the protection seller contractually agrees to compensate the protection buyer following the occurrence of a credit event.”).

³⁵ *Id.* at 26.

let's put a hedge fund on top of the insurance company and let's sell these derivative products to banks all around the world."³⁶

As discussed below, this argument proves too much, as many other entities, like banks, hedge funds, pension funds, and so on, used credit derivatives too, often disastrously but just as often fantastically, and therefore there is nothing special about the end users that justifies treating them as insurance. If anything, the fact that some insurance companies were harmed by them justifies different regulation on insurance companies.

IV. THE ARGUMENT AGAINST REGULATING CREDIT DERIVATIVES AS INSURANCE

This Part presents several conceptual and practical reasons why credit derivatives should not be regulated as insurance products or why sellers of credit protection should not be regulated as insurance companies. Notwithstanding the superficial appeal of the analogy between insurance contracts and credit derivative contracts, the policy justifications for special rules regulating insurance carriers and contracts do not obtain in the credit derivative context. Examining the rationale for insurance law and the important differences with credit derivatives will show this.

The two primary reasons for having a separate body of insurance law are the particular governance problems associated with insurance companies and the fact that insureds are typically unsophisticated individuals for whom insurance is essential and may be difficult to obtain in the event of certain individual characteristics. (These two justifications correspond with the two major features of insurance law – the regulation of insurance company investments and the regulation of sales to individuals.) Neither of these reasons justifies applying insurance law to credit derivatives.

Before addressing these reasons for insurance law, this Part addresses the reasons why policy arguments are needed in the first place. The primary reason insurance contracts are treated differently than other contracts (and “insurance law” is a separate body of law) is not because of their nature as “insurance” but rather because they are issued by insurance companies. This conclusion is evident from numerous problems that would

³⁶ See “President Barack Obama on ‘The Tonight Show with Jay Leno,’” N.Y. TIMES, Mar. 19, 2009, *available at* <http://www.nytimes.com/2009/03/20/us/politics/20obama.text.html>.

arise in extending the reach of insurance regulation to all contracts providing some kind of insurance.

A. PROVIDING INSURANCE IS NOT ENOUGH

1. Line-Drawing Problems

It could not be enough for a contract that insures against risk to be regulated as insurance or to bring the seller of that insurance within the ambit of (state-based) insurance regulation. *Every* contract has a degree of insurance embedded in it, and options and derivatives of all sorts, which are not considered insurance or regulated as such, are mostly about insuring or hedging against risk. If insurance law covered all contracts that are partially or completely about insurance, the line drawing problems about what is insurance would likely broaden the scope of insurance law to cover vastly more than it currently does. The result would be to add regulatory costs and uncertainty to a vast swath of the economy, with little or no expected benefit.

Consider a simple options contract, known as a forward contract: Farmer agrees to sell wheat in six months at a given price (a put option) and Baker agrees to buy wheat in six months at a given price (a call option). Both Farmer and Baker are purchasing price insurance from each other – insurance against a price rise (for Baker) and against a price drop (for Farmer) – by locking in a set price in advance. The contract is not regulated as insurance, and neither Farmer nor Baker are currently regulated as insurance companies, even though each may be providing insurance to the other. There are several reasons for this result. Both parties are presumably somewhat sophisticated, since they went in search of derivative hedging tools, or are intermediated by market professionals. These gatekeepers compete in competitive markets and are regulated by other laws and exchange-based rules to ensure fidelity to their clients’ interests and a suitability between client needs and products sold to them. In addition, the derivative contracts are likely made either on a competitive derivatives exchange or as the result of arms’-length negotiation.³⁷

³⁷ Nor is the result different if a third party sells the insurance to Farmer or Baker. For instance, an individual unconnected with the farming or baking business may believe that wheat prices will rise/fall in six months based on predictions about weather, changes in supply or demand, or other factors. This individual can enter into a forward contract with Farmer or Baker either directly or

There are innumerable contracts that provide the same type of insurance as a forward contract does and that are not regulated as insurance. Any hedging contract has an element of insurance. For instance, an investor who is long a particular security, commodity, or investment might want to reduce the risk of the position by entering into another investment or contract with a third party that moves the other way. The third party might be thought of as providing some insurance for the investor, but this is not how the contract is thought of or regulated. As discussed below, contracts like these have many purposes, and trying to sort contracts along this dimension is likely to be costly and highly imperfect, especially if done in *ex post* litigation, where litigant opportunism and hindsight bias will be problematic, or by regulators, who will face inevitable public choice problems in their definitional exercises.

One can imagine trying to sort between these two categories of contracts by inquiring into the minds of the contracting parties to see whether the contract was about insurance or something else. As noted above, this would require the mind of the parties to every contract to be examined to determine whether they are providing or seeking “insurance.” Regulators would need to know whether the investor was entering into the contract for insurance or hedging purposes. This is not generally the inquiry regulators make, perhaps because the question of knowledge is malleable and costly to enforce, especially given imperfect courts and a costly litigation system. Another dividing line could be the intent of the investor, but this too is an unhelpful and costly line to draw. It may be significantly over-inclusive, and it is susceptible to similar proof problems as knowledge. There may be mixed motives for all investments – return, hedging, speculation, and so on – that will be difficult to unpack accurately and without being subject to *ex post* bias, power grabs by regulators, and rent seeking by stakeholders of the firms in question.

Rather, some contracts tend to be primarily about insurance, while others have multiple functions, some of which might be about risk sharing. The former might fall within the ambit of insurance regulation, while the latter never do. But where is this line? Consider, for instance, equity investments in run-of-the-mill firms. As discussed above, any equity investment in any firm could be thought of as insurance in the same way

through an options exchange, such as the Chicago Mercantile Exchange. These exchanges are populated with relatively sophisticated parties and are covered by alternative regulatory regimes, including licensing requirements for brokers and dealers.

that credit derivatives are. After all, when a firm issues equity interests to investors, it is entering into a risk-sharing contract (on behalf of managers, creditors, and other stakeholders) with these investors. Equity holders, unlike debt holders, have no fixed claim on a firm’s assets, and therefore provide a source of funding that is less sensitive to downturns in performance than debt. A firm that has more equity on its balance sheet is, all else being equal, less risky than a firm that has less. So we could re-characterize a firm’s decision to issue equity (to lower its debt to equity ratio) as buying insurance (against a downturn in the firm’s affairs) and the investors buying the equity as selling insurance to the firm. Of course, no one thinks of equity in this way. But equity is as much about insurance as credit derivatives are.

One reason securities are not regulated as insurance is the fact that equity investments are regulated by a separate body of law – securities law – specifically designed to address the policy challenges of issuing and investing in securities. When Congress passed the securities laws in the 1930s, it could have simply called equity investments insurance and delegated regulation to state insurance law under the same theories as those calling for this treatment of credit derivatives. But this would have been a reach – although arguably insurance, equity securities are sufficiently different along numerous dimensions to justify a separate body of regulatory law.

Another reason equity might not be regulated as insurance is because of the particular characteristics of the contracts in question. As noted above, typical credit derivative contracts look like typical insurance contracts: one party makes periodic payments to another in return for a make-whole promise in the event of a future occurrence. This similarity is only a superficial one, however, since there are many other aspects of credit derivative contracts that are quite different. For instance, payments may not turn on actual losses, there may be no pooling of risk, the make-whole promise may be purely synthetic, and so on. In addition, it is hard to imagine regulatory treatment turning solely on the question of whether risk-sharing payments are made on a periodic basis (as in insurance contracts) or a lump sum basis (as in equity investments, forward contracts, and so on). This would elevate form over substance in an arbitrary way not anticipated by the parties, and would provide an easy roadmap to avoiding any regulation.

Finally, few if any investors making an equity investment think they are providing insurance. Rather, the investment is made for a whole host of reasons, including pure investment, speculation, hedging, and so

forth. Credit derivatives are used for these multiple reasons too, and this cuts strongly against trying to narrowly pigeonhole them as insurance.

2. Credit Derivatives Are Not Just About Reducing Risk

Another problem with regulating credit derivatives as insurance is that they are not just (or, even, primarily) used for “insurance” purposes. As noted above, a common use, but only about one-fifth of the current market, is the buying and selling of credit protection on loans, bonds, and other sources of indebtedness. It is doubtful that this use fits squarely within the regulatory definition of insurance or that insurance regulation would be beneficial to these markets, but importantly for definitional purposes swapping credit risk is only one of the many uses of these financial products.

Credit derivatives can be used to hedge risks independent of and unrelated to the original loan or bond being used as a reference entity. For example, a hedge fund that wants to reduce its exposure on, say, Russian wheat futures, may find a corporate bond risk whose risk offsets its commodity position favorably, and thus enter into a contract with a third party, who might be hedging Texas oil prices, who is willing to pay in the event the bond defaults. This transaction has nothing to do with the underlying bonds, since it only uses them as a reference for calculating a stream of payments. The transaction is akin to two individuals in China betting on whether I will crash my car. Neither of them is insuring me, but rather they are simply using the probability of me crashing my car (and the amount of damage that will result) as a reference for assigning risk among them. (The original debtor in credit derivative contracts is called a “reference entity,” a description that well captures this concept.) These bets are not considered insurance, because there is no privity with party suffering a loss (that is, the lenders in the case of a default by the borrower on the bond or me in the case I crash my car) and furthermore no proof required that an actual loss be suffered. Even if these bets were considered insurance, it would be impossible to regulate all of them in this way. Detecting them would be difficult and costly, and, even if possible, would simply direct the parties into transactions of similar risk-return combinations but other designs. In other words, if regulators deem one class of credit derivatives “insurance,” and thereby impose increased regulatory costs on that class, and deem another class of credit derivative contracts as “not insurance,” parties will naturally structure their transactions as the not-insurance kind. More on this later.

Credit derivatives are also used for speculation about credit risk. Investors can use credit derivative contracts to speculate on the default probability of a particular borrower. In general, there are no insurance contracts like this. Participants in insurance markets do not wager on whether an individual's probability of dying or crashing a car is rising or falling on a daily basis, as they do with firm risk in credit derivative markets. Insurance is based on probabilities at the macro level and across large numbers based on actuarial science, and, when it is based on individuals characteristics, it is done only at the point of origination and not for speculation purposes. But this is exactly what is done with credit derivative contracts. For example, an investor who believes that General Electric's credit quality is likely to worsen over the next few years can buy protection against default by GE on its debt. If the credit quality does deteriorate, the cost of protection will rise, and the investor will earn a profit. Similarly, if an investor believes that GE's credit quality is likely to improve, it can sell protection against default by GE. If the credit quality does improve, the cost of protection will fall, and the investor will earn a profit. Before credit derivatives this kind of speculation was extremely difficult, as it is practically impossible to short bonds or loans. The credit derivative market thus allows for information about debt quality to be processed in a market, perhaps with large gains to capital allocation efficiency. Like the hedging examples above, these transactions are not insurance in any meaningful sense. Nothing about the speculation contract requires that it be held for any period of time. An investor can buy or sell protection and hold it for an hour, a day, a year, or five years (the typical maximum length), depending on the profit that can be made from buying or selling at a particular time. The contract does nothing more than offer an opportunity to buy or sell later at a higher or lower price. In this way, credit derivatives can be, and are largely, about investment, not insurance.³⁸ In fact, they resemble secondary market transactions in equities, since they involve market-based trades about the fundamental value of a third party unrelated to the transaction in question.

A final (non-insurance) use of credit derivatives is arbitrage, of either the pure or regulatory variety. Pure arbitrage possibilities arise when there is temporary mispricing in markets that allows investors to engage in

³⁸ Life insurance could be used for investment in a way, but this is not its primary purpose or the way it is typically used. Moreover, state insurance regulators are not really concerned about regulating investment decisions by sophisticated parties.

paired transactions that earn sure profits, and thus eliminate the mispricing. Mispricing in debt securities might arise because the primary and secondary markets are traded on different markets: a firm's bonds might be trading on a bond market, while protection on those bonds might be trading on over-the-counter markets that are not perfectly correlated with the public bond markets. This difference might generate opportunities for an investor to buy/sell the underlying security in one market, while simultaneously buying/selling protection on it in the over-the-counter market using a credit derivative in a way that allows it to earn a return that is independent of the credit risk of the borrower. This kind of arbitrage opportunity has no insurance-like characteristics, and it is potentially quite useful in accurately pricing credit risk by removing temporary market inefficiencies. Participants in credit derivative markets, at hedge funds, insurance companies, and other large financial entities, describe this as a major driver of their transactions in credit derivative markets.³⁹

The other type of arbitrage – regulatory arbitrage – is related to insurance, since it is one of the main reasons insurance companies are involved in the credit derivatives markets. But the lesson here runs counter to any regulatory story – in fact, it is a product of regulation itself.

Regulatory arbitrage works like this. Banks are often the most efficient originators of loans, since they have relationships with lenders and the back office to underwrite and process loans. However, they are not always the most efficient holders of loans because of regulations that make holding risk more costly for them and regulations that make investing in credit risk difficult for insurance companies and other risk-sensitive investors. In other words, smaller banks, individual investors, insurance companies, pension funds, university endowments, foreign governments, and a whole host of other investors would like to participate in corporate debt markets, but cannot do so in the absence of financial instruments that allow large commercial banks to sell the risk, especially in ways, like securitization discussed above, that recharacterize the risk in ways that make individual investments in it appear less risky. If insurance companies, pension funds, or endowments can only invest in corporate debt rated AAA, banks, who are required to hold cash reserves on corporate debt rated below AAA will find a way of repackaging the debt so that some of the sub-AAA debt can become AAA debt – this is the securitization and tranching process described above.

³⁹ Interview with executive at insurance company, *supra* note 23.

The first set of regulations – capital adequacy rules in the Basle banking accords – require banks to hold a certain amount of cash in reserve for each dollar lent out. In short, banks have to hold cash, while other investors do not, which makes them inefficient holders of risk. The reason for this rule is because banks have average individuals as creditors through deposit taking, and given the moral hazard problems created by federal deposit insurance, banks would otherwise engage in socially inefficient risk taking. This cash, historically about 8 percent of the total value at risk, is not productive from the standpoint of the bank’s investors, so it would be more efficient for them to loan out the money, earn the fees on the origination, and then offload the risk, in whole or part, to other investors so that more of the bank’s cash can be put to use for its shareholders.

The reason insurance companies were involved heavily in these markets (primarily as protection sellers to banks that had originated loans), is because state law insurance regulations limit the kinds of investments that insurance companies can make. For instance, insurance companies are often restricted to investing in credits rated AAA by credit rating agencies. These credit rating agencies were in turn paid by the managers of credit derivative SPVs to rate the risks of investments in those SPVs, often to get a slice of them to be rated AAA to attract the monies held by insurance companies. As such, insurance companies became one of the largest investors in credit derivatives. For example, AIG (through its financial products business) invested nearly \$400 billion in providing various European financial institutions with “regulatory capital relief” through credit derivatives.

Credit derivatives help complete these markets by allowing the bank to offload the risk to investors who can more efficiently bear it, while still having the ability to earn fees from origination. A bank that makes a loan with a customer can now package the credit risk of that loan in a new entity, which then uses securitization to create risk slices that will be attractive to new classes of investors, and then sells off interests in the new entity. To be sure, the original bank could be thought of as buying insurance, since it is offloading to or sharing risk with others. But that description of the activity is a cartoon representation of the transaction. The bank is really engaging in regulatory arbitrage, but this redefinition is only superficial. The important point is that looking at what the bank is doing is only part of the story about whether regulation makes sense.

In addition, as discussed above, those buying interests in the bank’s credit risk are no different than investors in any firm. A CDO is just a business plan in which the proceeds from hundreds of credit risks from various lenders are pooled together to generate a series of cash flows. The

firm (or SPV) holding the interests in these future cash flows is not conceptually different from a firm that sells anything else, be it iPods, consulting services, or what have you, and then sells interests in the cash flows these sales generate. When a firm raises money from shareholders, it is buying insurance in the same way that a bank transfers some risk through a credit derivative contract. Although equity holders are not liable to make the seller of the risk whole in the event of some specific default, the equity investors are providing the firm with an opportunity to reduce its risk. Equity, like insurance, provides a cushion against a downturn. Of course, no one thinks of regulating securities as insurance despite the similarity along this dimension.

There are at least two important differences between these two types of risk-sharing mechanisms. First, on average the sellers of equity to regular firms are much less sophisticated than the sellers of risk protection to lenders. This, of course, cuts the other way from regulating credit derivatives as insurance. Second, the structure of the standard insurance contract (and the typical credit derivative contract) is different than the shareholder contract. Whereas in a credit derivative or insurance contract the party assuming the risk receives periodic payments in return for a promise to make the party selling the risk whole, in the shareholder contract, the sequence of payments is reversed: the party assuming the risk of default pays the money up front, while agreeing to receive future cash flows in the form of dividends, capital appreciation, or liquidation value at some time in the future. This alternative structure has important implications, which are discussed below, but it does not necessarily undermine the attempted analogy to insurance. After all, if insurance is defined as a contract in which risk is moved from one party to the other, the structure and terms of the contract are, all else being equal, irrelevant to whether risk is in fact being swapped.

The lesson to be learned from this use, which is also only superficially similar to classic insurance, is that any regulation of insurance company participation in credit derivative markets should focus on how insurance companies invest in credit derivatives. This is especially true since insurance companies are only a small fraction of the entire market in credit derivatives.

3. The Pooling Mismatch

Another reason insurance regulation is a bad fit for credit derivatives is that there is a conceptual difference in the function of insurance and that of credit derivatives. The premise of insurance is risk

pooling. Insurance companies try to spread or pool risk by attracting a sufficiently large number of diverse policyholders such that the law of large numbers will reduce the aggregate variance of claims. In this way, the total amount of risk can be shared by many and thus reducing its impact on any individual in the pool. Counterparties to derivative contracts do not usually do this.

In a single-name CDS contract, there are only two parties, so there is no pooling of risk. When a hedge fund sells protection to bank, it does not act like an insurance company that sells protection to an individual property owner. While the insurance company puts together a diversified portfolio of property owners to generate an actuarially predictable stream of liabilities, the hedge fund does not do this. Hedge funds may try to offset the risk of a particular CDS with other assets and liabilities in their portfolio, but they do not pool risk by writing protection on hundreds or thousands of firms based on predictions about default risk and correlation of risks. Or, to be more precise, they do not always and necessarily do this. These counterparties may be hedging risks and trying to reduce their overall risk exposure, but they are not doing so by pooling a lot of independent risks. Thus, the insurance component of the transaction looks more like simple hedging, which is not regulated as insurance.

The lack of pooling is a conceptual difference, but it may have a practical consequence. Insurance regulation requires insurance companies to hold significant capital reserves in part because if one insurance company fails, a significant amount of beneficiaries will lose. (Importantly, many of these beneficiaries will be average and unsophisticated citizens who are unable to bear the losses. This is the consumer protection angle of insurance regulation discussed below.) The same problem does not exist for credit derivatives generally, unless a single entity, like AIG or Countrywide, makes a multitude of credit derivative bets (that do not cancel or net out the risk of the sum) and the bets made are so large that it threatens the entity and its policy holders or depositors. Note, however, that in the rare cases in which this did or is likely to happen, the independent regulation of the insurance company or bank making the bets exists to ensure that the risks taken by the entity are not excessive. In other words, if the problem is that an insurance company, like AIG, took on excessive risks in credit derivative contracts, then the rules about what investments insurance companies can make should be reformed.

In more complicated credit derivative transactions, such as CDOs, there are multiple parties, and arguably more risk pooling. As discussed above, in a CDO, a new firm (an SPV) is created to sell protection to multiple lending banks, and numerous investors own shares in the newly

created SPV. In this case, the analogy to insurance pooling is more apt. One could view the multiple lending banks whose notes are pooled together in the SPV as the policyholders, while the SPV and its investors are the “insurance company.” While this analogy has more superficial appeal than the case of plain-vanilla CDS’, as discussed below, the policy arguments for insurance-like regulation do not obtain. So even in the case where there is risk pooling – a necessary conceptual component of insurance – there is no policy justification for insurance regulation. This is discussed below.

Even insurance commissioners admit that CDOs are not insurance for this reason.⁴⁰ There is a deep irony here. There is generally no risk pooling – an essential component of “insurance” – in CDS contracts, but these are the contracts that state insurance regulators and pundits consider insurance. In contrast, there is at least some risk pooling in more complex CDO contracts, but there is often no insurable interest in these transactions, so insurance regulators disavow any regulatory oversight of them. The reason for the line drawn by insurance regulators has to do with experience and thus expertise. Insurance regulators are used to dealing with entities that pool risk, are responsible for ensuring an adequate income stream to pay for future liabilities, and are contracting with every-day consumers who rely on the insurance company to make them whole in the event of large personal losses. This experience is obviously not transferable to a market in which none of these traditional aspects of insurance exists, nor are the key regulatory questions. This is explored in greater detail below.

4. Limits on the Reach of Regulation

The artificial distinction drawn by regulators between plain-vanilla CDS’ and more complicated credit derivative contracts points out a bigger problem with any attempt to regulate credit derivatives using an insurance framework. If (insurance) regulation is limited to cases where there is an insurable interest, the contract is not one of simple hedging, arbitrage, or speculation, and there is risk pooling, then this class of cases is like an empty set. If the set of regulated cases is limited, as regulators assert, to cases in which there is privity, there is no risk pooling. Conversely, if the set of cases is limited to where there is risk pooling, there is no privity, and thus the significant line-drawing problems discussed above arise.

More importantly, from a welfare and efficiency perspective, any regulation of one part of the market that does not cover the entire market

⁴⁰ See *infra* note 47 and accompanying text.

will simply redirect market activity to the unregulated market. Regulation means additional costs, and investors will try to avoid these costs if they can do so while achieving the same returns. The fact that credit derivative contracts are simply ways of creating a specific risk-return profile means that the same risk-return profile can be achieved in numerous ways that fall outside of any product (as opposed to institution-based) regulation. For example, if new insurance regulations cover only plain-vanilla CDS contracts, as proposed, investors can replicate the returns they would have achieved with a single-name, real-interest CDS by using a synthetic credit derivative contract that is by the regulators admission, unregulated. In a synthetic CDS contract, the parties do not actually interact with the borrower or lender and do not use the underlying debt instrument as anything more than a probability machine that determines future payoffs between the parties. There is simply no difference between a real and a synthetic CDS contract from the standpoint of these investors, and increased cost on the former will simply mean more of the latter. This fact poses a significant problem for regulators, since there are literally an infinite number of potential contracts and contract forms that can be used by investors to share and transfer credit risk.

Once one form of credit derivative is regulated, other forms will sprout up that will match exactly the same risk-return profile, but these new forms will be unregulated for one reason or another. As discussed below, a more sensible regulatory approach is to: (1) identify investors who are likely to make bad investment decisions on average for one reason or another, (2) ban them from particular forms of investment, (3) and require them to receive special disclosures or protections, or other paternalistic regulation. Lack of sophistication, for instance, provides a central justification for securities regulation, while market failures that may arise out of governance concerns provides the justification for insurance regulation. Both of these, however, are focused largely on the impact on particular investors, as opposed to the nature of the products being sold.

There is a case where synthetic derivative contracts may be used to reduce real risk, and therefore are more like insurance. If the original lending bank enters into a synthetic contract to hedge its risk, it is the same as if the bank enters into a standard credit derivative contract with a protection seller, since the bank is reducing its risk of loss on the default of the original debt. The bank in both cases is seeking regulatory relief from its capital adequacy requirements. But although this type of contract resembles a case where there is an insurable interest more closely, the party on the other side, (the one betting that the borrower will repay the loan) may not know that it is providing insurance of a sort; and even if it did, it

would still make more sense to regulate this contract as a matter of banking law. After all, the bank is engaging in regulatory arbitrage, which may be normatively good or bad from the perspective of banking policy. Banks hedging risk reduces the probability that they will default, and thus jeopardize the claims of depositors. On the other hand, it is possible that some banks used credit derivative contracts not to hedge risks but to increase profits by repackaging loans, moving them off of the bank's balance sheet, receiving regulatory relief, and then bringing the risk back onto the bank's balance sheet through mechanisms that were not transparent to regulators. Again, these issues are largely about banking law – that is, capital adequacy requirements, rules about relief from these requirements, banking oversight, compensation of bank executives, and so on.

5. Moral Hazard Problems and Solutions

One argument in favor of insurance regulation for credit derivatives is based on the fact that both insurance and credit derivative contracts are subject to moral hazard concerns. But the mere existence of moral hazard problems does not justify insurance regulation per se. Moral hazard arises in many contracts and situations that are not deemed insurance. In addition, there are alternative ways of reducing moral hazard short of full insurance regulation. For instance, regulation by other administrative agencies or resolving issues between parties by acting collectively through trade associations. As it turns out, contractual innovation and self-regulatory norms are already being deployed by the International Swap Dealers Association (ISDA) to remedy some of the moral hazard problems inherent in credit derivative markets. As discussed below, there remain some market failures, but none of them are especially redressable by insurance regulation alone.

Moral hazard problems arise whenever any risk is intermediated. Just as one is less likely to take care while driving if one has good insurance (especially with a low deductible), so too is a bank less likely to commit to an efficient level of due diligence or otherwise monitor a borrower if it is going to sell the risk to someone else (and not retain a first-loss position). In equilibrium, investors in the borrowers' credit risk have an incentive to price this potential shirking, and therefore the arrangers of the SPV would have an incentive to choose the credit risks to put in it wisely, lest they be required to offer greater returns to investors. In a frictionless world, in other words, the amount of due diligence would be priced by the market. The credit crunch revealed significant mispricing in

credit markets, but nothing that calls this equilibrium solution into doubt. With learning from the recent collapse, it is likely that this discipline will return to the market.

Even if it does not, there is not much that insurance regulation is likely to add to solve the problem. No regulators or private actors were aware of the mispricing problem, despite the fact that there were numerous regulators, including insurance regulators (AIG is an insurance company after all!) monitoring these markets closely, and despite the fact that investors were betting billions of dollars of their own money on these instruments. To simply declare that more regulation, and in particular more insurance regulation, is needed, is to simply declare the debate over.

Another type of moral hazard is the potential that the parties to credit derivative contracts might act in ways that destroy social value but increase the private value to the party. For instance, a buyer of protection, like a bank, might have incentives to force a borrower to default on a debt in order to collect on the credit derivative contract; even if it is not otherwise efficient for it to do so. This problem arises only because there may be technical defaults that would otherwise not lead to bankruptcy, but the bank buying protection could insist upon enforcing covenants against them now that its downside is limited by its purchase of protection. (The analogy to insurance here is that the buyer of insurance might willingly destroy an otherwise valuable asset to collect on an insurance premium in cases in which the asset has value but this value is less than the value of the insurance policy.) While this is possible, there are at least three things that limit its practical effect.

The first of these is the fact that private contracts take this problem into account, without the need for regulatory mandate. This is not to say that there are not market failures, but simply that this particular problem is not unknown or unremedied in credit derivative markets. As in insurance, where the problem exists too, buyers of protection voluntarily reduce the risk they will shirk because of the moral hazard problem by agreeing to bear some of the first losses that may arise from a default by the original borrower. In insurance, this is called a deductible, and the theory is that it reduces on the margin the incentive of the insured to engage in socially destructive behavior. Credit derivative contracts try to reduce this conduct too – the buyers of credit derivative protection routinely hold the first-loss position so as to signal to sellers of protection that the bank buying protection has some skin in the game and will not engage in this kind of destructive behavior. A deductible, being less than 100 percent of the risk, however, can never fully offset this risk, so there remains some moral

hazard problem. Part of this can be priced by the market, but industry norms can try to reduce it too. This is the next solution.

The second way moral hazard is reduced is through the fact that self-regulatory bodies, like the ISDA, are already aware of this possibility and are structuring industry norms and boilerplate contracts to mitigate these potential problems. The moral hazard issue arose for the first time in the case of credit derivative contracts written on the financial services and insurance firm Consecoco. Credit derivative contracts at that time required a payment from the seller to the buyer of protection in the event that the underlying reference entity—in this case, Consecoco—suffered a “credit event,” which included a restructuring of the reference entity’s bank loans. In 2000, Consecoco’s credit quality deteriorated and began to suffer liquidity problems, so it went to its borrowers in search of a restructuring agreement. The lenders agreed, including an extension of maturity, increased interest rates, and new covenants. The restructuring triggered payment under the existing credit derivative contracts.

This fact created a serious moral hazard problem. The original lenders to Consecoco, who had purchased protection against a credit event, were the ones who got to decide whether to restructure Consecoco’s debt, and thus whether a restructuring event transpired.⁴¹ The lenders could trigger payment simply by agreeing to extend the maturity of the loan or make other trivial changes to the loan that would cost them little (and would be readily agreed to by the borrower) and yet trigger potentially large payments from the sellers of protection. In fact, the situation under the then-prevailing boilerplate terms was much worse than that. Under the ISDA’s 1999 version of the boilerplate terms (called the “Definitions”), the buyers of protection could deliver any debt instrument of the same kind as that on which the lender or other party bought protection. Since Consecoco had a number of outstanding debt instruments of varying maturity, the bank triggering a restructuring credit event could choose the cheapest of these outstanding debt instruments, thereby making large profits on its self-triggering claim. Specifically, Consecoco had short-term bonds that were trading at about 90 percent of face value, while its long-term bonds were trading at about 60 percent of face value. This meant the original lender could declare default, and then buy long-term bonds at 60 cents to settle out its much more expensive short-term bonds. This exposed the sellers of

⁴¹ SATYAJIT DAS, CREDIT DERIVATIVES: CDOs AND STRUCTURED CREDIT PRODUCTS 101-103 (3d ed. 2005).

protection to large losses. The same result obtained in the restructuring of Xerox Corporation in 2002.⁴²

There were two reactions to the moral hazard issues raised by the Conseco and Xerox cases, both private. The market price of credit derivative contracts quickly adjusted to account for the increased risk of contracts including restructuring events as credit events. The contracts “including restructuring as a credit event was 10 to 20 [basis points per year] higher than for credit default swaps without the restructuring credit event.”⁴³

The other reaction was a modification of the boilerplate credit default swap documentation by ISDA. One possibility was that ISDA could simply eliminate restructuring as a credit event, but this was foreclosed by a decision of the Federal Reserve that protection from the risk of restructuring was essential to the transfer of credit risk essential to receiving regulatory relief under the Basle accords.⁴⁴ ISDA therefore issued a “Restructuring Supplement” that provided restructuring would not be a credit event in cases where there were fewer than four holders of the debt in question or where less than a super-majority of unaffiliated holders approved the restructuring.⁴⁵ In addition, the amendments to the boilerplate tried to reduce the potential arbitrage inherent in delivery options for debts with different maturities by requiring any settlement of credit derivative contracts to be made with debt contracts within 30 months of the restructured facility. Further changes to the boilerplate were made in 2003 to address market developments.⁴⁶ The idea with these changes was to reduce the moral hazard problems by contract.

The third non-regulatory way moral hazard is reduced is the presence of countervailing interests on the other side of the transaction that generate behaviors that may cancel out any possibility of abuse. Just as the buyer of protection has incentives to act in a socially inefficient way by

⁴² See Viktor Hjort, *The Xerox Debt Restructuring – A Moral Hazard Issue?*, Morgan Stanley, Fixed Income Research (July 19, 2002).

⁴³ DAS, *supra* note 41, at 103.

⁴⁴ See William Rhode, Fed Says No To Credit Restructuring, DERIVATIVES STRATEGY, Dec. 2000, <http://www.derivativesstrategy.com/magazine/archive/2001/1200shrt.asp>.

⁴⁵ See Restructuring Supplement to the 1999 ISDA Credit Derivatives Definitions, ISDA (2001); see also Donald A. Bendernagel & Oussama Nasr, *Legal Documentation and the Restructuring Debate*, CREDIT DERIVATIVES ISSUES AND OPPORTUNITIES, 2001, at 21.

⁴⁶ DAS, *supra* note 41, at 105.

destroying value to cash in on a credit-derivative contract, the seller of protection has incentives to act in exactly the opposite way. More specifically, if the original lender that bought protection refuses to make a loan that would increase the value of the debtor in order to cash in on a credit protection contract, the hedge fund that sold protection has the incentive to make the loan so not to have pay on the contract. These offsetting incentives only work under two conditions: both parties have sufficient access to capital to provide the loans necessary, and there are relatively symmetric incentives to act; otherwise there may be a socially inefficient level of lending (either too much or too little). Even if the other party to the transaction does not act to deliberately counteract the action, it will be well positioned to detect it and report any misconduct to the market (to impose reputational penalties) or the government (to impose civil or perhaps criminal penalties). (As a side note, insofar as antitrust-based uncertainties preclude collective action on the part of numerous sellers of protection, these rules should be rethought in this light.)

The more generic version of this moral hazard concern is the risk of sabotage. Individuals and entities that are not insurance companies (meaning: not regulated by insurance law) are not permitted to write certain insurance contracts, say, on an individual's life, out of concern that one of the parties will try to sabotage the contract. Or, to look at another way, where there is no symmetry or where the attack and defense would simply result in an arm's race of dead weight costs, the risk created by insurance contracts outweighs any gains. This is especially true when the value of the asset that is the subject of the contract is particularly valuable or difficult to value.

Although this logic might make sense for individuals and contracts like life insurance, the risk of sabotage is overstated in the world of credit derivatives. First, in this \$60 trillion industry, there has never been a reported case in which one party to a contract acted to deliberately sabotage an underlying borrower in order to cash in on a credit derivative contract. Second, the gains from sabotage are as great or greater in equity markets, currency markets, and a whole host of other markets where third parties are able to make large bets on economic outcomes. For instance, a malicious investor could take a large short position in Firm X, and then destroy an asset of Firm X, say by not loaning it money, blowing it up, spreading rumors about it, or any number of activities. This risk is real, but it is uncommon because other laws (for example, criminal law, tort law, and securities laws banning market manipulation) and norms restrain individuals from making socially destructive (but privately beneficial) decisions. However, market participants, observers, and regulators should

always keep a look out for the possibility of sabotage, since reporting it to the market and prosecutors is likely to provide sufficient deterrence.

* * *

As this section has showed, regulators are likely to find it very difficult to draw sensible lines or regulate financial products in a coherent and efficient manner under the rubric of insurance regulation. This means that regulation might more sensibly focus on investors instead of investments. Insurance law is based less on regulating insurance than on regulating insurance companies. In other words, we do not regulate insurance companies because they sell insurance, but rather regulate insurance contracts because they are sold by insurance companies. The right question to ask is not whether credit derivatives are “insurance,” but rather if they are sold or issued by “insurance companies.” This, of course, begs the question of what should count as an insurance company.

The next Part tries to answer this question by looking at the policy reasons for having a separate body of insurance law to regulate insurance companies. The policy reasons are uniquely applicable to insurance companies, not all firms that participate in credit derivative markets, and thus there is no good policy reason for applying insurance regulation.

B. THE POLICY REASONS FOR INSURANCE LAW DO NOT OBTAIN

Given the problems of defining what “insurance” is, it must be the case that the scope of insurance law is either quite arbitrary or based on other considerations. In fact, insurance regulation exists not to regulate insurance contracts per se, but rather to regulate contracts issued by insurance companies. For sophisticated or semi-sophisticate parties, the insurance companies are the problem, not the insurance.⁴⁷ Insurance companies are regulated differently than companies producing other products because of the unique governance problems associated with their production cycle and unique governance structure. Let us consider these in turn.

⁴⁷ There is another justification for insurance regulation that has to do with the consumer-facing nature of some insurance contracts. *See infra* p. 32.

1. Inverted Production Cycle

The first unique feature of insurance companies is the fact that they sell their products long before they are delivered to customers. This means the normal production cycle is inverted. The typical (that is, non-insurance) firm produces products and then sells them in return for cash. Payment and delivery are linked closely in time and there is an immediate feedback from customers. Insurance companies, however, have a much different production cycle that causes unique governance problems. An insurance company's customers are policyholders, who pay (in the form of premium payments) in advance for products (payments on claims) that come many years later, if at all.⁴⁸ This is important because the discipline on how cash can be spent that comes with having to sell valuable products or services in the market is missing or attenuated. Payments are made based on promises alone, and there is thus the risk that the cash reserves given in advance to the insurance company will be squandered on risky investments, and thus unavailable to pay off claims when they come due. This is the Ponzi-scheme problem discussed above. When there is continuous solicitation of investment by outsiders and a mismatched payment scheme (current investments pay liabilities of previous investors), there is a risk that managers will engage in too much risk when the liabilities that arise are greater than predicted. In these bad states of the world, insurance company managers have incentives to attract more capital on irrational terms to pay current liabilities owed to prior investors.

The inverted production cycle of insurance companies has another problem where there are competitive markets for insurance services. In an unregulated market, insurance companies are bound to compete heavily on price, and this may lead to under reserves such that future liabilities will not be covered by sufficient assets. There are two parts of this claim, so it is worth unpacking it.

First, competition among insurance companies is likely to focus primarily and perhaps excessively, on price. This is because the quality of the products insurance companies are selling (the other thing on which they compete) is identical or unobservable. The repayment of losses less the deductible is the same regardless of the insurer. There is some risk that the insurer will fail and be unable to repay the liabilities, but this is something that is, by its very nature, unobservable by the insureds. Reputation and longevity may be correlated with this risk, but these factors coupled with

⁴⁸ The payout for life insurance policies may obviously be decades away.

the governance problems noted above may simply lead to more risky investment choices by the insurer in the next period. Another unobservable component of potential competition is service quality – that is, the timeliness of payments, the cost of the claims process, the quality of customer service, and so on. Like the risk of default, these will be unobservable by the insureds in the period when they make their investment decision, since they happen only many years later and after premiums have already been collected. For these reasons, price is likely the primary way in which insurance companies would compete in an unregulated market.

Second, price competition for insurance products is different from price competition for non-insurance products, and, if unregulated, may lead to pricing at below marginal cost. Non-insurance firms have no incentive to price below marginal cost, since, as a consequence, every sale would lose the firm money. Insurance companies, however, have inverted production cycles, which translates that the costs of the product being sold are felt long after the cash is collected by the firm for the sale. This means myopic managers, hubristic managers, over-confident managers, or desperate managers may charge too little for new insurance premiums. Insurance involves extensive long-range forecasting and the potential for costs, which are realized only after sales, to be much higher than expected. Absent the immediate feedback loop of typical production cycles, the possibility of competition leading to destructive price wars is greater than for normal firms.

It is true that credit derivative contracts are somewhat based on future results and forecasting problems may arise. But this concern is ameliorated by several factors. Credit derivative contracts are generally much shorter in term than insurance contracts, lasting a maximum of five years, and very often held for much shorter than that. So although make-whole payments under the contract may occur in the future, the potential for error is reduced by the fact that forecasting need be made over a much shorter period. The risk is also priced much more frequently, since payments made by protection buyers are due quarterly. For most credit derivative contracts, the prices of buying and selling production are adjusted quarterly depending on the financial condition of the underlying borrower, and this generates the kind of frequent pricing data that is common in regular product markets. Even where it is not, the continuous pricing of the same debt in the market allows holders of risk to engage in pairwise transactions that allow them to rebalance their portfolio on short time horizons. In addition, the parties on both sides are highly sophisticated financial institutions (and their investors), and it is unlikely that any

outsider, like a regulator, could do a better job at estimating the future cash flows from particular debt instruments so as to improve pricing.

The pricing problem for insurance companies is exacerbated by the weak governance structure described below, which may encourage overly risky actions when policy payouts exceed expectations. In short, the managers of an insurance firm that charges too little, for whatever reason, and finds itself unable to meet claims as they come due, may be less constrained by creditors in the kind of response it will take. The managers may simply try to sell more policies to pay off existing claims from other policyholders with the hope of someday righting the ship. This potential that arises from price competition may turn an insurance company into a sort of Ponzi scheme.

This super risk preference situation is unlikely to arise in the case of non-insurance firms because of the discipline of product markets and because of the discipline of creditors when times are bad. For credit derivative firms (that is, SPVs holding the rights to the cash flows from various debt instruments), the probability of this arising is even lower. This is because investment by the SPV managers happens before any investment is made by shareholders in the SPV. A pot of cash is created and then sold, with an implicit promise that no more assets will be added that that particular pot. The possibility of super risk preferring managers is thus extremely unlikely.

2. Weak Governance Structure

The second unique feature of insurance companies is the weak corporate governance structure that is inherent in the insurance company model. Non-insurance firms are generally funded by a large number of diffuse shareholders and a small number of concentrated creditors, typically banks or other lenders. In this governance model, the shareholders are the residual claimants of firm value, and in good times it is in their interest that the managers operate the firm.⁴⁹ The diffuse nature of the holdings of equity, coupled with the business judgment rule, means that firms have a lot of slack in the risk they take during good times. When things turn for the worse, however, the interests of the shareholders are set aside and the concentrated interests of banks and other lenders take over

⁴⁹ Assuming, of course, the managers aren't acting in their own interest.

the decision-making process.⁵⁰ Loan covenants are enforced, contracts are rewritten by the lenders, managers are replaced, and a restructuring of the firm is undertaken. The reason creditors take effective control long before bankruptcy is because in the absence of these control rights, shareholders (and their agents, the managers) would have incentives to act ever more risky as the value of shares drops. Consolidated creditors can help preserve going concern value and reduce dangerous risk taking in near final periods. (It is, of course, irrelevant whether this concentration of creditors takes place when debt is issued (when times are good) or when it is consolidated (when times are bad), since concentration is less likely in insurance companies in either case because the creditors of the firm are its policy holders.)

Unlike non-insurance firms (but like banks), insurance firms are structured with both weak equity holders and weak creditors. Insurance firms (and banks) have shareholders that are indistinguishable from other firms, but their creditors are as diffuse and disinterested as their shareholders. An insurance company's creditors are its policyholders. Their large number makes coordination difficult, either for monitoring or action, and information costly and very unlikely to be obtained. In addition, policy holders are not investors (like many shareholders are), and therefore likely to be unaware of and unsophisticated about matters of corporate governance and finance. And unlike diffuse creditors of non-insurance firms (e.g., bondholders), the claims of policyholders cannot be and are not consolidated or concentrated in periods of distress. This means an increased threat of excessively risky decision making in bad times because the insurance company's creditors are diffuse instead of concentrated. Both in insurance and banking, where depositors are substituted for policy holders, this suggests the need for a prudential regulator to effectively consolidate the diffuse policyholders into a bank-like consolidated creditor to deal with the insurance company in bad times.

This governance problem is not present consistently, if ever, in credit derivative transactions. To see this, consider the simplest case of a plain-vanilla CDS. Remembering the analogy with insurance set forth above, the bank that lent the money to the borrower is the insured, and thus analogous to the policyholder in an insurance contract. In a single-name CDS contract, there is no pooling and therefore no diffusion of interest

⁵⁰ See, e.g., Douglas G. Baird and Robert K. Rasmussen, *Private Debt and the Missing Lever of Corporate Governance*, 154 U. PA. L. REV. 1209, 1209-11 (2006).

among the alleged insurer's "policy holders." In addition, banks that buy credit protection are nothing like individuals that buy auto or life insurance policies. Banks are sophisticated, repeat players, represented by counsel, capable of processing information about the riskiness of their counterparty, and have tremendous bargaining power.

It is possible for a protection seller to write many CDS contracts, and thus put any individual buyer of credit protection into the position of holding a small claim against the firm, say insurance company or hedge fund. Although this would not change the nature of the protection buyer and the ability of it to fend for itself, it does raise potential concerns about the management of the protection seller. But this just then puts the inquiry about governance back at that level, and tied to the nature of the seller of protection. If the protection seller has a weak governance model, like that of a bank or an insurance company, then it may be susceptible to this problem, but if it is an entity with strong governance in bad times, then the concern about too much risk on the part of managers (on behalf of shareholders) is much lower. Hedge funds, for instance, must return to the market frequently for capital (that is, they do not have capital lock in) and are funded by extremely sophisticated investors. They are decidedly not subject to this concern. In short, insurance companies with weak governance should be subject to regulation to avoid the social inefficiency that might arise from their governance structure, while non-insurance companies, with strong governance, are less worrisome.

The same result obtains even when we consider a more complicated credit derivative contract. The parties buying protection that have their default risk pooled into a CDO structure are large financial entities with much greater sophistication and risk-bearing ability than individuals buying typical insurance products. The risk that the sellers of protection will "pull a fast one" on them is much lower given this sophistication. In addition, the investors in the SPV holding the default risk (the analogous insurance company) are likewise large financial entities capable of making risk assessments, demanding and processing information, pricing risk, and wielding their bargaining power in the event a bad future state arises.

Moreover, the nature of the typical CDO structure is effectively a one-time game, in which credit risks are pooled and the cash flows sold off to investors. The sponsor and manager of the SPV does not continue to sell protection based on a pool of funds provided by investors (as in an insurance company), but rather makes the investments first (by choosing risks to pool), then goes to the market to sell cash flow rights to investors. This means that managers of the SPV do not really do much or can do

much in terms of risk alteration of the SPV once it has raised money from investors. The future course of the SPV is set, and the payouts are what they are. No future investments are made, no additional investors are brought in, and there can be no change in strategy for the firm. Therefore, there is less chance for abuse in the event the SPV payouts are less than expected. Governance quality is largely irrelevant in this model firm.

Applying this governance model to the insurance company model, it is as if the insurance company wrote all of its policies before raising money in the market. In that case, investors would worry less about the governance of the insurance company, since its job would simply be to process claims from the policies it had written – it would not take on new policies (and a new source of cash) on terms likely to be unfavorable to existing investors. There would still be some governance risk, however, since the decisions on what policies to pay out on, how much to compensate executives, and other firm costs still have to be made. In some of these, managerial interests may be aligned with those of investors, while in others they may diverge. Importantly, however, this residual governance risk is not present in the CDO case, since all of these decisions are made before the investment in the firm (for example, management fees) or are automated (for example, the amount of payouts). In short, any governance problems simply do not obtain in the typical structure of credit derivative contracts.

3. Consumer Protection

The third policy reason for a separate body of insurance law is the need for strong consumer protection. While the concern with the inverted-production-cycle and governance problems was basically insurance firms not charging insureds enough, the consumer protection concern is that insurance firms will charge too much. As mentioned above, the concern is based on the following syllogism: insurance is a critical product for most individuals; individuals are not sophisticated about insurance products or contracts; and therefore insurance companies will take advantage of customers by overcharging them. Accordingly, (the bulk of) state insurance law regulates the substance and terms of insurance policies (to make them simpler to understand and compare across firms), as well as regulating service and coverage issues (to make sure insurance firms do not back away from promises to pay). In other words, insurance is sometimes regulated as a specialty consumer product in which informational and bargaining power asymmetries are sufficiently large that social losses may be generated from an unregulated market.

The credit derivatives market described above has none of these characteristics or concerns. The parties to credit derivative transactions are all large financial institutions or other sophisticated investors with access to information, the ability to understand and process the information, bargaining power, and the ability to bear losses. This is in sharp contrast with insurance contracts entered into by average consumers, who have none of these attributes. It may be sensible for insurance regulators to try to reduce informational and bargaining power asymmetries between insurance companies and consumers, to provide oversight of claims management and customer service, to provide standardized contract terms that allow comparison shopping, and to even regulate rates, but these policies are unnecessary where the buyers and sellers of “insurance” are large financial institutions. In fact, if anything, the sellers of protection (the alleged insurance companies) may often be less sophisticated than the buyers of protection (the alleged insureds). For example, a small hedge fund run by a few investors may enter into a contract to sell protection to a large commercial bank. In this case, it is not at all clear where insurance-law-like consumer protection duties should run. After all, existing law will treat both the hedge fund and the bank as not needing the protection of the securities laws or other regulations.

In addition, as discussed above, standard-setting groups, like the International Swap Dealers Association, are already doing much of the work for credit derivative markets that insurance regulators do to protect average consumers. ISDA provides, among other things, standard form contracts (which innovate to respond to changes in the market), dispute resolution mechanisms and guidance, information, educational services, and so on.

C. INSURANCE LAW DOESN'T WORK WELL AND WON'T
UNIQUELY ADD MUCH

A final argument against treating credit derivative contracts as insurance is a practical one having to do with the value that insurance regulation, as currently constituted, might add to the market. In short, insurance law and its generation and enforcement regime is generally considered to be inefficient and in need of dramatic reforms, and is therefore not the most appropriate locus of authority for a regulation of a new financial innovation that spans numerous types of institutions and serves innumerable purposes, most of which have nothing to do with insuring against risk as it is commonly understood.

1. Jurisdictional Issues

As discussed above, insurance regulation is state based, meaning there are at least fifty different regulators and models of regulation. The state-based model is largely premised on the consumer-protection function of insurance regulation, since it is the state police power to protect citizens from abuse that justifies a local approach. While the merits of this model are debatable when it comes to providing efficient insurance services for health care or automobile risk, the global nature of modern financial markets makes the local argument much more difficult to make for credit derivatives. Do citizens of New Jersey need different protection in credit derivative markets than citizens of New York? In fact, insurance companies seemed to exploit this regulatory fracture by dividing up their businesses into discrete components that were regulated piecemeal by various state regulators. Given the ease of capital flows, the ability of firms to incorporate anywhere around the globe, and the fact that even transactions in the Cayman Islands can impact investors around the world, the idea of insurance regulators in a particular state controlling the global market seems fanciful. New York regulators, for instance, had authority over less than 10 percent of AIG’s operations, because of the corporate structure of AIG.⁵¹ As a result, one of world’s biggest insurers was, under the current system, able to be largely below the radar screen of its primary insurance regulators. Applying this dysfunctional model (which is based primarily on consumer issues, also which are the least applicable to this market) to new financial products makes little sense.

The choice of regulator, be it a question of a particular entity or a general jurisdictional choice (like federal or state), is based on regulatory expertise, incentives, and the expected consequences of the regulatory model on the ability to minimize the decision costs and error costs of regulation. State-based regulation might make sense if jurisdictional competition is likely to lead to the efficient regulation (that is, the race-to-the-top theory of state-based corporate law). But this is not the basis for state-based insurance law. State insurance law is not based solely on the state of incorporation of the insurance company, but rather the locus of policy holders. In any event, this model will probably not work for a market like credit derivatives. Financial markets are generally regulated by

⁵¹ See *This American Life* #382: *The Watchmen* (Chicago Public Radio broadcast Jun. 5, 2009), available at http://www.thislife.org/extras/radio/382_transcript.pdf.

federal agencies (or perhaps in the future by international ones) because financial products are sold in a global market, and any state-based regime could be avoided by simply opting into offshore regulatory regimes. If the goal of regulation is to encourage credit derivatives to be traded on exchanges as opposed to in over-the-counter markets, as argued below, having a single regulator to choose from is crucial to creating commodifiable products. If credit derivative-holding SPVs could opt into one set of regulations through choice, this might provide some federalism-esque benefits, but this is not the way insurance law operates.

Another factor influencing the choice of regulator is expertise. Here too, there is nothing about state insurance regulators that seems special or unique. State insurance regulators are used to dealing with insurance companies and insurance contracts, which, as described above, deal with issues of risk pooling, governance problems, consumer-facing contract issues, and the like. None of these obtain in credit derivative markets. Insurance commissioners are also generally concerned with counterparty risk—a real concern in credit derivative markets—but this is something bank regulators (like the FDIC, Federal Reserve, and Treasury), derivative regulators (like the CFTC), and securities regulators (like the SEC) are also especially concerned with. In addition, these latter regulators do not have the state overlap problem described above.

2. Substantive Law

Deeming credit derivatives to be “insurance” (or credit protection sellers “insurance companies”) would, under current law, have several consequences, none of which is likely to improve the efficiency of credit derivative markets.

i. Licensing

First, entities could not sell protection unless the seller was a licensed insurance company. All fifty states require a state-issued license before a firm may issue an insurance policy.⁵² Such a pre-screening requirement might make some sense as part of trading on a credit

⁵² See N.Y. INS. LAW. § 1102(a) (2006). “No person, firm, association, corporation or joint-stock company shall do an insurance business in this state unless authorized by a license in force pursuant to the provisions of this chapter, or exempted by the provisions of this chapter from such requirement.” *Id.*

derivatives exchange, say by requiring those buying or selling on the exchange to meet certain criteria, like margin requirements. However, certifications based on other factors unrelated to the exchange’s risk of default would add costs without any offsetting benefits, and would merely open up the possibility that regulators could extract rents from firms wanting to participate in these markets. As discussed below in the context of capital requirements, the licensing scheme for insurance companies is doable in part because there are only a handful of firms providing insurance in each state. In contrast, there are literally tens of thousands of investment funds that have sold or could sell credit protection in credit derivative markets, and this would make any licensing scheme prohibitive or meaningless for state regulators. It would also impose potentially large costs on funds who do not sell protection as a normal part of their investment strategy, but might find it efficient and sensible to do so in limited cases. Regulatory costs would therefore deter these funds from participating in the market, without any proof that the funds have imposed any costs on others.

ii. Duties

Second, the buyer and seller of protection would be subject to a duty to act with the utmost good faith, that is, something beyond the “morals of the marketplace.” This might make some sense for markets in which buyers and sellers are of widely differing sophistication, have access to different information, and have different bargaining power, but it makes much less sense when the parties on both sides of a transaction are similar giant financial institutions. In fact, the trend in the market is for large investors to opt out of these kinds of disclosure requirements and the like using waivers known as “Big Boy” letters. To impose fiduciary duties or other litigation-generating obligations on parties without the potential for opt out will increase uncertainty and costs without any obvious benefit from ex post judicial determinations of what were and were not good deals.

iii. Capital Reserves

Third, protection sellers would be required to maintain a certain amount of capital based on the risk inherent in its “insurance-based” business.⁵³ For instance, state insurance regulation requires every insurer to

⁵³ See N.Y. INS. LAW. §§ 1322, 1324 (2006 & 2009).

maintain certain specified amounts of capital on hand to reduce counterparty risk and to submit its risk-based capital levels to regulators on an annual basis.⁵⁴ The amounts of capital required vary by jurisdiction and entity, depending on the riskiness of the insurance company.⁵⁵

In the abstract, capital reserves are unobjectionable, since they are about reducing counterparty risk, and therefore about increasing the number of socially beneficial transactions. After all, these requirements are a common element of banking law and other areas where counterparty risk and the problem of runs and systemic risk are present. The question is then how much capital should be required, what is the best way to reduce counterparty risk, and which regulator has the incentive to figure these things out. There are several reasons why insurance regulators are not obviously the best candidate to fulfill this mission, and why the solution of requiring credit derivatives to be exchange-traded is a more elegant solution, albeit one fraught with problems too.

There are several problems with insurance regulators imposing this requirement on the credit derivative markets. For one, the number of entities and individuals writing protection on indebtedness is enormous, making any pre-screening regulation extremely costly. For example, there are over 15,000 hedge funds in the United States alone, each of which could be a participant in these markets.⁵⁶ The magnitude and complexity of the job of simply tracking each of these hedge funds—only one type of protection seller—would be beyond the capacity of any existing state

⁵⁴ See N.Y. INS. LAW. § 1402(a) (2006). “Before investing its funds in any other investments, every domestic insurer shall invest and maintain an amount equal to the greater of the minimum capital required by law or the minimum surplus to policyholders required to be maintained by law for a domestic stock corporation authorized to transact the same kinds of insurance, only in investments of the types specified in this section which are not in default as to principal or interest.” *Id.*

⁵⁵ The Model Insurance Act, for instance, provides for three “risk-based capital” levels: (i) mandatory control level risk-based capital (measured at .7 times authorized control level risk-based capital), (ii) regulatory action level risk-based capital (measured at 1.5 times authorized control level risk-based capital), and (iii) company action level risk-based capital (measured at 2.0 times authorized control level risk-based capital). NAT’L ASS’N OF INS. COMM’RS, MODEL LAWS, REGULATIONS, AND GUIDELINES vol. III (2009).

⁵⁶ See *Number of Hedge Funds up by Two Thirds in Two Years, Concludes PerTrac Study*, THE TRADE NEWS, Mar. 5, 2007, available at <http://www.thetradenews.com/hedge-funds/prime-brokerage/613>.

regulators. The job would be made even more overwhelming since selling protection is not necessarily a full-time job. Hedge funds and other sellers may hold the default risk for a very short time (a few months or less) and may engage in transactions only periodically or on a one-off basis. Unlike insurance companies that exist to provide risk sharing services, protection sellers are not necessarily in the business of holding debt risk. The fluid nature of market participants would make any licensing or ex-ante regulatory regime incredibly costly and drive many participants out of the market.

In addition, capital requirements did not work well if at all in preventing insurance companies, such as AIG, from investing aggressively and, as it turns out, dangerously in credit derivative markets. The state-based model was manipulated by AIG and others, and this possibility could only be expected to be worse if every credit derivative protection seller becomes a ward of insurance regulators. In other words, the job of regulation would get much more difficult without any obvious way of increasing the capabilities of regulators. This point is made even clearer by reiterating the point made above about how insurance regulators are not experts in financial markets in which most protection sellers participate. If insurance companies can avoid insurance regulation, it is very likely that hedge funds and other sophisticated and fast-moving private money funds will also be able to do so.

Moreover, capital adequacy requirements imposed by regulators (as opposed to margin requirements required by exchanges) generated the incentive for regulatory arbitrage described above. Firms subjected to these requirements had incentives to hold higher quality debt risk, which received lower capital charges, and to move debt risk off of their balance sheets and into bankruptcy-remote SPVs. Although this type of arbitrage is likely inevitable at some level, the current regulatory model for insurance proved ineffectual at preventing arbitrage that imposed systemic risk externalities on society.

Finally, insurance regulators are not experts about the amount, type, and structuring of capital requirements to reduce counterparty risk in non-insurance financial transactions. It is arguable that insurance regulators, representing the state, have incentives to determine the amount of social cost from the failure of an insurance company, since many of the social harms that would result would be paid for by a state-funded social safety net or would otherwise result in state-based harms. But the failure of a hedge fund or foreign bank or other protection seller may generate no social losses, because gains from bets on one side cancel out losses from bets on the other side, or are ones that are not clearly within the purview or

concern of state regulators. Systemic risk is something not felt completely or even directly by one state, and therefore a collective action problem may generate insufficient incentives to get the regulation to the efficient level.

It is important to note that, as proposed below, trading credit derivatives on an exchange would likely require some financial assurances akin to capital adequacy requirements on market participants, through margin requirements, and on the exchange, which would be the ultimate bearer of counterparty risk. For the reasons discussed below, the concerns here are much less than through regulatory capital reserve requirements. For one, exchanges, which act as a centralized counterparty, bear the entire risk of loss if a trading party defaults, and therefore have the best incentives in terms of setting up rules to ensure that traders are likely to pay for their losses.

iv. Disclosure

Fourth, being an insurance company would trigger a detailed disclosure requirement of any insurance business to state regulators. The state-based requirements track roughly those firms with publicly traded securities. Audited reports of the insurance company's financial and accounting situation must be made quarterly and annually.⁵⁷ These include disclosure of routine data, like the firm's balance sheet, income statement, and statement of cash flows, as well as more detailed information than generally required by securities laws, like a list of every asset owned by the firm, every asset acquired or sold during the relevant period, a report of all derivative transactions, and so on.

Although disclosure of the assets and risks of hedge funds and other private wealth pools may indeed be a socially valuable regulation, there is no obvious reason why this should be tied to a regulatory apparatus that is about only a very small part of what a hedge fund may be doing or may have done. As noted above, there are potentially tens of thousands of separate legal entities participating in credit derivative markets at any time, and requiring each of them to make disclosures to insurance regulators upon engaging in such activity is highly problematic.

As a basic principle, disclosure regulation should be implemented and monitored by regulators that cover the primary activity of the regulated entity. It is for this reason that the SEC is the agency responsible for the

⁵⁷ See N.Y. INS. LAW §§ 307(a), 308 (2006) (requirements for filing an annual financial statement and a quarterly financial statement).

disclosure by firms issuing securities and by brokers and dealers that buy and sell securities, and the reason that insurance regulators are responsible for the disclosure by insurance firms. To require an investment fund to make disclosures to insurance regulators solely because it engaged in a credit derivative transaction will impose potentially large costs on funds based on potentially a single or small number of transactions. If the disclosure rules are tied instead to how often a fund trades or how many trades it makes in these markets, the funds will inevitably try to avoid these costs by making the decisions on whether to sell protection based in part on the arbitrary triggers. For example, if ten incidents of protection selling trigger a disclosure obligation, we shouldn't be surprised to see funds selling protection nine times.

If the reason for insurance-based disclosure rules is because of the lack of disclosure to other regulators – hedge funds have little or no disclosure obligations under the securities laws – this is not an argument for disclosure to insurance regulators, it is an argument for a securities law disclosure regime. The regulator that can best calibrate what kinds of disclosure are cost justified, what form disclosures should be made in, and what to do with the disclosed information, if anything, is the regulator that should require and monitor disclosure. For one, it is not clear what insurance regulators would do with the disclosures, especially if the bulk or almost all of it is about activities that are unquestionably not insurance.

v. Contract Regulation

Fifth, state law generally requires insurance companies to submit insurance contracts, known as “policies,” to state regulators for pre-approval before any policies can be sold using the contract. For example, in New York, contracts for life, accident, and health insurance are subject to prior regulatory approval.⁵⁸ This requirement would layer possible fifty different state law requirements on top of existing private contracting in the over-the-counter credit derivative markets. There are several problems with such a requirement.

Most obviously, as noted above, there is already a quasi-regulator, the ISDA, that provides industry-wide boilerplate contracts for credit derivative transactions. As the Conesco and Xerox examples above

⁵⁸ Under New York law, life, accident and health and annuity policy forms are subject to prior regulatory approval. *Compare* N.Y. INS. LAW. § 1102(a), *with* N.Y. INS. LAW § 1108(a) (2006).

illustrate, ISDA is already incorporating best practices into these standard contracts, as well as modifying terms that cause problems regulators would be concerned about, such as manipulation, externalities, and contractual unfairness that may have arisen from any bargaining power asymmetries, mistakes, or the like. As such, contract regulation is likely redundant, and would in any event be replacing a highly knowledgeable set of regulators with one without any experience with credit derivatives.

Paternalistic contract regulation to protect one party or the other is also unnecessary because of the sophistication of the parties to these contracts. There is no obvious systematic bias in favor of one party or the other in these contracts, and the typical arguments that may justify contract form and substance regulation – for example, information or bargaining power asymmetries – do not obtain or point always in one direction.

A final point has to do with the fact that the parties to credit derivative contracts are not tied to physical locations in the way that insureds are, and therefore any state-based regime will inevitably invite avoidance through incorporation choice or choice of law provisions. This may be viewed as normatively good or bad (the old race to the top versus race to the bottom debate), but even where it might be thought of as generating efficient contract forms that private parties would choose in any event, it would take us simply to the current ISDA model. After all, if there were a more efficient set of contracts that could be written – that is, the one that parties freely choosing would choose anyway – it would exist or will exist under the current quasi-regulatory regime.

vi. Price Control

Finally, states impose substantive restrictions on the prices that can be charged by insurance companies. Regulation of prices varies widely by state and by the type of insurance, but a few common themes are apparent. There are generally three types of regulation: pre-approval, “file and use,” and “use and file,” with the strictness of the regulation decreasing accordingly. For example, New York law requires prices for workers’ compensation and automobile insurance to be approved in advance by regulators, while rates for property and casualty insurance are subject only to a pre-issuance filing policy.⁵⁹ The general regulatory touchstone is that rates shall not be too high, too low, discriminatory, or anticompetitive.⁶⁰

⁵⁹ Compare N.Y. INS. LAW §§ 2305(b), 2310(a), 2344 (2006), and N.Y. COMP. CODES R. & REGS. tit. 11 §§ 161.1-161.12 (2009), and N.Y. INS. LAW §§

Whatever sense price regulation makes for contracts written by insurance companies, it makes much less sense for credit derivative products. Price regulation of insurance contracts is premised on the pooling of large numbers of individuals and on the governance problems described above. In contrast, credit derivative prices are based on the idiosyncratic risk associated with particular firms. This is something that is difficult to price in the abstract or to know when prices are too high, too low, discriminatory, or the like. In addition, because the price of risk is traded in markets, the idea of using regulators, especially ones without any expertise or experience in this area, to set prices is nonsensical.

There is one area where the pricing of credit derivatives was erroneous. As shown in recent research, price models used ubiquitously by buyers, sellers, credit rating agencies, and other participants in the markets systematically mispriced various tranches of risk.⁶¹ It turns out that highly rated tranches were underpriced, meaning they were riskier than buyers and sellers thought, and unrated tranches were underpriced, meaning they were less risky than thought.⁶² Importantly, however, no one was aware of this problem, even though everyone had strong incentives to be so. In addition, this kind of error is now known, and parties to these contracts do not need regulators to inform them about these errors. Other pricing issues might arise in the future, but market participants have incentives to identify such issues. The problem was not that the market for setting prices was biased in one way or the other, but rather simply a mistake in assumptions. Regulators are not well positioned to remedy these kinds of problems absent a crystal ball that no one believes they possess.

Moreover, if various states are competing with each other to offer market participants pricing regulations that fit their needs, the jurisdictional choice point made above will obtain – contracts will migrate to those states that offer the pricing rules that the parties would have come to anyway.

2303, 3231(d) (2006).

⁶⁰ Compare N.Y. INS. LAW §§ 2305(b), 2310(a), 2344 (2006), and N.Y. COMP. CODES R. & REGS. tit. 11 §§ 161.1-161.12 (2009), and N.Y. INS. LAW §§ 2303, 3231(d).

⁶¹ See Joshua D. Coval, et al., *Economic Catastrophe Bonds*, 99 AM. ECON. REV. 628 (2009) (showing how AAA-rated tranches contained very little to no idiosyncratic risk, but large and underappreciated amounts of systematic risk).

⁶² See *id.*

V. AN ALTERNATIVE TO INSURANCE REGULATION

It is understandable why people mistakenly analogize credit derivatives to insurance: insurance is about risk sharing and diversification, and this is what credit derivatives are about as well. Insurance companies were also big players in credit derivative markets. But other contracts are about these things as well, and there were many other types of entities that participated in these markets. In addition, credit derivatives are about many other things than risk sharing. In fact, as shown above, credit derivatives may have started as a risk-sharing or risk-transferring mechanism, but their primary use was and is speculation, hedging, and other non-insurance-like functions. Moreover, even where the insurance analogy is most apt, it does not follow that the current insurance regulatory regime is the best available for credit derivatives, assuming additional regulation is needed.

There may be a case for more regulation, premised on the failure of the market to adequately address counterparty risk issues, but insurance law has little to add. A simple rule requiring derivative contracts to be traded on an exchange in most cases will do most of the work required.

As noted above, a credit derivative does not eliminate risk for the original bearer of it, but simply trades default risk for counterparty risk. In other words, the risk in a loan that the borrower will not repay is traded for the risk that the seller of default protection will not pay in the event the borrower does not. This counterparty risk was bigger than anyone thought; firms no one thought would fail, like AIG, failed by taking on too much risk. This led to a cascade of failures of brokers, like Lehman Brothers, and other intermediaries, which in turn led to huge collateral calls and a general constriction of credit flows. Quite simply, the mispricing of and realization of counterparty risk caused the credit crunch.⁶³

Fortunately there is a somewhat simple solution to reducing counterparty risk – an exchange. Using a centralized exchange, like the Chicago Mercantile Exchange, eliminates the counterparty risk, replacing it with the risk of default of the exchange. If A and B have a contract that exposes A to a net of \$100 in risk to B, this risk can be eliminated if A and B both trade through a centralized clearinghouse or exchange. A will now have a \$100 liability to the clearinghouse, while B will have a \$100 credit

⁶³ See JOHN B. TAYLOR, GETTING OFF TRACK: HOW GOVERNMENT ACTIONS AND INTERVENTIONS CAUSED, PROLONGED, AND WORSENERED THE FINANCIAL CRISIS (Hoover Inst. Press 2009) (showing how the credit crunch was not caused by a liquidity shortage but by an increase in counterparty risk).

with the same. If A defaults on its ability to pay B, B can still be paid by the clearinghouse. As long as the clearinghouse is solvent, the counterparty risk for B is eliminated.

The risk-reducing qualities of an exchange can be seen more clearly when the number of trading parties increases. Consider the case where A is owed \$100 by B, B is owed \$90 by C, and C is owed \$80 by A. In this case, A has a net risk exposure of \$180 to B and C, since if they both default, A is owed \$100 from B and owes \$80 to C. If these three liabilities are managed through an exchange, however, A’s risk exposure to B and C is reduced to zero. In this scenario, A is owed \$20 from the exchange, and B and C each owe the exchange \$10. Thus, A’s risk to B and C has been eliminated, and the netting of liabilities has reduced the magnitude of the overall amounts owed to much more manageable sums. So long as the exchange can ensure, say through margin requirements, that B and C can make good on their \$10 (about 10% of the total bets), the market stays together.

It is for this reason that the clearinghouse model is used for other derivative markets, like commodities markets, futures markets, and currency markets. Of course, the clearinghouse must be solvent and for this it needs sufficient scale of operations and the ability to impose rules on trading parties that help reduce the risk that they will not be able to make due on their contracts. This last point is precisely about the locus of regulatory authority – who knows better how to regulate the leverage or other characteristics of market participants? Since the clearinghouse, typically a for-profit enterprise, stands to lose personally and dramatically in the event of a failure, it has arguably better incentives along this dimension than government regulators, who are not betting their own money and who, perversely and ironically, may see increased funding from any failures.

Given these benefits, the question is why the exchange did not arise as a natural part of the market. One answer might be that an exchange has some elements of a public good, since it reduces the potential for systemic risk by decreasing the likelihood of a credit crunch from the failure of a single firm, and public goods are chronically under supplied by the market. But the story here is more complicated, because the analysis above suggests that it is in each individual firm’s interest to reduce risk in this way. Moreover, the collective action problems that typically cause the market to under supply public goods do not obtain in this context, since there were only about eight major intermediary market makers, and they were all located in New York City.

There must be, therefore, some benefit to writing contracts off exchange that exceeded any benefits from counterparty risk reduction, at least in expectation before the actual counterparty risks, like the failure of AIG, were known. One possibility is that the brokerage houses arranging over-the-counter credit default contracts are able to earn higher profits for writing specialized contracts than they could for simply dealing in standardized contracts, as typically required for exchange-traded contracts. There is less pricing transparency in over-the-counter contracts, since they are written for a specific buyer and seller in a one-off fashion, and there is also more work that arguably goes into writing these contracts. And, private firms do not bear the full costs of the over-the-counter system, since bankruptcy law limits the downside risk to any investor to what they invested. Under this view, brokerage firms are able to capture the private benefits of idiosyncratic, over-the-counter contracts, while externalizing the risks of systemic meltdown of the entire system.

In this way, the government's initial efforts to encourage the trading of credit derivatives on an exchange is a sensible reform. Firms have resisted this to date,⁶⁴ because nothing has changed the private incentives with respect to systemic risk – in fact, the rash of bailouts of private firms have arguably exacerbated the problem. In addition, there are multiple competing exchanges, including the CME and ICE exchanges, and academics have shown that exchanges need a great amount of scale to be able to adequately reduce counterparty risk.⁶⁵ The government may be rightfully worried about choosing one exchange as the preferred or exclusive exchange, but the need for scale may force some collective choice to be made.

As noted above, the virtue of the exchange model is that it bakes into a private-ordering system many of the laudable aspects of the insurance law regime. Specifically, capital requirements, disclosure, pricing transparency, and general oversight of risk are all functions that exchanges provide, since exchanges are on the hook for losses arising from the failure

⁶⁴ See Jacob Bunge, *NY Fed Examines Slow Progress in CDS Clearing*, WALL ST. J. MARKETWATCH, Apr. 1, 2009, available at <http://www.marketwatch.com/story/ny-fed-examines-slow-progress-cds>.

⁶⁵ See Darrell Duffie & Haoxiang Zhu, *Does a Central Clearing Counterparty Reduce Counterparty Risk?* (Rock Center for Corporate Governance, Working Paper No. 46; Stanford Graduate Sch. of Bus. Research, Paper No. 2022, 2009), available at http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1348343 (showing how reductions in counterparty risk is highly correlated with scale and the ability to net with other derivative contracts).

of any of the market participants. Moreover, insofar as there are multiple exchanges competing to act as a clearinghouse, there will be competition in law making, which will increase the chances of efficient rules being created. In the private model, there is also less chance of regulatory capture or a public choice distortion, because rival exchanges can always arise to offer market participants alternatives. This assumes, however, that entry is relatively unrestricted, something that is not necessarily true in a world where scale is so important and perhaps difficult to achieve quickly. Insofar as this is true, some oversight of the exchange(s) may be required to simply ensure that they are not subject to these shortcomings. A first guess at a sensible regulator of the exchange(s) would be one of the existing regulatory bodies that deals with exchanges (e.g., the SEC or CFTC) or the regulators that deal with banks and systemic risk (e.g., the Federal Reserve or the Treasury Department).

VI. CONCLUSION

This Essay has shown that the simple argument that some credit derivatives help banks and other providers of debt share risk with other investors is not sufficient for credit derivative contracts in general to be deemed “insurance.” A separate body of insurance law exists not because the underlying contracts are insurance, but rather because typical insurance contracts are sold by insurance companies. It has also shown that the policy justifications for regulation of insurance companies—an inverted production cycle, weak corporate governance in bad times, and unsophisticated insureds—do not obtain in the context of credit derivative markets or apply to parties to credit derivative contracts. Finally, it has shown how an exchange for credit derivative contracts can provide most if not all of the substantive regulation insurance regulators can provide, at lower cost and in a more efficient manner. There remain unsolved problems with the exchange solution, including issues of scale and bilateral netting, but this is a subject for another day.

THE ROAD FROM “TWIN PEAKS” – AND THE WAY BACK

Michael W. Taylor*

This article explores the fragmented regulatory structure of financial markets in the United States in light of the current financial crisis. Two approaches for regulatory reform that originated in the United Kingdom are presented. The first approach is the creation of a unified regulatory agency responsible for regulating all the main segments of the financial services industry. The second, also known as the “Twin Peaks” approach, is to structure regulation around two agencies, one responsible for the safety and soundness of all financial firms and the other for regulating their sales practices. This article describes the debate in the UK prior to the creation of one unified regulatory agency, the Financial Services Authority (FSA). Next, it explores justifications for a single regulator, such as the FSA, followed by a discussion of the rejection of the “Twin Peaks” approach in the UK. Subsequently, the debate regarding the role of a central bank, like the Bank of England in the UK, is discussed. Then US regulatory reform is reviewed in terms of the lessons of the British experience of creating a single regulatory agency. Finally, the concluding section describes how some variation of the “Twin Peaks” alternative would prove to be more successful than the single regulator approach.

I. INTRODUCTION

The Global Financial Crisis has put the spotlight on the United States’ complex and fragmented regulatory structure as an issue of global systemic importance. The failure of large investment banks like Bear Stearns and Lehman Brothers has put into question the adequacy of the regulation of large non-bank financial intermediaries. The lack of consolidated supervision of the AIG group, with its Financial Products

* Adviser to the Governor, Central Bank of Bahrain; formerly Head of Banking Policy, Hong Kong Monetary Authority; Senior Economist, International Monetary Fund; Reader in Financial Regulation, ICMA Centre, University of Reading.

Division falling under the authority of the Office of Thrift Supervision (OTS) while the insurance company was regulated at State level, further illustrates the systemic problems created by regulatory fragmentation. Finally, the Commodities Futures Modernization Act, which deliberately excluded the regulatory authority of both the SEC and the CFTC from the Credit Default Swaps market, resulted in a failure to ensure adequate regulation of that market with implications for the global financial system.

In a message clearly directed to US policy-makers, the Group of Thirty, a think tank comprising some of the most distinguished figures from international finance, has recommended in its report on *Financial Reform: A Framework for Financial Stability*: “Countries should reevaluate their regulatory structures with a view to eliminating unnecessary overlaps and gaps in coverage and complexity, removing the potential for regulatory arbitrage, and improving regulatory coordination.”¹ This reevaluation has now begun, with the structure of US regulation being seriously re-examined for the first time in over a generation. Although the 1999 Gramm-Leach-Bliley Act dismantled the structural barriers between commercial and investment banking and between banking and insurance, it did not result in significant structural change to the complex and overlapping authorities of US regulatory agencies.² However, in March 2008, the Bush administration unveiled a plan for a major structural reform of regulation,³ while more recently the Obama administration has proposed a similar, but less radical reform, to Congress.⁴

The U.S. debate on regulatory structure has lagged behind in other countries of the Organisation for Economic Cooperation and Development (OECD) by over a decade.⁵ By the end of the last century many of these

¹ GROUP OF THIRTY, FINANCIAL REFORM: A FRAMEWORK FOR FINANCIAL STABILITY 10 (2009), available at <http://www.group30.org/pubs/recommendations.pdf>.

² Heidi Mandanis Schooner & Michael Taylor, *United Kingdom and United States Responses to the Regulatory Challenges of Modern Financial Markets*, 38 TEX. INT’L L.J. 317, 327-29 (2003).

³ U.S. DEP’T OF THE TREASURY, THE DEPARTMENT OF THE TREASURY BLUEPRINT FOR A MODERNIZED FINANCIAL REGULATORY STRUCTURE (2008), available at <http://www.treas.gov/press/releases/reports/Blueprint.pdf>.

⁴ U.S. DEP’T OF THE TREASURY, FINANCIAL REGULATORY REFORM, A NEW FOUNDATION: REBUILDING FINANCIAL SUPERVISION AND REGULATION (2009) available at http://www.financialstability.gov/docs/regs/FinalReport_web.pdf.

⁵ Schooner & Taylor, *supra* note 2, at 320. For a discussion of reform elsewhere in the OECD, *see id.* at 340-44.

countries had already embarked on major reorganizations of their institutional structures of financial regulation.⁶ These reform initiatives were presented as a response to the challenge of regulating today’s increasingly integrated financial markets in which the traditional distinctions between banking, securities, and insurance had become blurred.⁷ Moreover, with the dismantling of the structural regulations that had previously segmented the financial industry, diversified financial conglomerates had emerged, necessitating a group-wide perspective to ensure their effective regulation.⁸

Two broad approaches emerged in response to these challenges.⁹ The first, and most high profile, was the approach adopted in the United Kingdom that created a unified regulatory agency responsible for regulating all three of the main segments of the financial services industry for both financial soundness and consumer protection purposes.¹⁰ The alternative approach, which had originated in the U.K. but was not adopted there, was to structure regulation around two agencies, one responsible for the safety and soundness of all financial firms and the other for regulating their sales practices.¹¹ This “Twin Peaks” approach was adopted first in Australia and later in the Netherlands.¹² Variations of it are also to be found in Spain, France and Canada.

This essay attempts to distil some lessons for the United States from the U.K.’s reforms and especially the factors that led to the creation of a single, unified regulatory agency, the Financial Services Authority

⁶ *Id.* at 340-44.

⁷ *Id.* at 340.

⁸ *Id.* at 323.

⁹ See Richard K. Abrams & Michael W. Taylor, *Issues in the Unification of Financial Sector Supervision* 22-23 (Int’l Monetary Fund, Working Paper No. 00/213, 2000), available at <http://www.imf.org/external/pubs/ft/wp/2000/wp00213.pdf>. Within these two broad forms there is also scope for substantial variation. See *id.* at 21-24.

¹⁰ For a general review of the background to the U.K.’s reforms, see Eilis Ferran, *Examining the United Kingdom’s Experience in Adopting the Single Regulator Model*, 28 BROOK. J. INT’L L. 257 (2003).

¹¹ MICHAEL TAYLOR, “TWIN PEAKS”: A REGULATORY STRUCTURE FOR THE NEW CENTURY 10-11 (Ctr. for the Study of Financial Innovation) (1995).

¹² SELECT COMMITTEE ON ECONOMIC AFFAIRS, BANKING SUPERVISION AND REGULATION, 2008-9, H.L. 101-I, at 34.

(FSA).¹³ The radicalism of the U.K.'s approach should not be underestimated. Not only did it involve the merger of nine pre-existing regulatory agencies¹⁴ into one but, most controversially, it involved the decision to remove the responsibility for bank regulation from the Bank of England, the U.K.'s central bank, and to transfer it to the FSA.¹⁵ Although unified regulators had been previously created elsewhere, most notably in Scandinavia, none had involved the removal of bank regulation authority from the central bank.¹⁶

Critics of the U.K.'s arrangements at the time of the FSA's creation charged that the separation of bank regulation from the central bank's lender of last resort (LoLR) responsibilities was highly risky. It was argued that without the detailed institutional knowledge that derives from bank regulatory authority, the Bank of England would be unable to perform its LoLR responsibilities adequately. The subsequent experience of the run on the British mortgage bank Northern Rock in September 2007 seemed to confirm these critics. However, as this essay will argue, this conclusion overlooks the range of possible alternatives to the U.K.'s reforms, and particularly the Twin Peaks model. A "Twin Peaks" separation of prudential (safety and soundness) and consumer protection regulation would have offered a number of advantages over the FSA, including in relation to crisis management arrangements. The essay concludes by

¹³ For an assessment of the FSA more generally, see Howell E. Jackson, *An American Perspective on the U.K. Financial Services Authority: Politics, Goals & Regulatory Intensity* (Harvard, John M. Olin Ctr. for Law, Econ. & Bus., Discussion Paper No. 522, 2005), available at http://www.law.harvard.edu/programs/olin_center/papers/pdf/Jackson_522.pdf.

¹⁴ Clive Briault, *The Rationale for a Single National Financial Services Regulator* 6 (Fin. Servs. Auth., Occasional Paper No. 2, 1999), available at <http://www.fsa.gov.uk/pubs/occpapers/OP02.pdf>. These agencies include the Securities and Investments Board, the Personal Investment Authority, the Investment Management Regulatory Organisation, the Securities and Futures Authority, the Supervision and Surveillance Division of the Bank of England, the Building Societies Commission, the Insurance Directorate of the Department of Trade and Industry, the Friendly Societies Commission, and the Registrar of Friendly Societies. *Id.* at 6 n.1.

¹⁵ *Id.* at 7.

¹⁶ See Michael Taylor & Alex Fleming, *Integrated Financial Supervision: Lessons from Northern European Experience* 17 (World Bank, Working Paper 2223, 1999), available at http://info.worldbank.org/etools/docs/library/50180/TaylorFleming_1999.pdf.

drawing some conclusions from the British experience that might be considered by policy-makers in the U.S.¹⁷

II. THE DEBATE IN THE UK PRIOR TO THE FSA

What is striking about the policy debate within the U.K. prior to the formation of what became the FSA, is just how little attention was given to the possibility of creating a single integrated financial regulator. For several years prior to the election of a new Labour government in May 1997, there had been discussion of the need to reform the U.K.’s regulatory system, but the ideas being debated stopped short of proposing to create a single regulatory agency with a mandate that covered the entire banking, insurance and investment industries.¹⁸ The concept only came to prominence on May 20, 1997 with an announcement to the House of Commons by the new Chancellor of the Exchequer that the government intended to create a single regulatory authority for the banking and securities industries. The announcement itself came as a surprise to many observers and showed signs of having been rapidly prepared. This impression arose not only because the statement was vague concerning matters of detail, but also because it did not address some more fundamental issues, such as whether the prudential regulation of insurance companies would be included in the scope of the new regulator.¹⁹

Prior to this announcement, the British regulatory system combined institutional and functional regulation, similar to the system created by the

¹⁷ For further discussion regarding the Twin Peaks model in a U.S. context, see Heidi Mandanis Schooner, *Regulating Risk Not Function*, 66 U. CIN. L. REV. 441 (1998); Howell E. Jackson, *A Pragmatic Approach to the Phased Consolidation of Financial Regulation in the United States*, (Harvard Law Sch., Working Paper No. 09-19, 2008), available at <http://ssrn.com/abstract=1300431>; Cynthia Crawford Lichtenstein, *The Fed’s New Model of Supervision for “Large Complex Banking Organizations”*: *Coordinated Risk-Based Supervision of Financial Multinationals for International Financial Stability* (Boston College Law Sch. Research Paper No. 89, 2006), available at <http://ssrn.com/abstract=882474>.

¹⁸ Schooner & Taylor, *supra* note 2, at 320.

¹⁹ Ensuring the solvency of insurance companies had been the responsibility of the Department of Trade and Industry although it was briefly transferred to HM Treasury before the FSA was established. In July 1997, i.e. two months after the original announcement, the government confirmed that this function would also form part of the responsibilities of the new agency.

Gramm-Leach-Bliley Act in the United States.²⁰ Banks were regulated by the Bank of England under the Banking Act 1987²¹ with respect to their safety and soundness, while insurance companies were subject to solvency regulation under the Insurance Companies Act of 1982²² by a department of the Treasury (a function which was previously discharged by the Department of Trade and Industry). Sales practice (“conduct of business”) regulation was in the hands of a network of self-regulating organisations (SROs) which were also responsible for the safety and soundness regulation of non-bank financial intermediaries such as securities brokers and dealers and investment managers.

The SRO system was established by the Financial Services Act 1986²³ which had been described as “self-regulation within a statutory framework” both by its chief architect²⁴ and the Conservative government that enacted it. This system had been designed to provide an all-encompassing investor protection regime for securities, mutual funds, and other forms of collective investment through a number of “Self-Regulating Organizations” overseen by a quasi-governmental body, the Securities and Investments Board (SIB).²⁵ The SROs administered the sales practice regime and were responsible for ensuring that the users of financial services (generally speaking, securities brokering and dealing; futures brokering and dealing; investment management; financial advice; and sales practices relating to collective investment schemes like personal pensions and life insurance) were subject to appropriate levels of consumer protection. This system applied a functional approach to the regulation of investment services, products, and advice. If a service or product was being offered, it needed to be regulated by the relevant SRO, no matter what the nature of the firm offering the service.

The Financial Services Act was initially administered by no fewer than five separate SROs: The Securities Association (TSA) for Stock

²⁰ Schooner & Taylor, *supra* note 2, at 324-25.

²¹ Banking Act, 1987, c. 22, § 1 (Eng.) (repealed 2001).

²² Insurance Companies Act, 1982, c. 50, § 3 (Eng.) (repealed 2001).

²³ Financial Services Act, 1986, c. 60, § 8 (Eng.) (repealed 2001).

²⁴ L.C.B. Gower, “*Big Bang*” and *City Regulation*, 51 MOD. L. REV. 1, 11 (1988).

²⁵ The SIB exercised powers that were transferred to it under the Financial Services Act by the Secretary of State for Trade and Industry (a government minister). However, the SIB itself was in the unusual position of being a company limited by guarantee and not a department of government. A similar structure was subsequently adopted for the Financial Services Authority.

Exchange brokers and dealers; the Association of Futures Brokers and Dealers (AFBD) for dealers in futures and options; the Investment Management Regulatory Organisation (IMRO) for asset management and mutual funds; the Life Assurance and Unit Trust Regulatory Organisation (LAUTRO) for collective investment schemes marketed by insurance companies; and the Financial Intermediaries, Managers and Brokers Regulatory Association (FIMBRA) for independent financial advisers, many of whom acted as agents of the insurance companies.²⁶ During the later years of the self-regulatory system's existence some streamlining took place: the TSA and AFBD merged, as did LAUTRO and FIMBRA, thus reducing the number of SROs to three. Nonetheless the system was criticized for its complexity and opacity to the consumer, especially as it gave rise to what was described as an " 'alphabet soup' of regulatory agencies."²⁷ At the same time, the financial services industry criticised the system for not being genuinely self-regulatory, and for imposing an inappropriate regulatory burden on the interprofessional ("wholesale") money and capital markets. The SIB developed its own rulebook and required the SROs to adopt "equivalent" standards. This resulted in a lesser role for practitioner input and greater uniformity in the SRO rulebooks than had originally been intended.²⁸

However, the SRO system was most thoroughly discredited in the eyes of opposition lawmakers by what became known as the "pensions mis-selling scandal."²⁹ It had been the policy of the Conservative government in the mid-1980s to encourage more personal provision for retirement, rather than relying on occupational or state-provided schemes. Approximately eight million personal pensions were sold in the UK between 1988 and 1995.³⁰ The SRO system was intended in part to provide protection for individuals who entered into one of these personal savings schemes; in the words of John Major (then a junior minister but later

²⁶ See the account given by DAVID F. LOMAX, *LONDON MARKETS AFTER THE FINANCIAL SERVICES ACT 78 (1987)*.

²⁷ Taylor, *supra* note 11, at 7.

²⁸ These criticisms were recognized in a report issued by Andrew Large when he assumed the Chairmanship of the SIB in 1993. ANDREW LARGE, *FINANCIAL SERVICES REGULATION: MAKING THE TWO TIER SYSTEM WORK* (London: Securities and Investments Board, 1993).

²⁹ PETER CARTWRIGHT, *CONSUMER PROTECTION IN FINANCIAL SERVICES* 152 (1999).

³⁰ Michael Taylor, *The Policy Background in* BLACKSTONE'S GUIDE TO THE FINANCIAL SERVICES. & MARKETS ACT OF 2000 14 (Michael Blair ed., 2000).

Prime Minister) the Financial Services Act would “safeguard people against the unscrupulous overselling of personal pensions.”³¹

Personal pension plans were mainly offered by insurance companies (regulated by LAUTRO) which employed a sales staff with a large commission element in their remuneration. By 1993 it had emerged that a significant number of public sector employees, including teachers, nurses and the employees of former state-owned industries such as coal mining, had been encouraged by these salespeople to switch from their occupational schemes to personal pension plans. As the employer-provided plans were defined benefit, whereas the personal plans were defined contribution, this arguably placed these individuals at a potentially serious financial disadvantage.³² A report commissioned by the SIB suggested that as many as 1.5 million pensions had been mis-sold with compensation costs amounting to some £4 billion. Many of those affected were a core constituency of the Labour party – public sector workers – and hence the issue became highly politicised with the opposition party using it as a stick with which to beat the government.³³

Before winning the 1997 General Election Labour, party spokesmen had committed the party to end what they termed “City self-regulation.”³⁴ One of the few definite policy commitments to emerge from their pledge was the intention, once in government, to abolish the two-tier system of SIB and SROs.³⁵ In its place they undertook to establish a single, statutory regulatory agency for securities and investments. Thus the commitment to end City self-regulation might be narrowly construed as the commitment to replace the system created by the Financial Services Act.

At the same time, however, there were indications that the Labour party also considered the Bank of England to be part of the City’s “self-regulatory” system, even though its powers to regulate banks derived from

³¹ Michael Taylor, *Fin. Svcs. & Mkts. Act: The Policy Background-II*, 31 *AMICUS CURIAE* 4, 6 (2000).

³² Whether the individuals were disadvantaged and to what extent depended on a number of actuarial assumptions and assumptions about investment returns. The intricacies of these issues were, however, drowned out in the subsequent political debate.

³³ *See, e.g.*, 318, *PARL. DEB., H.C. (6th ser.)* (1998) 716, 718. (It continued to be used by Labour ministers against their Conservative opposite numbers even after the change of government).

³⁴ Mike O’Brien, *Labour’s Proposals for Regulation into the 21st Century*, 5 *J. FIN. REG. & COMPLIANCE* 115, 115-17 (1997).

³⁵ *LABOUR PARTY, LABOUR’S BUSINESS MANIFESTO* (1997).

a separate statute (the Banking Act 1987) and even though, unlike the SROs, it was a government agency. Labour suspicion of Britain’s central bank ran deep, with some commentators suggesting that it can be traced to the Bank’s role in the sterling crisis of 1931 that had helped to bring down a minority Labour administration headed by Ramsay Macdonald.³⁶ This fuelled Labour suspicions that the Bank of England was too closely aligned with the Conservative party, in which the financial interests of the City of London had a major influence. Thus when the British government considered the introduction of statute-based bank regulation in the mid-1970s some members of the governing Labour party proposed establishing a banking commission independent of the Bank of England to exercise regulatory powers. These proposals were rejected by the Cabinet after the then Governor of the Bank of England fought a rearguard action to ensure that it became the bank regulator.³⁷ Nonetheless, in subsequent years the Bank was to show itself a reluctant regulator which above all wished to maintain its traditional, informal relationship with the leading financial institutions in the City. Against this background it was possible to present it as part of the City’s “self-regulating” system and as merely the chief spokesman for a “cosy club.”

Nonetheless, it is doubtful whether the Bank of England’s responsibility for regulating the banking sector would have come under renewed scrutiny had it not been for two incidents in the first half of the 1990s. The first was the failure of the Bank of Credit and Commerce International (BCCI), which went into insolvent liquidation once it became clear that it had been a vehicle for a massive fraud.³⁸ Although the bank only had branches in the U.K. (its holding company was registered in Luxembourg), the group’s “mind and management” had been in London and hence there was a case for the Bank of England having taken the lead in ensuring that the group as a whole was subject to consolidated supervision. In a subsequent investigation conducted by Sir Thomas (later Lord Justice) Bingham, a senior judge, the Bank was found to have adopted

³⁶ MICHAEL MORAN, *THE POLITICS OF BANKING: THE STRANGE CASE OF COMPETITION & CREDIT CONTROL* 120 (2d ed., 1986).

³⁷ *Id.*

³⁸ H.M. STATIONARY OFFICE, *INQUIRY INTO THE SUPERVISION OF THE BANK OF CREDIT & COMMERCE INT’L.* (1992).

an excessively narrow interpretation of its powers and, partly for that reason, it had not been sufficiently proactive in regulating BCCI.³⁹

In BCCI's case, the Bank of England could argue in its defence that it was a bank regulator not a fraud investigator. No such defence was available in relation to the second episode – the failure of Barings merchant bank in early 1995.⁴⁰ Barings had been part of the City of London's "aristocracy," a centuries old merchant bank that had for generations been at the heart of the City's establishment to the extent of providing several Governors of the Bank of England. Barings had failed once before, in 1890, as the result of speculation in railroad construction in South America.⁴¹ It had been then bailed out by the Bank of England, at that time still a privately owned corporation.⁴² One hundred and five years later, Barings failed again, this time due to the poorly controlled activities of a futures trader based in Singapore who took large unhedged positions in the Singapore and Osaka futures exchanges.⁴³ This was the first of several episodes involving what came to be called "rogue traders" in the years that followed.⁴⁴ The episode was damaging to the Bank of England since it appeared that Barings had enjoyed a relatively light touch regulatory regime and thus provided an illustration of the operation of a so-called "self-regulatory system," at least as far as it applied to members of the City's establishment.⁴⁵

³⁹ 212 PARL. DEB., H.C. (6th ser.) (1992) 575. According to a statement given to the House of Commons by the then Chancellor of the Exchequer Norman Lamont, the Bingham report "argues that the Bank was slow to impose on BCCI an appropriate supervisory regime, and concludes that the Bank continued for too long to rely on the Luxembourg authorities to play the leading role." *Id.*

⁴⁰ H.M. STATIONARY OFFICE, REPORT OF THE BOARD OF BANKING SUPERVISION INQUIRY INTO THE CIRCUMSTANCES OF THE COLLAPSE OF BARINGS (1995).

⁴¹ See JOHN GAPPER & NICHOLAS DENTON, ALL THAT GLITTERS: THE FALL OF BARINGS 2 (1996).

⁴² *Id.*

⁴³ *Id.* at 28-29.

⁴⁴ A phrase that was originally coined to describe the Barings trader, Nick Leeson, which he used as the title of his subsequent book: NICK LEESON & EDWARD WHITLEY, ROGUE TRADER: HOW I BROUGHT DOWN BARINGS BANK AND SHOOK THE FINANCIAL WORLD (1996).

⁴⁵ See Gordon Brown, Ch. of the Exch., Statement to the H.C. on the Bank of Eng. (May 20, 1997), ("SIB will become the single regulator underpinned by statute. The current system of self-regulation will be replaced by a new and fully statutory system, which will put the public interest first, and increase public

Despite these episodes, Labour entered office in 1997 without a clear commitment to removing the Bank of England’s responsibility for bank regulation. Nor was there any indication of the possibility that a single financial regulator was on the policy agenda. What changed this situation was the new government’s announcement in its first few days in office that it would grant the Bank of England independence to set interest rates. Although this policy was not featured in the Labour party’s manifesto, central bank independence had been debated extensively in Britain since the early part of the decade.⁴⁶ On occasion in this debate the question of the central bank’s regulatory powers had arisen without, however, any definitive conclusion being reached. Nonetheless, once the decision was taken to create an independent central bank, a new Bank of England Act was required and this seems to have provided the pretext for a re-examination of the Bank’s role as bank supervisor.⁴⁷

The decision to remove banking supervision from the Bank of England appears to have been taken opportunistically. Before the start of each parliamentary year in Britain, each government department must put in “bids” for parliamentary time for the passage of legislation that it considers essential. The successful bids are then included in the government’s annual legislative program announced to parliament in the “Queen’s Speech.” In 1997 Treasury ministers wished to introduce two major bills – one to grant the Bank of England its independence, the other to abolish the “two tier” system of SIB and SROs created by the Financial Services Act. However, the new government had an ambitious policy agenda and a crowded legislative timetable, resulting in the Treasury being granted the time for only one major bill. According to the director of the Association of British Insurers, speaking the year after the event:

The Treasury team had failed to secure in the first Queen’s Speech legislation to abolish the two tier system under the Financial Services Act and Markets Act. However, a separate decision had been taken to give the Bank of England independence in respect of conducting monetary

confidence in the system.” (From the context it appeared that he considered the Bank of England to be part of the self-regulatory system). *Id.*

⁴⁶ See Michael Taylor, *Central Bank Independence.: The Policy Background*, in BLACKSTONE’S GUIDE TO THE FINANCIAL SERVICES AND MARKETS ACT OF 2000 at 10 (Michael Blair ed., 1998).

⁴⁷ See Ferran, *supra* note 10, at 271-72.

policy and this did require legislation. It seems that an opportunist decision was taken at this stage to move towards a single regulator because the legislation to give the Bank of England independence in respect of monetary policy could be used for any other purpose relevant to the Bank of England.⁴⁸

One of these “other purposes” was the transfer of responsibility for bank regulation from the Bank of England to the SIB, which then became the nucleus of the FSA. In other words, the parliamentary timetable rather than a reasoned policy debate seems to have triggered the decision to move to a single regulator. This also would have been consistent with an apparent about-turn in government policy after the Governor of the Bank of England apparently had been assured there were no immediate plans to strip the Bank of its bank regulatory function.

There have also been allegations that the concept of a single financial regulator had been developed within the Treasury before the change of government and had been inspired as much by Treasury rivalry with the Bank as by any policy considerations.⁴⁹ It was certainly the case that government ministers saw a single financial services regulator as an alternative centre of power to the Bank and hoped that the FSA would assume the Bank’s role as overseer of the City’s interests. Since the government was committed to establishing Bank of England independence in respect of monetary policy, it is also possible that removing its banking supervision function was seen as a way of preventing it from becoming “an over-mighty subject.” Whatever the exact motivation, it is clear that the momentous implications of an opportunistic and essentially political decision may not have been fully appreciated by government ministers who were still new to power after their party’s unusually long period out of office.⁵⁰

⁴⁸ See Ferran, *supra* note 10, at 271.

⁴⁹ See SIR MARTIN JACOMB, RE-EMPOWER THE BANK OF ENG. 2-4 (Centre for Policy Stud.) (2009) available at http://www.cps.org.uk/cps_catalog/Re-empower%20the%20Bank%20of%20England.pdf.

⁵⁰ It is important in this regard that the Labour party had been out of power for 18 years and few of its lawmakers had experience of government.

III. JUSTIFYING THE SINGLE REGULATOR

The decision to create a single financial regulator had to be justified after the fact. Government ministers and the FSA itself put forward a series of justifications for the creation of a single regulator. They fell into two broad categories: those relating to market developments and those relating to the purported effectiveness and efficiency of a single regulatory agency.

The argument that market developments justified a single financial regulator became known as the "blurring the boundaries" argument. In his statement to the House of Commons on May 20, 1997, Britain's Chancellor of the Exchequer argued that:

At the same time, it is clear that the distinctions between different types of financial institution--banks, securities firms and insurance companies--are becoming increasingly blurred. Many of today's financial institutions are regulated by a plethora of different supervisors. This increases the cost and reduces the effectiveness of the supervision.

There is therefore a strong case in principle for bringing the regulation of banking, securities and insurance together under one roof. Firms now organise and manage their businesses on a group-wide basis. Regulators need to look at them in a consistent way. That would bring the regulatory structure closer into line with today's increasingly integrated financial markets. It would deliver more effective and efficient supervision, giving both firms and customers better value for money, and would improve the competitiveness of the sector and create a regulatory regime to genuinely meet the challenges of the 21st century.⁵¹

The argument was further developed in a document issued by the Treasury the following year:

The existing arrangements for financial regulation involve a large number of regulators, each responsible for different

⁵¹ 294 PARL. DEB., H.C. (6th ser.) (1997) 510.

parts of the industry. In recent years there has been a blurring of the distinctions between different kinds of financial services business: banks, building societies, investment firms, insurance companies and others. This has added further to the complexity of financial regulation. The Government believes the current system is costly, inefficient and confusing for both regulated firms and their customers. It is not delivering a standard of supervision and investor protection that the public has a right to expect. We are therefore establishing a single, statutory regulator for the UK financial services industry with clearly defined regulatory objectives and a single set of coherent functions and powers.⁵²

However, it was left to the FSA itself to provide the most extensive justification for its own existence. While the FSA was still under construction, it published a paper written by one of its own officials, Clive Briault, who set out to defend the single regulator concept.⁵³ He did so by first invoking the “blurring of boundaries” argument:

The securitisation of traditional forms of credit (including mortgages, credit card outstandings and commercial loans) and, with the growth of options, increasingly elaborate ways of unbundling, repackaging and trading risks, have weakened the distinction between equity, debt and loans, and even between banking and insurance business (where, for example, credit derivatives bear many of the characteristics of an insurance product and insurance companies offer short-term deposit-like products).⁵⁴

This development, Briault explained, had an important consumer protection dimension in that the disappearance of a neat conjunction between a particular type of firm and a limited range of products being supplied by that firm means that it is difficult to regulate on a functional basis, since the

⁵² H.M. STATIONARY OFFICE, FIN. SVCS. & MKTS. BILL: A CONSULTATION DOCUMENT, pt. 1 (1998).

⁵³ See Briault, *supra* note 14.

⁵⁴ *Id.* at 13-14.

traditional functional approach no longer matches the structure of either firms or markets.⁵⁵

Accordingly, a single financial regulator was essential to provide adequate consumer protection when financial products could no longer be neatly slotted into the traditional contractual forms which have underpinned the functional approach to regulation.⁵⁶ Trying to regulate the sale and marketing of products on a functional basis would result in inadequate consumer protection, either because similar products would become subject to different levels of consumer protection or the regulatory agencies disputed jurisdiction over certain types of product.⁵⁷

The blurring the boundaries argument also related to the formation of financial conglomerate groups. The emergence of financial conglomerates (usually defined as a group which undertakes at least two of the activities of banking, securities or insurance) resulted from mergers and acquisitions that occurred most frequently between banks and securities firms and between banks and insurance companies.⁵⁸ In some cases they also involved the purchase of fund managers by banks and by insurance companies.⁵⁹ These combinations were permitted as the result of the dismantling of structural barriers – which in the U.K. had been mainly informal and non-statute based – in the course of the 1980s.⁶⁰ In response to these and similar developments elsewhere in the G10, the Tripartite Group of banking, securities, and insurance supervisors argued in a 1995 report that a “group-wide” perspective was required to obtain an adequate supervisory overview of these financial conglomerates.⁶¹ Nonetheless, as long as regulation remained structured along traditional institutional/functional lines, obtaining such a group-wide perspective would be difficult.

The British solution was to adopt the lead regulator concept.⁶² The lead regulator would be responsible for taking a consolidated view of the

⁵⁵ *Id.* at 14.

⁵⁶ *See id.*

⁵⁷ *See id.*

⁵⁸ *Id.* at 12-13.

⁵⁹ Briault, *supra*, note 14, at 13.

⁶⁰ TRIPARTITE GROUP OF BANK SEC. & INS. REG., THE SUPERVISION OF FIN. CONGLOMERATES at i (1995). Subsequently the Tripartite Group was renamed the Joint Forum.

⁶¹ *Id.* at i-ii.

⁶² *See* GEORGE ALEXANDER WALKER, INT’L BANKING REGULATION: LAW, POLICY AND PRACTICE 258 (Kluwer Law Int’l) (2001).

capital adequacy and liquidity of the consolidated group; taking a similarly group-wide view of more qualitative factors such as the calibre of senior management and the high-level systems and controls of the financial conglomerate; and co-ordinating and encouraging the exchange of information among the relevant regulatory bodies, both routinely and in the event of an emergency. Typically, since most such groups were headed by a bank, the Bank of England usually assumed this responsibility, which was similar to the Fed's umbrella supervisor role created by Gramm-Leach-Bliley.⁶³ In contrast to the U.S. arrangements, however, the Bank of England's role was largely extra-statutory and was the result of a framework of Memorandums of Understanding (MoUs) between the Bank of England and the functional regulators.

Although Briault claimed that the lead regulator concept had worked well, he nonetheless stressed that countries that had moved towards a single regulator had "done so in part because, with the growth in the number of multiple-function firms, the need for communication, co-ordination, co-operation and consistency across specialist regulatory bodies had become increasingly acute and increasingly difficult to manage efficiently."⁶⁴ If such firms were the rule rather than the exception (in contrast to the situation in the past) then new institutional arrangements were required to ensure that they were subject to more efficient oversight. Briault cited statistics to show that many firms were now subject to multiple regulators: eight firms (including HSBC, Halifax, Abbey National and the Royal Bank of Scotland) were authorised to conduct all five of the main regulated activities ("deposit-taking, insurance, securities and corporate finance, fund management, and advising on or selling investment products to retail customers").⁶⁵ A further 13 firms were authorised to conduct four of these activities, and more than 50 other firms were authorised for three of these five functions.⁶⁶

The efficient supervision of financial conglomerates was only one dimension of the superior efficiency claimed for the single regulator. It was also argued that it would allow scarce supervisory resources to be deployed more effectively; an example concerned the development of specialist teams to review firms' internal risk management models that had become an integral part of regulation during the 1990s. In the pre-FSA

⁶³ Schooner & Taylor, *supra* note 2.

⁶⁴ Briault, *supra* note 14, at 15.

⁶⁵ Briault, *supra* note 14, at 13.

⁶⁶ *Id.*

system, several different regulators had needed to build their own specialist model review teams, but individuals with the requisite skills were in high demand which made it difficult for regulatory agencies to recruit them in sufficient numbers.⁶⁷ By centralizing the available resources, a single regulator seemed to offer a way out of this impasse. Similarly, it was also argued that the creation of a single support infrastructure (e.g. IT system) would lead to significant cost savings as the duplication and overlap resulting from the nine pre-existing regulators was eliminated. The argument that a single regulator would be more cost effective was vital in selling the concept to the financial industry. It was therefore not surprising that Briault made much of this argument:

Economies of scale and scope should arise because a single regulator can take advantage of a single set of central support services (human resources, information services, financial control, premises etc); introduce a unified statistical reporting system for regulated firms; operate a single database for the authorisation of firms and the approval/registration of individuals; avoid unnecessary duplication or underlap across multiple specialised regulators; introduce a consolidated set of rules and guidance; tackle problems of co-ordination, co-operation and communication more effectively within a single entity and under a unified management structure than might be possible across separate specialist entities; offer a single point of contact to both regulated firms and to consumers (through a single complaints handling regime and a single compensation scheme); and adopt a more effective and focused approach to areas of common interest to most regulated financial activities (for example, handling Year 2000 issues and turbulence in international financial markets).⁶⁸

These arguments – consumer protection arrangements that were better suited to the characteristics of new financial instruments, improved oversight of financial conglomerate groups, and cost savings and efficiencies from a common regulatory platform – were at the heart of the

⁶⁷ See Taylor, *supra* note 11, at 6.

⁶⁸ Briault, *supra* note 14, at 18.

case constructed for the single financial regulator. The difficulty was that exactly these same arguments had been made in favor of an alternative regulatory structure – the so-called Twin Peaks model. It was therefore also necessary for the defenders of the single regulator to explain why this structure would be superior to the Twin Peaks alternative.

IV. THE REJECTION OF THE TWIN PEAKS ALTERNATIVE

Unlike the single regulator, a Twin Peaks structure had been actively debated in the U.K. prior to the 1997 reform, and it had attracted support from a number of influential figures both in the industry and in regulation.⁶⁹ It was, however, strongly opposed by the Bank of England which regarded the proposals as primarily an attempt to divest it of its regulatory responsibilities.

Twin Peaks proposed that, instead of being structured around the traditional tripartite distinction of banking, securities and insurance, the institutional structure of regulation should in future comprise two regulatory agencies, a Financial Stability Commission and a Consumer Protection Commission.⁷⁰ The first would be responsible for ensuring the stability of the financial system as a whole, mainly through the application of prudential regulations.⁷¹ The second would be charged with ensuring that firms deal with their (retail) customers in a fair and transparent manner.⁷² The two Commissions would be responsible for discharging their mandate irrespective of the legal form of the firms that they regulated.⁷³

The source of the “blurring the boundaries” argument can be traced to the Twin Peaks proposals, which also placed heavy emphasis on the need to ensure proper group-wide supervision of financial conglomerates.⁷⁴

⁶⁹ See Jill Treanor, *Regulators Back Taylor's Twin Peaks Theory*, THE INDEP., Oct. 29, 1996, available at <http://www.independent.co.uk/news/business/regulators-back-taylors-twinpeaks-theory-1360780.html>.

⁷⁰ Taylor, *supra* note 11, at 1.

⁷¹ *Id.*

⁷² *Id.*

⁷³ *See id.*

⁷⁴ *Id.* at 4. However, the earliest occurrence of this argument is to be found in a working paper published the year before Twin Peaks. See Claudio E.V. Borrio & Renata Filosa, *The Changing Borders of Banking: Trends and Implications*, 3, 16 (Bank for Int'l Settlements, Working Paper No. 23, 1994), available at <http://www.bis.org/publ/work23.pdf?noframes=1>.

The case for Twin Peaks also invoked the economies of scale that would result from (the admittedly more limited) regulatory consolidation that it also involved. Thus, because the arguments for a single regulator and for Twin Peaks were almost identical, it was necessary for the FSA’s defenders to show that theirs was the superior solution. The crux of the argument concerned the separation of prudential and conduct of business regulation that was the main feature of the Twin Peaks model; the defenders of a single regulator argued that the separation was not so clear cut as the Twin Peaks model presupposed.⁷⁵

The first strand of this argument was to contest the claim, central to the Twin Peaks analysis, that there were two relatively, clearly distinguishable regulatory objectives – financial stability on the one hand and consumer protection on the other. This case for treating these two objectives as interlinked is well summarized by Davies and Green:

The ultimate argument for financially sound and prudentially well regulated financial institutions is that they are then able to provide financial services and investment opportunities to consumer and businesses which those customers may use with confidence. A breakdown in consumer protection, whether in banking, investment or insurance products, may itself precipitate a wider loss of confidence in types of product or firms. There is therefore no necessary conflict between the two aims of regulation. In the long run they are aligned.⁷⁶

Closely related to this was the claim further claim that, in practice, prudential and conduct of business (sales practice) regulation required examination of very similar issues, and therefore that there would be significant overlap between the Twin Peaks agencies.⁷⁷ Briault put the point with characteristic clarity:

[T]here is a considerable overlap – both conceptually and in practice – between prudential and conduct of business regulation. Both have a close and

⁷⁵ See Briault, *supra* note 14, at 25.

⁷⁶ HOWARD DAVIES & DAVID GREEN, GLOBAL FIN. REGULATION: THE ESSENTIAL GUIDE 192 (2008).

⁷⁷ Briault, *supra* note 14, at 25.

legitimate interest in the senior management of any financial institution subject to both of these types of regulation, in particular because of the crucial roles of senior management in setting the “compliance culture” of a firm, in ensuring that management responsibilities are properly allocated and cover comprehensively the business of the firm, and in ensuring that other internal systems and controls are in place. The detail of some of these systems and controls may indeed be specific to either prudential or conduct of business considerations, but many of them will be more general.⁷⁸

In short, a single regulator was superior to a Twin Peaks structure because many of the same supervisory judgments would arise in considering prudential and sales practice regulation. There seemed little point in having two regulators reaching essentially duplicate judgments of broadly similar matters. Since there is substantial overlap between the two regulatory objectives and, in practice, prudential and conduct of business regulation will focus on the same fundamental issues, they were best administered by a single regulatory agency.

The Global Financial Crisis has created a very different perspective on this argument. The British Government’s own White Paper on regulatory reform after the crisis has concluded that the system places too much weight on “ensuring that systems and processes were correctly defined rather than on challenging business models and strategies” and on “conduct of business regulation of the banking sector rather than prudential regulation of banking institutions.”⁷⁹ Even the FSA’s senior management has acknowledged that the agency neglected prudential supervision.⁸⁰ In the words of the report on the banking crisis produced by the FSA’s current chairman, Lord Turner, the agency’s regulatory practices resulted in “[a] balance between conduct of business and prudential regulation which, with the benefit of hindsight, now appears biased towards the former.”⁸¹ Turner repeated this admission to a committee of the British House of Lords which

⁷⁸ *Id.*

⁷⁹ H.M. TREASURY, REFORMING FINANCIAL MARKETS, 2009, Cm. 7667, at 56.

⁸⁰ See FIN. SERV. AUTH., THE TURNER REVIEW: A REGULATORY RESPONSE TO THE GLOBAL BANKING CRISIS 87 (2009), available at http://www.fsa.gov.uk/pubs/other/turner_review.pdf.

⁸¹ *Id.*

referred in its final report to the “widely held perception that, in recent years, the FSA has emphasized conduct-of-business supervision at the expense of prudential supervision.”⁸²

This situation was especially apparent in the FSA’s supervision of the mortgage bank Northern Rock which was the first British casualty of the crisis.⁸³ The bank had received numerous contacts from the FSA concerning a consumer protection initiative (“treating customers fairly”), but supervision of capital and liquidity had been deficient and the bank had been placed in a category that subjected it to one major prudential meeting once every three years. The FSA’s own report on Northern Rock stated that “some of the fundamentals of work on assessing risks in firms (notably some of the core elements related to prudential supervision, such as liquidity) have been squeezed out.”⁸⁴

The House of Lords Committee thought it could identify the reason why the FSA had failed to give sufficient attention to prudential regulation:

Conduct-of-business is important and politically sensitive, and its results are easy to measure. In contrast, prudential supervision, while arguably more important, is conducted privately; its success is less easily measured, and, most of the time, it has a lower political impact than conduct-of-business supervision though in times of crisis such as the present its political impact, its effect on businesses, individuals and the economy, is very much greater than conduct-of-business supervision. It is natural and rational for a supervisor with responsibility for both activities to concentrate on the one with the greater immediate political sensitivity.⁸⁵

In other words, the argument that there were synergies between prudential and conduct of business regulation overlooked the distinct possibility that

⁸² SELECT COMMITTEE ON ECONOMIC AFFAIRS, BANKING SUPERVISION AND REGULATION, 2008-09, H.L. 101-I, at 33.

⁸³ FINANCIAL SERVICES AUTHORITY, THE SUPERVISION OF NORTHERN ROCK: A LESSONS LEARNED REVIEW ¶¶ 8, 27 (2008), *available at* http://www.fsa.gov.uk/pubs/other/nr_report.pdf.

⁸⁴ *Id.* at ¶ 36.

⁸⁵ SELECT COMMITTEE ON ECONOMIC AFFAIRS, BANKING SUPERVISION AND REGULATION, 2008-09, H.L. 101-I, at 33.

one type of regulation would come to dominate within a single regulator and that this would likely to be consumer protection given the realities of the political process.⁸⁶ *Twin Peaks* had predicted that this outcome was likely, and used it as one of the arguments against creating a single regulatory agency.⁸⁷

Thus the argument that there was a natural synergy between prudential and consumer protection regulation has been discredited by events both before and during the crisis. If the purported synergies were really as strong as was claimed, then the multiple reviews of Northern Rock's systems for handling consumer issues should have thrown out evidence that the bank's business strategy was dangerously flawed. They did not. Nor is this outcome really surprising. Although there may be some overlap of the relevant judgments at the margin, they ultimately involve quite fundamentally different matters. Weaknesses of internal control systems for dealing with consumer issues may be indicative of more general weaknesses in internal control within the institution as a whole, and this could indeed raise matters of prudential concern. But it is doubtful that these kinds of findings will demonstrate that the bank's management is following a deeply flawed and highly risky business strategy which is likely to end in failure. To reach this conclusion it is necessary to ask different questions to those a consumer protection regulator might ask.

V. THE ROLE OF THE CENTRAL BANK

One dimension of the *Twin Peaks* structure that had been actively debated before the decision to create the FSA was the role of the central bank. In a number of speeches and articles, the Bank of England's senior management defended the Bank's role as a bank regulator against the proposed *Twin Peaks* structure.⁸⁸ The central bank, it was argued, needed to be concerned with the financial condition of the banking system, as this was the conduit through which its monetary policy was transmitted to the wider economy. As Governor Eddie George argued in a speech given in 1994, before the *Twin Peaks* debate began, the soundness of banks and the central bank's ability to conduct monetary policy were intimately related:

⁸⁶ *See id.*

⁸⁷ Taylor, *supra* note 11, at 15.

⁸⁸ Howard Davies, *Financial Regulation: Why, How and By Whom*, in BANK OF ENG. Q'LY BULL. 107, 111 (1997).

Monetary and financial stability are inter-related. It is inconceivable that the monetary authorities could quietly pursue their stability-oriented monetary policy objectives if the financial system through which policy is carried on – and which provides the link with the real economy – were collapsing around their ears. The liabilities of banks in particular are money, and you cannot be concerned with preserving the value of money without being concerned also with preserving public confidence in money in this broader sense. Equally though, the financial system is much less likely to be collapsing around the ears of the monetary authorities in an environment of macro-economic stability than in one of exaggerated boom and bust and volatile asset values. This inter-relationship means that, whatever the precise institutional arrangements for financial regulation and supervision, central banks necessarily have a vital interest in the soundness of the financial system.⁸⁹

Moreover, banks were a “special” type of financial intermediary: as Sir Howard Davies, at that time still the Deputy Governor of the Bank of England,⁹⁰ said in early 1997 “in our view, there is still a reasonably clear distinction to be made between banks and other financial institutions, and their prudential soundness, or lack of it, can have rather different implications for the rest of the market.” As a result, he continued,

Of course it may be argued that the distinctive characteristics of banks, and their potential to create systemic risk—which central banks can counteract—does not necessarily mean that the central bank should act as their regulator. I agree. But there are significant synergies to be had from maintaining an institutional link between the two functions, and the burden of proof rests, I think,

⁸⁹ E. A. J. George, Governor, Bank of England, *The Bank of England – Objectives and Activities*, The Capital Market Research Institute, Frankfurt, Germany (Dec. 5, 1994).

⁹⁰ Shortly afterwards he was appointed the first Chairman and Chief Executive of the FSA. His views on the specialness of banks underwent a subsequent change.

with those who wish to make the case for disturbing that relationship.⁹¹

The main “synergy” that arose from retaining banking supervision within the central bank was with the Bank’s role as lender of last resort (LoLR). It was argued that the information acquired in the capacity of the bank supervisor was essential to the central bank performing the lender of last resort function, and that therefore the best arrangement was for LoLR and banking supervision to be located in the same institution. Following the Northern Rock experience, a number of commentators have reached the conclusion that this argument was correct. As Professor Willem Buiter argued in evidence to the House of Commons Treasury Select Committee:

The notion that the institution that has the knowledge of the individual banks that may or may not be in trouble would be a different institution from the one that has the money, the resources, to act upon the observation that a particular bank needs lender of last resort support is risky. It is possible, if you are lucky, to manage it, but it is an invitation to disaster, to delay, and to wrong decisions. The key implication of that is that the same institution—it could be the FSA or it could be the Bank of England—should have both the individual, specific information and the money to do something about it.⁹²

Against these arguments, proponents of separation argued that theoretical considerations and empirical evidence indicated that central banks with banking supervisory responsibilities tended to err on the side of laxity in monetary policy; as Goodhart and Schoenmaker argued in a widely cited paper, monetary policy aimed to be countercyclical, whereas regulatory policy was pro-cyclical.⁹³ Concerns were also expressed that banking supervision “failures” – which it was generally accepted were

⁹¹ DAVIES, *supra* note 88, at 110.

⁹² *Id.*

⁹³ C. A. E. Goodhart & D. Schoenmaker, *The Institutional Separation Between Supervisory and Monetary Agencies*, in *THE CENTRAL BANK AND THE FINANCIAL SYSTEM* 341 (1995)..

almost inevitable – would damage the reputation and credibility of the central bank as a monetary policy institution.⁹⁴

Twin Peaks further argued that “[a]s the once-clear demarcation lines between types of financial markets and institutions are broken down, the Bank’s role appears increasingly anomalous.”⁹⁵ In other words, owing to the changing nature of the financial system, banks could no longer be considered the unique source of systemic risk that traditionalists insisted that they remain. In consequence of these developments, it became increasingly difficult to argue that banks were “special” in the sense that they were uniquely, systemically important.⁹⁶ Many large non-banks were now “too interconnected to fail,” a phrase that was coined when Bear Stearns teetered on the brink of failure in March 2008.⁹⁷ On the traditional view, Bear Stearns would not have been considered systemically important; however, the episode confirmed the argument of *Twin Peaks* that “the rise of the OTC markets means that we must extend our concept of what constitutes a systemically important firm.”⁹⁸

Yet if the concept of a systemically important firm was extended in this way, it was by no means obvious that the central bank was the right institution to regulate these firms. Twin Peaks acknowledged the possibility that the functions of the FSC could be performed by the central bank and that LoLR was an important issue.⁹⁹ However, on balance it rejected the case for the central bank also performing the role of prudential regulator of the new, broader category of systemically important firms.¹⁰⁰ In the first place, a broader concept of systemically important firms meant that the central bank would need to interact with various institutions that were not its traditional counterparties (a prediction that has come to pass following the Federal Reserve’s expansion of its facilities in the wake of the financial crisis).¹⁰¹ Secondly, the expertise necessary to regulate investment banks and insurance companies does not naturally reside in central banks.¹⁰² As *Twin Peaks* identified, a major problem for central banks is in finding a place for such regulatory specialists in organizations

⁹⁴ *Id.* at 341

⁹⁵ Taylor, *supra* note 11, at 13-14.

⁹⁶ *Id.* at 4.

⁹⁷ *Id.*

⁹⁸ *Id.* at 5.

⁹⁹ *Id.* at 14.

¹⁰⁰ *Id.* at 13-14.

¹⁰¹ Taylor, *supra* note 11, at 5.

¹⁰² *Id.* at 6, 12.

where they will have few opportunities for career progression.¹⁰³ Nonetheless, Twin Peaks also recognized that close links would be needed between the central bank and the FSC.¹⁰⁴ Although it was comparatively sketchy about the nature of those links, apart from proposing overlapping board membership, the need for close coordination between the central bank and the prudential regulator was an important component of the Twin Peaks structure.¹⁰⁵

The FSA's relationship with the Bank of England was, in theory, also to be a close one.¹⁰⁶ Yet when the FSA was established, very little attention was given to the need for institutional linkages between the regulator and the central bank.¹⁰⁷ Instead, given the prominent role played by ex-Bank of England staff in the FSA, the professional relationships between former colleagues were supposed to guarantee cooperation between the two institutions. However, once this generation of officials had either retired or left the FSA, there was no institutional mechanism to ensure close collaboration between the two institutions. More recently, the British government has announced the formation of a Financial Stability Council which can be seen as a belated attempt to build the stronger institutional linkages between the Bank and the FSA that were required from the outset.

Briault acknowledged that in the "multi-faceted" relationship between Bank and FSA, close cooperation and regular information flows would be essential. These would need to occur both routinely for those aspects of financial stability in which the central bank has an interest for the setting of monetary policy, and in exceptional circumstances for more specific and detailed information relating to the position of financial institutions for whom support operations are being considered (where the fiscal authority is also likely to have a close interest). "The UK Memorandum of Understanding... provides an important underpinning to the necessary exchange of such information under the new arrangements in the UK."¹⁰⁸

The Memorandum of Understanding to which Briault refers was between the Treasury, Bank of England, and FSA and it supposedly created

¹⁰³ *Id.* at 12.

¹⁰⁴ *Id.* at 14.

¹⁰⁵ *Id.*

¹⁰⁶ *Id.* at 13-14.

¹⁰⁷ Taylor, *supra* note 11, at 13.

¹⁰⁸ Briault, *supra* note 14, at 33.

the framework for both information exchange and for crisis management.¹⁰⁹ These are referred to under the MoU as the “Tripartite Authorities” and the Bank of England’s responsibilities are summarised as contributing “to the maintenance of the stability of the financial system as a whole.”¹¹⁰ The FSA has the responsibility of authorising and supervising individual banks.¹¹¹ HM Treasury has responsibility for the institutional structure of the financial regulatory system, and the legislation behind it.¹¹² In a crisis, the Financial Services Authority would, according to the Memorandum of Understanding, be responsible for monitoring “the health of institutions that fall within its regulatory remit” and for ensuring, “as far as is appropriate in the circumstances, continuing compliance with regulatory standards.”¹¹³ However, the Bank of England would remain in charge of “official financial operations ... in order to limit the risk of problems in or affecting particular institutions spreading to other parts of the financial system”.¹¹⁴

The MoU also established a Joint Crisis Management Committee, chaired by the Chancellor, for dealing with what the MoU referred to generically as “support operations.”¹¹⁵ It did not, however, clearly distinguish between those operations that relate to emergency liquidity assistance and those that would involve solvency support.¹¹⁶ In both cases the Treasury sat at the apex of a pyramid with both the Bank and FSA in subordinate roles.¹¹⁷ This contrasts with the practice of most other countries in crisis management, which is to ensure that as long as the issue remains one of liquidity the central bank will be in the lead.¹¹⁸ It alone has (or should have) the information and the ability to react sufficiently

¹⁰⁹ GORDON BROWN ET AL., MEMORANDUM OF UNDERSTANDING BETWEEN H.M. TREASURY, THE BANK OF ENGLAND AND THE FINANCIAL SERVICES AUTHORITY 4-5 (2009), *available at* <http://www.bankofengland.co.uk/financialstability/mou.pdf> [hereinafter MoU].

¹¹⁰ *Id.* ¶ 2.

¹¹¹ *Id.* ¶ 3(i).

¹¹² *Id.* ¶ 4(i).

¹¹³ *Id.* ¶ 17(iii).

¹¹⁴ *Id.* ¶ 2(iv).

¹¹⁵ GORDON BROWN ET AL., *supra* note 109, at ¶ 14.

¹¹⁶ *Id.* ¶ 17(iii).

¹¹⁷ *Id.* at ¶ 4, 10, 13.

¹¹⁸ Dong He, “Emergency Liquidity Support Facilities”, Appendix in CHARLES ENOCH et al (eds) BUILDING STRONG BANKS THROUGH SURVEILLANCE AND RESOLUTION (IMF 2002), p.134

promptly to emerging problems.¹¹⁹ In this case the FSA's role would be clearly established as that of a handmaiden to the Bank, under an explicit obligation to provide it with any and all information required by for the discharge of its duties.¹²⁰ Only in the event that the issue becomes one of providing solvency support should the Treasury have taken the lead, with both the Bank and the FSA in supporting roles.¹²¹ The subordinate role to which the Bank was assigned in the MoU provides support to those who argue that the post-1997 arrangements were designed to reduce the Bank's status.

In practice, however, the arrangements envisaged by the MoU were rarely tried in practice and the Joint Crisis Management Committee rarely met. The House of Commons Treasury Select Committee, in reviewing the Northern Rock experience, concluded that "in terms of information exchange between the Tripartite authorities, the system might have ensured that all the Tripartite authorities were fully informed. However, for a run on a bank to have occurred in the United Kingdom is unacceptable, and represents a significant failure of the Tripartite system."¹²²

VI. LESSONS OF THE BRITISH EXPERIENCE FOR UNITED STATES REGULATORY REFORM

Reviewing the lessons of the British experience, and considering the current regulatory reform debate in the United States, I offer the following conclusions:

A. THE CONCEPT OF A SINGLE REGULATOR HAS NOT BEEN DISCREDITED, BUT ITS LIMITATIONS HAVE BEEN EXPOSED

The U.K. was not the first country to establish a unified regulatory agency outside the central bank: that honor belongs to the Scandinavian countries, with Norway (1986) as the pioneer followed by Denmark (1988) and Sweden (1991).¹²³ In these countries an important consideration was

¹¹⁹ *Id.* at ¶ 2, 4.

¹²⁰ *Id.* at ¶ 6.

¹²¹ GORDON BROWN ET AL., *supra* note 109, at ¶ 14.

¹²² HOUSE OF COMMONS TREASURY COMMITTEE, *THE RUN ON THE ROCK*, 2007-08, H.C. 56-1, at 107.

¹²³ Taylor & Fleming, *supra* note 16, at 6-7.

the “economies of scale” argument.¹²⁴ As relatively small countries with financial systems dominated by a small number of financial conglomerate groups, combining all regulatory functions within a single agency appeared to offer numerous efficiency benefits.¹²⁵ Moreover, since the central bank was not involved in banking supervision in any of these countries, the powerful – and sometimes emotive – issue of the central bank’s powers did not arise.¹²⁶

There continues to be a case for single regulatory agencies in comparatively small countries where the economies of scale gains are significant.¹²⁷ It is expensive to establish regulatory agencies with their associated support services and infrastructure, and therefore minimizing overhead costs is a worthwhile ambition. However, in larger countries, especially those with a large and complex financial system, any potential efficiency gains are far outweighed by the inefficiencies of combining too many regulatory functions in a single agency. As noted above, the FSA has struggled with the combination of prudential and conduct of business regulation and its past performance suggests that it was simply tasked with too many functions to perform all of them adequately.

B. PRUDENTIAL AND CONDUCT OF BUSINESS REGULATION DON’T MIX

Despite the claims of the FSA’s supporters that there are substantial synergies between prudential and conduct of business regulation, the crisis has shown the limits of these synergies. While some of the relevant supervisory judgements do overlap, especially on such matters as internal controls and the probity of management, prudential regulation needs a different focus. The factors influencing the financial soundness of an institution and the likelihood that it might fail go far beyond those of concern to a consumer protection regulator. Moreover, as the House of Lords Select Committee on Economic Affairs observed,

There is also a cultural difference between conduct-of-business and prudential supervision. Conduct-of business supervision is often performed by lawyers. Prudential

¹²⁴ *Id.* at 25.

¹²⁵ *See id.*

¹²⁶ *Id.* at 17.

¹²⁷ *Id.* at 31.

supervision is largely an economic activity, particularly at the macro level. It seems likely that either a lawyerly or an economic approach would dominate in a supervisory body that performed both prudential and conduct of business supervision, and that this dominance would reduce the effectiveness of the dominated half of the organisation.¹²⁸

The function that receives the greatest emphasis will be that having the greatest political saliency: this means that in normal times, when bank failures are rare, consumer protection regulation is likely to be the main focus of agency attention. Although the FSA has now increased the resources it devotes to prudential regulation,¹²⁹ the above analysis suggests that this is likely to be a relatively short term development, remaining in place only as long as political attention is focused on the fall-out from the crisis.

C. IT IS ESSENTIAL TO ACHIEVE A BALANCE BETWEEN THE FINANCIAL STABILITY AND CONSUMER PROTECTION OBJECTIVES

Because of the circumstances in which the U.K.'s regulatory reforms took place – as a reaction to perceived regulatory failures in consumer protection – it was perhaps inevitable that this aspect of regulation should have been their main focus. The overarching desire on the part of the FSA's architects was to establish a strong consumer protection regulator that would be independent of the industry.¹³⁰ However, one consequence of the consumer protection focus was that the financial stability objective did not receive the attention that it either warranted or deserved.¹³¹ Fortunately, the Northern Rock episode has provided the impetus to restore some balance to the post-1997 arrangements. The Bank of England has now been given both formal statutory responsibility for financial stability and for handling bank resolutions under a new legislative framework, the Banking Act 2009.¹³²

¹²⁸ SELECT COMMITTEE ON ECONOMIC AFFAIRS, BANKING SUPERVISION AND REGULATION, 2008-09, H.L. 101-I, at 33.

¹²⁹ *Id.* at 32.

¹³⁰ *Id.* at 30, 32, 52-53.

¹³¹ *Id.* at 31-32.

¹³² *Id.* at 30-31.

However, as noted earlier, there are still a number of aspects of the “Tripartite system” where reforms are still needed. In addition, although charged with the formal statutory responsibility for maintaining financial stability the Bank of England lacks most of the policy tools it needs for this task. Hence, even now, the rebalancing is only partly finished.

A second dimension of the financial stability focus concerns what is now termed “macroprudential” regulation. When the U.K.’s arrangements were put in place, prudential regulation was conceptualized in terms of ensuring the soundness of individual institutions. As the financial crisis has made clear, however, ensuring the soundness of individual firms is a necessary, but not sufficient condition for ensuring financial stability. While in one respect the comparative neglect of financial stability issues under the U.K.’s post-1997 arrangements was due to an oversight, it also reflected the fact that what is now called the macroprudential perspective had not at the time gained the prominence that it now enjoys. As noted above, an unfinished aspect of the U.K.’s attempt to re-balance its regulatory system concerns the additional macroprudential powers that should be assigned to the Bank of England.

D. POLITICALLY MOTIVATED REFORMS OR THOSE MOTIVATED BY A DESIRE TO “PUNISH” THE CENTRAL BANK ARE COUNTERPRODUCTIVE

There is at least some circumstantial evidence for concluding that part of the motivation for the U.K.’s reforms was to “punish” the central bank or to “cut it down to size.”¹³³ However, as the subsequent British experience shows, there is no plausible alternative to having a central bank with an extensive mandate and the ability to intervene to mitigate a crisis. The FSA’s architects appear to have believed that it would be possible to create a rival center of power to the Bank, without realizing the reality that without significant financial muscle of its own, the FSA was destined to play a subsidiary role in any crisis. Only the central bank has the ability to play the role of LoLR and this fact means that it must play a unique role in any financial safety net arrangement. The members of the U.S. Congress who have recently criticized the Federal Reserve for its actions in stemming the crisis need to reflect on whether there are any viable alternatives. The British experience suggests that there are not.

¹³³ Jacomb, *supra* note 49, at 3-4.

E. SOME “OVERLAP” AND “DUPLICATION” OF REGULATORY FUNCTIONS IS UNAVOIDABLE

As should now be apparent, the U.K.’s regulatory reforms were inspired, to a very large extent, by the desire to eliminate the perceived duplication and overlap of regulatory authority resulting from the Financial Services Act system in particular. While the Act had indeed created a system that was excessively complex – especially from the point of view of the individual consumer – this factor arguably received too much attention in the resulting reforms.

A particularly clear example was the decision not to give the Bank of England its own powers to gather information from the financial sector (banks in particular). It was therefore reliant on the FSA to provide it with the data it required to perform its “financial stability” function.¹³⁴ The thinking appeared to be that if the FSA was to be the banking supervisor, the Bank of England should have only a general role in relation to overall financial stability, and did not require the ability to gather institution-specific information.¹³⁵ Because one stated objective of the 1997 reforms was to reduce regulatory duplication and overlap – a major selling point with the industry – only the FSA was given information-gathering powers.

This decision ignored the experience of many other countries where the central bank was not itself the prudential regulator, and indeed the Bank of England’s own history before it assumed the statutory responsibility for bank regulation in 1979.¹³⁶ In its role as lender of last resort it had been able to exert significant moral suasion over the banking sector, and the Discount Office was able to obtain information from banks on a purely informal basis.¹³⁷ Other central banks also enjoy substantial information gathering powers of their own.¹³⁸ For example, the Bank of Japan’s information-gathering ability includes the power to conduct bank examinations, notwithstanding that this duplicates the function of the Financial Services Agency. These precedents should have shown that even without the formal statutory responsibility for banking supervision, the

¹³⁴ BROWN, *supra* note 109, at ¶ 8.

¹³⁵ *Id.* at ¶ 6.

¹³⁶ See Schooner & Taylor, *supra* note 2, at 629-32.

¹³⁷ *Id.* at 614-15.

¹³⁸ INSTITUTE FOR MONETARY AND ECONOMIC STUDIES, FUNCTIONS AND OPERATIONS OF THE BANK OF JAPAN 80 (2004), available at www.boj.or.jp/en/type/exp/about/data/fobj01.pdf

central bank still needed to have access to substantial amounts of institution-specific information and ideally its own capacity to go about gathering that information.

F. CRISIS MANAGEMENT PREPAREDNESS MATTERS

Finally, insufficient attention was given to crisis management arrangements. For at least two decades prior to the formation of the FSA, the U.K. had not experienced any episodes of serious financial distress. This may have bred a certain degree of complacency about the need for adequacy crisis management preparedness and planning. Although the Memorandum of Understanding was drawn up between the Treasury, Bank of England, and FSA, the arrangements envisaged were rarely tried in practice and the Joint Crisis Management Committee rarely met. The arrangements also assumed that the Treasury would be the glue that held this system together, thus involving it in the minutiae of crisis management decision-making – a role that it was ill-equipped to perform and one that hampered the ability to reach quick decisions in an environment where time was of the essence.¹³⁹ It is therefore necessary to ensure that the central bank’s freedom of manoeuvre is not excessively constrained by any arrangements that are put in place.

VII. CONCLUSION: THE RETURN TO “TWIN PEAKS”

In the aftermath of the Global Financial Crisis there has been a revival of interest in the Twin Peaks model. The experience of the U.K. during the financial crisis has strengthened the arguments of the FSA’s critics who challenged the viability of a single regulatory agency in a financial centre as large and diverse as the U.K. The British Conservative party, which at the time of writing is still in opposition but is expected to win the election due in 2010, has now adopted the policy of abolishing the FSA and introducing a division between prudential and conduct of business regulation with the former being returned to the Bank of England.¹⁴⁰ Similarly, in the U.S., the Twin Peaks concept has received attention in

¹³⁹ Peter Hayward “The Financial Sector – The Responsibilities of Public Agencies” in CHARLES ENOCH et al, *supra* note 118.

¹⁴⁰ George Osborne, *Foreword*, in FROM CRISIS TO CONFIDENCE: PLAN FOR SOUND BANKING (2009).

evidence given to Congressional committees¹⁴¹ and as a major source of the Bush administration's proposals of 2008.

As can be seen from the above analysis, when Britain adopted its single regulator structure in 1997, it did not do so due to a conscious rejection of the Twin Peaks alternative. Rather, the British government's decision seems to have owed more to the legislative timetable and the apparent simplicity of the single regulator in avoiding some of the complexity, duplication and opacity which had been a focus of the criticisms of the previous system. The single regulator's very simplicity may well have been a factor in its favour; but the apparent simplicity of the structure was deceptive as it resulted in some of the complexities of financial regulation and crisis management being neglected.

The Twin Peaks alternative might, arguably, have avoided some of the design flaws of the post-1997 arrangements. In particular, it would have avoided trying to set up a rival center of power to the Bank of England, thereby creating crisis management arrangements that were far too unwieldy. Because the Bank and the FSA were assigned equal status in the Tripartite arrangements, the active role of the Treasury was essential to hold the ring and to ensure a cooperative relationship between the two agencies. By contrast, a specialist prudential regulator might have been established more clearly under the Bank of England's wing, and as a result could have enjoyed much closer links with the central bank than did the FSA. There are a variety of precedents for this possible arrangement: the relationship between the Bank of France and the Commission Bancaire, or between the Finnish Central Bank and that country's Financial Supervision Agency could have been potential models.¹⁴² In these structures, although the regulatory agencies are governed by boards separate from those of the central bank, their staff are central bank employees and extensive use is made of shared facilities, information technology platforms and databases.

Nonetheless, although Twin Peaks has its attractions, it is necessary to be cautious about trying to introduce too much neatness and tidiness into regulatory structures. The objectives of financial regulation can be neatly packaged into two, but the range of regulatory functions is far more diverse. At least six (or possibly seven) regulatory functions can be

¹⁴¹ See *Enhancing Investor Protection and the Regulation Of Securities Markets: Hearing Before S. Comm. on Banking, Housing and Urban Affairs*, 11th Cong. 36-39 (2009) (testimony of Mr. John Coffee, Professor, Columbia Law School).

¹⁴² See Abrams & Taylor, *supra* note 9, at 23-24.

identified: financial system stability; crisis management; the prudential regulation of systemically important firms; the prudential regulation of firms that are not systemically important; sales practice regulation; and market conduct regulation.¹⁴³ (Competition policy is a possible seventh regulatory function although it applies in many sectors other than financial services.) At its most basic, the problem of designing a regulatory structure is one of deciding which of these functions belong together in the same agency. The single regulator concept tried to combine most of these functions within one agency. That has been shown to be a step too far. But there are many possible configurations between this option and the current highly fragmented regulatory system in the United States.

¹⁴³ See C. A. E. GOODHART ET AL., FINANCIAL REGULATION: HOW, WHO AND WHERE NOW? (1998).

WHITHER THE DUTY OF GOOD FAITH IN UK INSURANCE CONTRACTS

John Lowry*

This article explores the current state of the law in the United Kingdom concerning the duty of good faith in insurance contracts. Recent case law provides that the duty of disclosure by insureds is constantly being refined. It argues that due to the current fragmentation of the law, future reform should be focused on creating a consistent regime for insurance contracts. Such regime should be flexible enough to encompass both consumer and commercial insurance, while demonstrating certain and clear objectives. The first part examines the duty of disclosure by an insured as formulated by Lord Mansfield CJ. The second part analyzes the case law that followed Carter v. Bohem, which developed the notion of good faith and expanded it into a duty of utmost good faith. Third, the discomfort of the UK courts and UK law reform agencies over the severity of the insured's duty along with the injustices that result when insurers avoid a policy for non-disclosure is explored. Fourth, recent judicial opinions that attempt to alleviate the position of the insured are assessed. The fifth and concluding part of this article briefly examines the 2009 [UK] Consumer Insurance (Disclosure and Representations) Bill published by the English and Scottish Law Commissions in December 2009. It constructs an alternative model which takes account of recent developments in Australian law. It is argued that the focus should be on balancing the economic costs of reform with the benefits of a more balanced regime which does not create a distinction between consumer and business insureds.

* Professor of Law and Vice Dean of the Faculty of Laws, UCL. I owe a debt of gratitude to the anonymous referees for their helpful comments. Liability for any errors, however, is mine alone. I would also like to thank the editors of the Journal for affording me the opportunity to amend the article immediately prior to its publication so as to include reference to UK reform proposals published in December 2009.

Insurance contracts are highly unusual in that they are founded upon the doctrine of *uberrimae fidei*.¹ The consequence is that the general contractual duty borne by parties to avoid misrepresentation is extended and reinforced by the additional obligation to disclose all material facts that would induce the insurers to underwrite the risk. This was first laid down by Lord Mansfield C.J. in *Carter v. Boehm*,² and his formulation of the disclosure duty is partially codified in the Marine Insurance Act 1906.³ Lord Mansfield was at the time attempting to import into English commercial law the civil law notion of good faith, but this ultimately proved unsuccessful and only survived for a very limited class of transactions, including insurance.⁴

The effect of non-disclosure by the insured entitles the insurer to avoid the contract *ab initio*, notwithstanding the absence of any fraudulent intent. The economic consequences are severe and disproportionately harsh. The policy becomes valueless so that the insured loses the financial safeguard that the policy was designed to provide should the losses caused by the insured risks come to pass.⁵ This is not to say that the rationale underlying the disclosure duty and the remedy for its breach is in any way obscure. The role of underwriters is to assess risk and if there are material factors known only to the insured, then the insured must disclose them. The reason for the duty is clear where the underwriter is not in a position to

¹ For the common law position governing the general law of contracts *see, e.g.*, *Keates v. Cadogan*, (1851) 138 Eng. Rep. 234. For a review of the policy considerations underlying the general contractual position *see* J. BEATSON, ANSON'S LAW OF CONTRACT 263-64 (28th ed. 2002).

² (1766) 97 Eng. Rep. 1162, 1165.

³ *See* Marine Insurance Act, 1906, 6 Edw. 7, c. 41, §§ 17-20 (Eng.).

⁴ *See* Lord Mustill's speech in *Pan Atl. Ins. Co. Ltd. v. Pine Top Ins. Co.*, [1994] 1 A.C. 501 (H.L.), and Lord Hobhouse's speech in *Manifest Shipping Co. Ltd. v. Uni-Polaris Shipping Co. Ltd.*, [2003] 1 A.C. 469. *See also* Potter L.J.'s observations in *James Spencer & Co. Ltd. v. Tame Valey Padding Co. Ltd.*, QBENI 97/1118 CMS1 (A.C. April 8, 1998) (Smith Bernal Reporting Ltd. for Lawtel). It is noteworthy that Lord Mansfield was familiar with the civilian tradition (*see* Sir WILLIAM HOLDSWORTH, 5 A HISTORY OF ENGLISH LAW 147 (1966)). This is not to suggest that there is a universal view among civilian jurisdictions on the meaning and scope of good faith. *See* R. ZIMMERMANN & S. WHITTAKER, GOOD FAITH IN EUROPEAN CONTRACT LAW (Cambridge, CUP, 2000). *See also* M. Bridge, *Doubling Good Faith*, NEW ZEALAND BUSINESS LAW QUARTERLY, 2005, at 426.

⁵ *HIH Cas. & Gen. Ins. Co. v. Chase Manhattan Bank* [2003] UKHL 6 (statement of Lord Hobhouse).

discover for himself or herself the circumstances that may impact upon the risk. For example, one case would be where the insured seeks to take out a life policy having received a death threat. Or a vessel the underwriter is asked to insure may be on the high seas and, therefore, not available for inspection.⁶ In such circumstances there can be no quarrel with the fairness of the insured's duty of good faith. But in the case law subsequent to *Carter v. Boehm*, the boundaries of the insured's duty were expanded beyond, it is suggested, what Lord Mansfield originally envisaged. For example, it has been held that the failure of an insured to disclose criminal convictions and his precarious financial position when applying for fire insurance will enable the insurer to avoid the policy following a fire caused by an electrical fault.⁷

In practical terms, the issue which pervades the duty of good faith can be reduced to the following question: how can the ordinary insured, whether acting in a private or commercial capacity, untutored in the niceties of insurance law, be expected to know what particular circumstances are material and would, therefore, influence the prudent underwriter?⁸ The sheer breadth of the insured's duty together with the all-or-nothing consequence of avoidance, therefore, rightly gives rise to legitimate concern. In a series of recent cases, decided over the last ten years or so, the English courts have been steadily refining the disclosure duty while, at the same time, laying considerable emphasis on the mutuality of the requirement of good faith by giving content to that borne by insurers. This process is a rebalancing exercise.⁹ As such, it involves a tacit

⁶ See NEW ZEALAND LAW COMMISSION, No. 46: SOME INS. LAW PROBLEMS 2 (1998).

⁷ *Quinby Enter. Ltd. v. Gen. Accident Ltd.* [1995] 1 N.Z.L.R. 736. *But cf.* Waller L.J.'s reasoning in *North Star Shipping Ltd. v. Sphere Drake Ins. plc* [2006] EWCA (Civ) 378; *see also* *Doheny v. New India Assurance Co.* [2004] EWCA Civ. 1705; *O'Kane v. Jones* [2003] EWHC 3470 (Comm.); *James v. CGU Ins. plc*, [2002] Lloyd's Rep. I.R. 206; *March Cabaret Club & Casino Ltd. v. London Assurance*, [1975] 1 Lloyd's Rep. 169.

⁸ See Lord Mansfield's formulation of the duty and the Marine Insurance Act, 1906, 6 Edw. 7, c.41, § 18.

⁹ Indeed, this is in line with various calls for reform which have long gone unheeded by the legislature. *See infra* note 15. That said, the UK Financial Ombudsman Service (FOS) has sought to mitigate the harshness of the duty, at least in so far as it applies to consumer insureds and small businesses with a group annual turnover of less than £1 million. For commercial insureds, however, who

unraveling of case law spanning over a century in which the courts had adopted an unequivocal stance in permitting avoidance for non-disclosure across the range of insurance, both consumer and commercial, without regard to notions of fairness, proportionality, or whether there was actual inducement.¹⁰ Admittedly the anxiety of modern judges has generally been directed towards relieving the position of the consumer or private insured, but the limits of this shift of emphasis are not entirely clear. It begs the question whether commercial insurance is also being targeted. Such doubt carries the danger of undermining the very certainty that should represent the cornerstone of commercial law in this respect and the economic implications are potentially significant.

Taken in the round, it is possible to distill several strands of reasoning from the modern case law. First, focusing solely upon the content of the insured's pre-contractual duty of disclosure, the courts have sought to limit its scope by refining the conditions, most notably the requirement of inducement, that must be met before the insurer may justifiably avoid the policy for non-disclosure. Further, in relation to avoidance, it is noteworthy that in recent times some judges have had recourse to notions of good faith, conscience, and fairness when assessing whether insurers may exercise the remedy. But judicial thinking in this respect is not entirely consistent for it has been suggested that, as with the remedy of rescission for misrepresentation, the rights of insurers are unfettered by such considerations.¹¹ Another strand of reasoning that has emerged has been directed towards the contours of waiver. An insurer has every opportunity to ask specific questions of the applicant for insurance in the proposal form. Typically, those questions will be directed towards claims history or health where the application relates to life or sickness insurance. Nonetheless, even where such questions are raised, the insured is not relieved from his or her duty to volunteer any further material circumstances that fall outside the scope of them. Any defense that an insured might seek to raise based on waiver is, in the orthodox view, doomed to failure.¹² However, this has been challenged recently in a

are outwith the FOS jurisdiction, it has been left to the courts to alleviate their position.

¹⁰ See, e.g., the judgments in *Lambert v. Co-operative Ins. Soc'y*, [1975] 2 Lloyd's Rep. 485. See also *infra* notes 82-91.

¹¹ See the judgment of Mance L.J. in *Brotherton v. Aseguradora Colseguros S.A.* [2003] EWCA Civ. 705. See *infra* note 197.

¹² See, e.g., *Roselodge, Ltd. v. Castle*, [1966] 2 Lloyd's Rep. 113, 132.

powerful dissent delivered by Rix L.J., whose reasoning may portend future developments.¹³ In this regard, considerable emphasis is now being paid to the importance of the presentation of risk. The modern view is that insurers should not be content to play a passive role during the disclosure process but should be prepared to make necessary enquiries about the risk to be underwritten.¹⁴ Finally, the question of whether or not the insured's duty of good faith continues post-contractually and again triggers at the claims stage has also attracted considerable judicial scrutiny of late.

This article is in five parts. It first examines the scope of the insured's duty of disclosure originally formulated by Lord Mansfield C.J.. Secondly, it considers the case law subsequent to *Carter v. Boehm* in which the notion of good faith was developed and expanded into a duty of so-called *utmost* good faith. The third part will outline the unease expressed by the courts and by the law reform agencies, over the harshness of the insured's duty and the injustices that result when insurers avoid a policy for non-disclosure.¹⁵ The fourth part will assess recent judicial inroads into the orthodox position that appear to be aimed at alleviating the position of the insured. In this respect, the starting point will be *Pan Atlantic Insurance Co. v. Pine Top Insurance Co.*¹⁶ Although the House of Lords took the opportunity to settle the legal position relating to the insured's duty of disclosure, it did not quell the debate surrounding the perceived iniquities of the insurer's remedy. Finally, against the backdrop of modern English case law, together with key developments in Australia, a model will be

¹³ See *WISE Ltd. v. Grupo Nacional Provincial S.A.* [2004] EWCA Civ. 962. See *infra* note 168-87.

¹⁴ This is not a novel development but reflects the view expressed by Lord Mansfield C.J. in *Carter v. Boehm*, (1766) 97 Eng. Rep. 1162. However, by the mid-nineteenth century, the point seems have faded from judicial thinking when addressing the scope of the disclosure duty.

¹⁵ See, e.g., LAW REFORM COMMITTEE, *CONDITIONS AND EXCEPTIONS IN INSURANCE POLICIES*, 1957, Cmnd. 62; THE LAW COMMISSION, *INSURANCE LAW – NON-DISCLOSURE AND BREACH OF WARRANTY*, 1980, Cmnd. 8064; JOHN BIRDS, NAT'L CONSUMER COUNCIL, *INSURANCE LAW REFORM: THE CONSUMER CASE FOR A REVIEW OF INSURANCE LAW* (1997). Although the UK Department of Trade and Industry (now renamed the Department for Business, Innovation and Skills (DBIS)) issued a draft Insurance Bill following the Law Commission's report in 1980, nothing came of it, the government of the day being satisfied that the industry's response via the Association of British Insurers' Statement of Practice struck the appropriate balance for the consumer-insured.

¹⁶ [1995] 1 A.C. 501 (H.L.).

proposed that might serve to inform the English and Scottish Law Commissions' current re-examinations of insurance law that identifies non-disclosure as a key issue.¹⁷ It will be seen that the current state of the law is fragmented and complicated and that future reform should be directed towards constructing a coherent regime for insurance contracts that meets the objectives of certainty and clarity, while being sufficiently flexible to encompass both consumer and commercial insurance.

I. THE ORIGINS OF THE DISCLOSURE DUTY AND ITS STATUTORY CODIFICATION

A. *CARTER V. BOEHM*

An enduring legacy of the Seven Years War is that it left us with a landmark decision which contains the most quoted passage in U.K. insurance law. In *Carter v. Boehm*, the issue of non-disclosure came to court as a result of the Governor of Sumatra, George Carter, effecting a policy of insurance on Fort Marlborough, a trading fort, against the likelihood of a French attack.¹⁸ His decision to insure was vindicated when, in April 1760, the fort was seized by the French.¹⁹ The Governor's claim was disputed by the underwriter and in 1766, Lord Mansfield C.J.,

¹⁷ The English and Scottish Law Commissions are statutory independent bodies created by the [UK] Law Commissions Act 1965 c. 22, to keep the law under review and to recommend reform where it is needed. The insurance contract law reform project was announced on 14 October 2005, the first "issues paper" on misrepresentation and non-disclosure was published at the end of September 2006. See The Law Commission, *Insurance Contract Law, Misrepresentation and Non-Disclosure* (Sept. 2006), available at http://www.lawcom.gov.uk/docs/insurance_contact_law_issues_paper_1. See *infra* notes 305-11. On December 15 2009 the Law Commissions published their joint report and draft Bill to reform the law on what a consumer-insured must disclose to the insurers prior to the conclusion of the policy; see CONSUMER INSURANCE LAW: PRE-CONTRACT DISCLOSURE AND REPRESENTATION, 2009, *Cm 7758*, discussed *infra* note 307 *et seq.*

¹⁸ *Carter*, 97 Eng. Rep. at 1163. Park notes that *Carter* is a seminal case ("... from it may be collected all the general principles which the doctrine of concealments, in matters of insurance, is founded, as well as all the exceptions. . . ."). JAMES ALLAN PARK, *A SYSTEM OF THE LAW OF MARINE INSURANCES* 193 (Thomas & Andrews 1800) (1787).

¹⁹ *Carter*, 97 Eng. Rep. at 1163.

presiding at Guildhall, heard the consequent action.²⁰ The underwriters had sought to avoid the contract on the basis that the Governor had failed to disclose the fort's weakness and its vulnerability to an attack by the French.²¹ Their defence failed, but Lord Mansfield took the opportunity to formulate the duty of good faith which has come to represent a cornerstone of English insurance law:

The special facts, upon which the contingent chance is to be computed, lie most commonly in the knowledge of the insured only; the under-writer trusts to his representation, and proceeds upon confidence that he does not keep back any circumstance in his knowledge, to mislead the under-writer into a belief that the circumstance does not exist, and to induce him to estimate the *risqué*, as if it did not exist.

The keeping back such circumstance is a fraud, and therefore the policy is void. Although the suppression should happen through mistake, without any fraudulent intention; yet still the under-writer is deceived, and the policy is void; because the *risque* run is really different from the *risque* understood and intended to be run at the time of the agreement.

.....

Good faith forbids either party by concealing what he privately knows, to draw the other into a bargain, from his ignorance of that fact, and his believing the contrary.

But either party may be innocently silent, as to grounds open to both, to exercise their judgment upon. *Aliud est celare; aliud, tacere; neque enim id est celare quicquid reticeas; sed cum quod tuscias, id ignorare emolumenti tui causa velis eos, quorum intersit id scire.*²²

This formulation begs the question as to why insurance contracts are exceptional in requiring a positive duty of disclosure. Although the disparity of knowledge between the parties has been proffered as the explanation, this is hardly a satisfactory rationale in itself given that in

²⁰ *Id.* at 1162-63.

²¹ *Id.* at 1163.

²² *Carter*, 97 Eng. Rep. at 1164.

other common contracting situations, the parties similarly lack equality of information. The explanation for this exceptional feature of insurance must, therefore, lie elsewhere. The investigation begins by considering Lord Mansfield's reasoning. It then goes on to review the jurisprudence surrounding the good faith requirement and how it evolved into a duty of so-called utmost good faith.

B. THE RATIONALE

Lord Mansfield explained that the policy considerations underlying the duty are the prevention of fraud and the furtherance of good faith: it therefore fulfils a prophylactic role.²³ He based it upon the concept of "concealment," but over time this developed beyond deliberate concealment so as to encompass all non-disclosure, however innocent, of a material fact. In *Carter v. Boehm*, the underwriter had argued that the insured had been fraudulent in failing to disclose the fort's vulnerability to attack.²⁴ This contention was unsuccessful, it being held that the underwriter must be taken to have realised that the Governor, by insuring, obviously apprehended the possibility of an attack.²⁵ By underwriting the risk, the insurer thereby assumed knowledge of the state of the fort.²⁶ It was stressed that the underwriter, sitting in London, was in a better position than the Governor to stay informed about the fortunes of the war and so it was not a matter within the private knowledge of the Governor only, but was, in fact, in the public domain.²⁷ Lord Mansfield concluded that a verdict in favour of the underwriters would have had the effect of turning a rule against fraud into an instrument of fraud.²⁸ He proceeded on the basis that good faith was a mutual duty, not an obligation borne solely by

²³ This is a narrower view than that expressed over a century later by Channell J in *Re Yager & Guardian Assurance Co.*, [1912] 108 L.T. 38 (K.B.), to the effect that the rationale underlying the disclosure duty is not the need to prevent harm to the insurer as such, but the need for a true and fair agreement whereby risk is transferred. *Id.* at 44-45.

²⁴ *Carter*, 97 Eng. Rep. at 1163.

²⁵ *Id.* at 1167.

²⁶ *Id.*

²⁷ *Id.*

²⁸ *Id.* at 1169.

insureds, and being mutual the problem of unfairness was shared between the parties.²⁹

Lord Mansfield also placed emphasis on the need for insurers to demonstrate reliance. He explained that the underwriter “proceeds upon confidence that [the insured] does not keep back any circumstance in his knowledge, to mislead the under-writer into a belief that the circumstance does not exist, and to induce him to estimate the risque, as if it did not exist.”³⁰ Herein we see the assimilation of non-disclosure with misrepresentation. The good faith duty converts non-disclosure into misrepresentation because an insured who fails to disclose a material fact is effectively misrepresenting the true state of affairs. Lord Mansfield’s choice of language is critical: it traverses the two vitiating factors and reliance and inducement lies at the heart of both. For misrepresentation, the consequence is therefore the same as with pure non-disclosure, namely avoidance of the contract *ab initio*.³¹

Both in *Carter v. Boehm*, and in subsequent cases, Lord Mansfield sought to limit the scope of the insured’s duty by, for example, stressing the need for underwriters to be proactive in ascertaining facts material to the

²⁹ Indeed, Lord Mansfield was scathing in his condemnation of the underwriter’s defense:

The underwriter, here, knowing the governor to be acquainted with the state of the place; knowing that he apprehended danger, and must have some ground for his apprehension; being told nothing of neither; signed this policy, without asking a question. If the objection ‘that he was not told’ is sufficient to vacate it, he took the premium, knowing the policy to be void; in order to gain, if the alternative turned out one way; and to make no satisfaction, if it turned out the other: he drew the governor into a false confidence . . .

. . . If he thought that omission an objection at the time, he ought not to have signed the policy with a secret reserve in his own mind to make it void; if he dispensed with the information, and did not think this silence an objection then; he cannot take it up now, after the event.

Id. at 1169.

³⁰ *Carter*, 97 Eng. Rep. at 1164.

³¹ The insurers remedy in this respect is codified by the Marine Insurance Act, 1906, 6 Edw. 7, c. 41, § 20 (Eng.).

risk. In *Noble v. Kennoway*, the insured vessel had arrived safely at Labrador but prior to being unloaded, it was used for fishing.³² Subsequently, the vessel was taken by privateers while it was unmanned.³³ The owners' claim for the value of the cargo was met with the defence by the underwriters that they were not liable because of the delay in unloading.³⁴ The insured argued that this was a trade usage in this particular port because of the lack of warehousing.³⁵ Lord Mansfield, finding for the insured, reasoned that every underwriter was presumed to know the practices of the trade he insures and if he does not know then it is his duty to inform himself of it.³⁶ He returned to the point in *Mayne v. Walter*, where the insured's claim for the loss of supercargo seized by the French was met with the defence that he should have disclosed the existence of a French ordinance prohibiting Dutch ships carrying the supercargo of any country at war with France on pain of it being taken as prize.³⁷ Lord Mansfield said that if both parties were ignorant of the relevant fact, "the underwriter must run all risks: and if the [underwriter] knew of such an edict, it was his duty to inquire, if such supercargo were on board."³⁸ He went on to note that "[i]t must be a fraudulent concealment of circumstances, that will vitiate a policy."³⁹ This has been termed the narrow Mansfield rule.⁴⁰ Reflecting upon his original formulation, Lord Mansfield appears to have come around to the view that the duty is limited insofar as it must strike a balance between the parties so as to achieve some symmetry between them. Indeed, by the early nineteenth century, emphasis was being placed on Lord Mansfield's clear admonition that underwriters have a distinct investigative role to play in the disclosure process. For example, in *Friere v. Woodhouse*, a marine insurance case, Burrough J. said, "what is exclusively known to the assured ought to be communicated; but what the underwriter, by fair inquiry and due diligence, may learn from ordinary sources of information need not be

³² 99 Eng. Rep. 326, 326-27.

³³ *Id.* at 326.

³⁴ *Id.*

³⁵ *Id.*

³⁶ *Id.* at 327.

³⁷ *Mayne v. Walter*, (1782) 99 Eng. Rep. 548, 548-49. PARK, *supra* note 18, at 196.

³⁸ PARK, *supra* note 18, at 196.

³⁹ *Id.*

⁴⁰ See R. A. Hasson, *The Doctrine of Uberrima Fides in Insurance Law - A Critical Evaluation*, 32 M.L.R. 615, 618 (1969).

disclosed.”⁴¹ The material information could have been discovered by the underwriter from *Lloyd’s List*.⁴²

It is striking that throughout his judgments on the issue of non-disclosure Lord Mansfield avoided the terminology of “utmost” good faith. Yet section 17 of the Marine Insurance Act 1906, the preamble of which declares it to be a codifying statute, states that insurance is *uberrimae fidei*.⁴³ It goes on to provide that a contract of insurance is a contract based upon the duty of *utmost* good faith which, if broken, entitles the other party to avoid the contract.⁴⁴ Section 17 does not, therefore, precisely mirror the language of Lord Mansfield’s formulation which, as seen above, draws the distinction between deliberate concealment and misrepresentation (bad faith) and innocent (good faith) mistaken non-disclosure. The provision must, therefore, be seen as synthesising not Lord Mansfield’s views, but rather the dominant view emerging from the case law decided during the latter half of the nineteenth century to the effect that the underwriter is a passive recipient of information supplied by the insured when presenting the risk. As such, this loses sight of the more restrictive views expressed not only in *Carter v. Boehm*, but which was also in the judgments found in *Noble v. Kennoway*, *Mayne v. Walter*, and *Friere v. Woodhouse*.⁴⁵ It was certainly not within the mandate of Sir Mackenzie Chalmers, who drafted the Digest upon which the 1906 Act was based, to correct the significant body of case law he sought to codify.⁴⁶ The question that arises, therefore,

⁴¹ *Friere v. Woodhouse*, (1817) 171 Eng. Rep. 345, 345. See also *Gandy v. The Adelaide Marine Ins. Co.*, 6 Eng. Rep. 746, 757 (Q.B.).

⁴² *Friere*, 171 Eng. Rep. at 345.

⁴³ It has long been settled that the Marine Insurance Act, 1906, 6 Edw. 7, c. 41, §§ 17-18 (Eng.) are of general application in insurance law. See, e.g., *P.C.W. Syndicates v. P.C.W. Reinsurers*, (1996) 1 W.L.R. 1136, 1140 (A.C.); *Australia and New Zealand Bank, Ltd. v. Colonial and Eagle Wharves Ltd.*, (1960) 2 Lloyd’s L. Rep. 241, 251 (Q.B.D.); *Cantiere Meccanico Brindisino v. Janson*, (1912) 3 K.B. 452, 467. More recent cases have also proceeded on the assumption that the provisions apply to all types of insurance. See *Pan Atlantic Ins. Co. Ltd. v. Pine Top Ins. Co. Ltd.*, [1994] 1 A.C. 501, 518 (H.L.); *Lambert v. Co-operative Ins. Soc’y*, (1975) 2 Lloyd’s Rep 485, 487.

⁴⁴ Marine Insurance Act, 1906, 6 Edw. 7, c. 41, § 17 (emphasis supplied).

⁴⁵ *Friere v. Woodhouse*, (1817) 171 Eng. Rep. 345; *Mayne v. Walter*, (1782) 99 Eng. Rep. 548; *Noble v. Kennoway*, (1780) 99 Eng. Rep. 326. See also Lord Ellenborough’s closely reasoned judgment delivered in *Haywood v. Rodgers*, (1804) 4 East 590.

⁴⁶ See note 45 *infra* for conflicting case law.

is when did the disclosure duty metamorphose into something requiring *utmost* good faith?

II. THE EVOLUTION OF *UBERRIMAE FIDEI*

In broad terms, a hallmark of much of the case law subsequent to *Carter v. Boehm* is the expansive approach that was taken towards Lord Mansfield's formulation of the disclosure duty. It is not proposed to examine the merits of this case law in terms of whether the scope of Lord Mansfield's judgments were misconstrued,⁴⁷ but, as commented above, it was such that by the end of the nineteenth century, synthesising it required section 17 of the 1906 Act to declare insurance contracts to be of "utmost" good faith. Perhaps surprisingly, the suggestion that an insured must be of *utmost* honesty (as if there may be lower degrees of honesty) as represented in this statutory declaration was not seen as being particularly controversial or novel. As Lord Herschell, who originally took charge of the Bill when it was introduced in the House of Lords in 1894, explained, its purpose was to reproduce as exactly as possible the state of the existing law.⁴⁸

Tracing the antecedents of "utmost" good faith is an intriguing exercise, for it has no equivalent in the civil law.⁴⁹ Indeed, in *Mutual and Federal Insurance Co. Ltd. v. Oudtshoorn Municipality*,⁵⁰ the Supreme Court of Appeal of South Africa, expressing the view that the effect of the Pre-Union Statute Revision Act 43 of 1977 was to make South African insurance law governed by Roman-Dutch law, was moved to observe that "*uberrimare fides* is an alien, vague, useless expression without any meaning in law...our law of insurance has no need for *uberrimae fides* and the time has come to jettison it."⁵¹ Its origins can, however, be discerned in U.K. case law decided during the latter half of the nineteenth century.⁵² For example, in *Bates v. Hewitt*, the court paid little heed to Lord

⁴⁷ See Hasson, *supra* note 40.

⁴⁸ *Durant v. Durant*, 1 Haggard Eccl. Rep. 733.

⁴⁹ See M.A. Millner, *Fraudulent Non-disclosure* 74 S.A.L.J. 177, 188 (1957).

⁵⁰ *Mut. and Fed. Ins. Co Ltd. v. Oudtshoorn Mun.* 1985 (1) SA 419 (AD).

⁵¹ *Id.* at 433F.

⁵² See A. D. M. FORTE, *Good Faith and Utmost Good Faith* in A.D. M. Forte (ed), *GOOD FAITH IN CONTRACT AND PROPERTY LAW*, (Hart Publishing 1999). Joubert J.A. cites *Dalglisch v. Jarvis*, (1850) 2 Mac. & Gord. R. 231, as the decision in which the term *uberrimae fides* is first used in *Mut. and Fed. Ins.*, (1) SA at 431I.

Mansfield's views expressed over a century earlier.⁵³ The claimant had not informed the insurer that the insured vessel, the *Georgia*, had been a Confederate cruiser.⁵⁴ The *Georgia* was well known to the British public, and when the ship came to Liverpool for breaking she attracted considerable interest in both the press and in the House of Commons.⁵⁵ The defendant underwriter admitted he had known of the ship's history but that at the time of underwriting it was not present in his mind.⁵⁶ The jury found that the underwriter was ignorant of the vessel's notoriety at that particular time, although they did go on to express the view that when the risk was presented, he did have the means available for identifying the ship.⁵⁷ The court held that the claimant was in breach of his duty of disclosure.⁵⁸ Curiously, the judges in the case went to considerable lengths to explain that they were merely applying a long established principle. Lord Cockburn CJ stated that a proposer of insurance "is bound to communicate to the insurer all matters which will enable him to determine the extent of the risk against which he undertakes to guarantee the assured."⁵⁹ Shee J., while admitting that the underwriter might through his own investigations have discovered the material fact about the *Georgia*'s history, concluded, however, that he was under no duty to make such enquiries.⁶⁰ This fails to sit with Lord Mansfield's notion of an insurer's constructive knowledge - a critical factor in his finding in *Carter v. Boehm*. It also fails to sit with *Friere v. Woodhouse*, in which it will be recalled, the court, applying Lord Mansfield's formulation of the duty, had no hesitation in finding that underwriters had a pro-active role to play during the disclosure process.⁶¹ Nonetheless, towards the close of the nineteenth century, the consensus of judicial opinion was such that determining if the duty of disclosure has been discharged requires something more than merely exacting a duty of honesty from the insured. This came to the fore in *Life Association of Scotland v. Foster*, in which the term "utmost good faith" is adopted by Lord President Inglis: "Contracts of Insurance are in

⁵³ *Bates v. Hewitt*, (1867) 2 L.R.Q.B. 595; 1867 WL 9866.

⁵⁴ *Id.* at 604.

⁵⁵ *Id.* at 595.

⁵⁶ *Id.* at 604.

⁵⁷ *Id.*

⁵⁸ *Id.* at 599.

⁵⁹ *Bates*, 2 L.R.Q.B. at 604-05.

⁶⁰ *Id.* at 611.

⁶¹ *Friere v. Woodhouse*, (1817) 171 Eng. Rep. 345. *See also* *Noble v. Kennoway*, (1780) 99 Eng. Rep. 326.

this, among other particulars, exceptional, in that they require on both sides *uberrima fides*. Hence, without fraudulent intent, and even *bona fides*, the insured may fail in the duty of disclosure.”⁶²

This decision is followed soon after by *Ionides v. Pender*, in which Blackburn J. was moved to assimilate the prevailing view into the so-called “prudent insurer” test.⁶³ While the judge accepted that “it would be too much to put on the assured the duty of disclosing everything which might influence the mind of an underwriter,” he nevertheless conceded that “a concealment of a material fact, though made without any fraudulent intention, vitiates the policy.”⁶⁴ Blackburn J. concluded by stating that “all should be disclosed which would affect the judgment of a rational underwriter governing himself by the principles and calculations on which underwriters do in practice act.”⁶⁵

⁶² (1873) 11 M. 351, 359. It might be the case that this generation of judges was perhaps placing too much reliance on Park’s *Law of Marine Insurance* published in 1787, rather than tracing Lord Mansfield’s reasoning first hand. For example, chapter 10 of Park’s treatise states that “the learned judges of our courts of law, feeling that the very essence of insurance consists in a rigid attention to the purest good faith, and the strictest integrity, have constantly held that it is vacated and annulled by any the least shadow of fraud or *undue concealment*.” PARK, *supra* note 18, at 174 (emphasis supplied).

⁶³ (1874) 9 L.R.Q.B. 531. This was incorporated by Chalmers into § 18(2) of the Marine Insurance Act, 1906, 6 Edw. 7, c.41 (Eng.), which provides, “[e]very circumstance is material which would influence the judgment of a prudent insurer in fixing the premium, or determining whether he will take the risk.”

⁶⁴ *Ionides*, 9 L.R.Q.B. at 537, 539.

⁶⁵ *Id.* at 539. A further opportunity to put forward his view on the scope of the duty of disclosure was taken by him, now Lord Blackburn, in *Brownlie v. Campbell*, (1880) 5 App. Cas. 925, in which he noted:

... [i]n policies of insurance, whether marine insurance or life insurance, there is an understanding that the contract is *uberrima fides*, that if you know any circumstance at all that may influence the underwriter's opinion as to the risk he is incurring, and consequently as to whether he will take it, or what premium he will charge if he does take it, you will state what you know. There is an obligation there to disclose what you know; and the concealment of a material circumstance known to you, whether you thought it material or not, avoids the policy.

Id. at 954. Material facts are typically categorised as either those relating to

Blackburn J.'s formulation is encapsulated in sections 17 and 18(1) and (2) of the 1906 Marine Insurance Act. More particularly, section 18(1) lays down the overriding pre-contractual duty of disclosure while section 18(2), which provides that "every circumstance is material which would influence the judgment of a prudent insurer in fixing the premium, or determining whether he will take the risk," gives content to the governing principle of utmost good faith declared by section 17.⁶⁶ The view that insurance required nothing less than utmost good faith, or the idea of comparative degrees of honesty, thus became firmly entrenched in English insurance law.

Shortly after the 1906 Act received Royal Assent, the opportunity to explore the insured's duty of disclosure came before the Court of Appeal in *Joel v. Law Union and Crown Insurance Company*.⁶⁷ As a means of overcoming the practical difficulties of proof which a duty based solely upon utmost good faith could give rise to,⁶⁸ Fletcher Moulton L.J. superimposed a requirement of reasonableness. The judge explained that:

There is, therefore, something more than an obligation to treat the insurer honestly and frankly... There is the further duty that he should do it to the extent that a *reasonable man* would have done it; and, if he has fallen short of that by reason of his bona fide considering the matter not material, whereas the jury, as representing what a reasonable man would think, hold that it was material, he has failed in his duty, and the policy is avoided. This further duty is analogous to a duty to do an act which you undertake with reasonable care and skill, a failure to do which amounts to negligence, which is not atoned for by any amount of honesty or good intention. The disclosure must be of all you ought to have realized to be material, not of that only which you did in fact realize to be so.⁶⁹

physical hazard or those relating to moral hazard. See JOHN LOWRY & PHILIP RAWLINGS, *INSURANCE LAW: DOCTRINES AND PRINCIPLES* 93-99 (2d ed. 2005).

⁶⁶ Marine Insurance Act, 1906, 6 Edw. 7, c. 41, § 18 (Eng.).

⁶⁷ [1908] 2 K.B. 863.

⁶⁸ See PETER MACDONALD EGGERS ET AL., *GOOD FAITH AND INSURANCE CONTRACTS* ¶ 3.10 (2004).

⁶⁹ *Joel*, [1908] 2 K.B. at 883-84 (emphasis added). This passage occurs in a reserved judgment, and there is no indication that Vaughan Williams or Buckley

The insured is thus under a duty to disclose material facts, irrespective of whether he or she appreciated their materiality.⁷⁰ This combined test can, and does give rise to unjust results. For example, in *Horne v. Poland*, the insured's policy was voidable due to his failure to disclose that he was an alien having come to this country at the age of twelve, and that he had changed his name from that of Euda Gedale to Harry Horne.⁷¹ Lush J., having noted that the applicable principle of law had been stated by Fletcher Moulton L.J. in *Joel*, added:

If a reasonable person would know that underwriters would naturally be influenced, in deciding whether to accept the risk and what premiums to charge, by those circumstances [i.e., that he came from a country where his countrymen were not as careful and trustworthy as Englishmen], the fact that they were kept in ignorance of them and indeed were misled, is fatal to the plaintiff's claim.⁷²

Similarly, in *Becker v. Marshall*, Salter J., also applying the test laid down by Fletcher Moulton L.J., held on an issue of concealment as to foreign origin and change of name, that while the claimant:

... in good faith did not realise that these were things material to be disclosed...the average business man, the average reasonable man, would not have taken that view, and...that a reasonable man, the average reasonable man, would have disclosed and would have known that it was necessary to disclose.⁷³

L.J.J. dissented, although they base their concurring judgments on different grounds.

⁷⁰ See EGGERS ET AL., *supra* note 68, ¶¶ 3.10-3.11.

⁷¹ [1922] 2 K.B. 364, 364.

⁷² *Id.* at 367. The decision must now be viewed as running counter to the Race Relations Act, 1976, c.74 (Eng.).

⁷³ (1922) 11 Lloyd's List L.R. 114, 119 (K.B.). Fletcher Moulton L.J.'s reasoning was further explained by Roskill J in *Godfrey v. Britannic Assurance Co. Ltd.*, [1963] 2 Lloyd's Rep 515, 529 (Q.B.D.):

Curiously, for McNair J. in *Roselodge v. Castle*,⁷⁴ Fletcher Moulton L.J.'s formulation represents nothing less than a pure application of Lord Mansfield's rule. The dispute arose out of the rejection by the insurers of the insured's claim who, as diamond merchants, had insured diamonds against all risks.⁷⁵ The insurers' defence was founded upon the non-disclosure of two alleged material facts. First, that the principal director of the insured company had been convicted of bribing a police officer in 1946 and second, that the insured's sales manager had been convicted of smuggling diamonds into the United States in 1956.⁷⁶ According to one of the expert witnesses called by the insurer, a person who stole apples when aged 17 is much more likely to steal diamonds at the age of 67 even if he had led a blameless life for 50 years, than someone

[T]here is a point here which often is not sufficiently kept in mind. The duty is a duty to disclose, and you cannot disclose what you do not know. The obligation to disclose, therefore, necessarily depends on the knowledge you possess. I must not be misunderstood. Your opinion of the materiality of that knowledge is of no moment. If a reasonable man would have recognized that it was material to disclose the knowledge in question, it is no excuse that you did not recognize it to be so. But the question always is, Was the knowledge you possessed such that you ought to have disclosed it?

Such was the momentum of this approach that the 2nd (Hailsham) edition of the Laws of England, Volume 18, prepared by Scott L.J., stated that:

Materiality is a question of fact, not of belief or opinion. The assured does not therefore discharge his duty by a full and frank disclosure of what he honestly thinks to be material; he must go further and disclose every fact which a reasonable man would have thought material... If, however, the fact, though material, is one which he did not and could not in the particular circumstances have been expected to know, or if its materiality would not have been apparent to a reasonable man, his failure to disclose it is not a breach of duty.

Id. at 586(3). This passage is repeated in the 3rd (Simonds) ed., vol. 22, 360.

⁷⁴ [1966] 2 Lloyd's List L.R. 113, 131-32 (QBD).

⁷⁵ *Id.* at 113.

⁷⁶ *Id.*

who had led a totally blameless life.⁷⁷ This did not convince McNair J. who held that the 1946 conviction was not a material fact, having “no direct relation to trading as a diamond merchant.”⁷⁸ Having examined the authorities, with particular emphasis being given to *Horne v. Poland*, the judge concluded:

In my judgment, on this review of the authorities the judgment of Lord Justice Fletcher Moulton in *Joel's case* contains, if I may respectfully say so, a correct statement of the law on the topic. It has the merit...of emphasizing that even under the present practice of admitting expert evidence from underwriters as to materiality, the issue as to disclosability is one which has to be determined as it was in Lord Mansfield's day by the view of the Jury of reasonable men.⁷⁹

As conceded by Fletcher Moulton L.J., the duty does not require the insured to disclose that of which he or she is ignorant, unless the insured ought to have known of such circumstances in the ordinary course of business.⁸⁰ But, nevertheless, from the insured's perspective the disclosure duty laid down by the Marine Insurance Act of 1906, as explained by the subsequent case law, is particularly harsh and, it is suggested, represents an overly expanded view of Lord Mansfield's original formulation which was premised upon the notion of “concealment.” From a contemporary standpoint, it is hardly surprising that by the second half of the twentieth century both the courts and the law reform bodies were questioning whether such a strict approach was necessarily appropriate for all classes of insurance.

⁷⁷ *Id.* at 132.

⁷⁸ *Id.*

⁷⁹ *Id.* at 131.

⁸⁰ See also *Lambert v. Co-operative Ins. Soc'y Ltd.*, [1975] 2 Lloyd's Rep. 485, 490 (A.C.). See *infra* notes 63-66. The private insured will not, of course, be deemed to possess constructive knowledge. See *Economides v. Commercial Union Assurance Co. plc*, [1998] Q.B. 587, 601-03 (Simon Brown, L.J.).

III. CRITICISMS OF THE DUTY OF DISCLOSURE AND THE REMEDY FOR BREACH

Although modern judges inevitably follow the substantial line of case law on the Marine Insurance Act of 1906 §§ 17 and 18, they have not been timid in expressing their unease over the rigours of the disclosure duty. For example, in *Anglo-African Merchants Ltd. v. Bayley*, Megaw J. queried whether the insured should be bound to disclose that which he does not appreciate to be material.⁸¹ Further, in *Lambert v. Co-operative Insurance Society, Ltd.*, all three judges in the Court of Appeal took the opportunity to criticise the prudent insurer test.⁸² Lawton and Cairns L.J.J. went so far as to call for Parliamentary intervention to address the injustices caused by the harshness of the duty.⁸³ Briefly, the facts concerned an insured, Mrs. Lambert, who claimed under a household “all risks” policy that she and her husband had held for some nine years.⁸⁴ Neither at the commencement of the policy nor on its subsequent renewals had the insurers asked whether they had any criminal convictions.⁸⁵ When a claim was made for £311.00, representing the value of items of jewellery that had been lost, the insurers avoided liability on the basis that the criminal convictions of Mr. Lambert for, amongst other things, handling stolen cigarettes and stealing shirts, had not been disclosed.⁸⁶ In fact he was in prison at the time of the claim and could not, therefore, have been responsible for the loss.⁸⁷

In the event, Mrs. Lambert’s appeal against the decision of the trial judge who found in favour of the insurers failed.⁸⁸ MacKenna J., delivering the leading judgment in the Court of Appeal, took the law to be that stated by the Law Reform Committee in its 1957 report,⁸⁹ namely that the “question in every case is whether the fact not disclosed was material to the risk, and not whether the insured, whether reasonably or otherwise,

⁸¹ [1970] 1 Q.B. 311, 319.

⁸² [1975] 2 Lloyd’s Rep. 485, 491-93 (A.C.).

⁸³ *Id.* at 492-93.

⁸⁴ *Id.* at 486.

⁸⁵ *Id.*

⁸⁶ *Id.*

⁸⁷ *Id.*

⁸⁸ *Lambert*, [1975] 2 Lloyd’s Rep. at 491.

⁸⁹ LAW REFORM COMMITTEE, CONDITIONS AND EXCEPTIONS IN INSURANCE POLICIES, *supra* note 15.

believed or understood it to be so.”⁹⁰ Nevertheless, the judge went on to express considerable sympathy for Mrs. Lambert and stated that he hoped the insurers “would act decently if, having established the point of principle, they were to pay her. It might be thought a heartless thing if they did not, but that is their business, not mine.”⁹¹ As a matter of principle, the decision in *Lambert* is, of course, correct, although nowadays a consumer-insured would be able to refer the issue to the UK Financial Ombudsman Service rather than the courts, which, as will be seen, does not follow the strict law.

Considerable anxiety has also been expressed over the need for an insured to disclose allegations of dishonesty which, in fact, are false. The conundrum which arises here was identified by Forbes J. in *Reynolds and Anderson v. Phoenix Assurance Co. Ltd.*, who explained that the rule applied only to unfounded allegations.⁹² If the allegation was true, the insured was bound to disclose that he had committed the fraud and disclosure of the allegation added nothing.⁹³ Forbes J. noted that “the only occasion on which the allegation as an allegation must be disclosed is when it is not true. This appears to me to be a conclusion so devoid of any merit that I do not consider that a responsible insurer would adopt it and nor do I.”⁹⁴

However, against this, the view of Colman J. in *Strive Shipping Corp. v. Hellenic Mutual War Risks Association (Bermuda) Ltd.*,⁹⁵ reflects the orthodox approach taken towards the disclosure duty:

If an allegation of criminal conduct has been made against an assured but is as yet unresolved at the time of placing the risk and the evidence is that the allegation would have influenced the judgment of a prudent insurer, the fact the

⁹⁰ *Lambert*, [1975] 2 Lloyd’s Rep. at 489. MacKenna J. was particularly influenced by the opinion of the Privy Council in *Mutual Life Insurance Co. of New York v. Ontario Metal Products Co. Ltd.*, [1925] A.C. 344, 351-52, to the effect that the test, as laid down in the Marine Insurance Act, 1906, 6 Edw. 7, c. 41, § 18, is whether the non-disclosed fact would have influenced a reasonable insurer to decline the risk or to have stipulated for a higher premium.

⁹¹ *Lambert*, [1975] 2 Lloyd’s Rep. at 491.

⁹² [1978] 2 Lloyd’s Rep. 440, 460 (Q.B.D.).

⁹³ *Id.*

⁹⁴ *Id.*

⁹⁵ [2002] EWHC 203.

allegation is unfounded cannot divest the circumstances of the allegation of the attribute of materiality.⁹⁶

But, having held that the allegation was material, Colman J. nevertheless mitigated his finding by holding that for the insurers to persist at trial in taking the point, in the face of evidence that pointed to the suggested facts being totally false, would be contrary to their obligation of good faith.⁹⁷ It is noteworthy that recently, the Court of Appeal in *North Star Shipping Ltd. v. Sphere Drake Ins. plc* expressed sympathy for Forbes J.'s views and urged the insured to argue that allegations of dishonesty which were unrelated to the risk were immaterial.⁹⁸ Ultimately though, it felt constrained by authority to reject the contention which Waller L.J. stated he otherwise "might be tempted to follow."⁹⁹ However, on the issue of whether the impecuniosity of the insured was a material fact, Waller L.J. stated that, "the non-payment of premium is either material on its own or not, and since it seems to go to the owner's credit risk, and not to the risk insured, I would have thought it was not material."¹⁰⁰ In so finding, Waller L.J. admitted that he was placing a significant limitation on section 18(2) of the 1906 Marine Insurance Act given that this was plainly a material fact which went to the decision of a prudent underwriter whether or not to underwrite the risk.¹⁰¹

With respect to the insurers right of avoidance, the judiciary has also displayed considerable tenacity in its condemnation of the results which necessarily flow from the exercise of the remedy. To take just one recent example, in *Kausar v. Eagle Star Ins. Co. Ltd.*, Staughton L.J. stated:

⁹⁶ *Id.* In *North Star Shipping Ltd. v. Sphere Drake Insurance plc*, [2005] EWHC 665, Colman J. again took this view. *See also* *Brotherton v. Aseguradora Colseguros S.A.* (No. 2) [2003] EWCA Civ. 705 (Mance, L.J.) (discussed *infra* notes 105 and 132); *The Dora*, [1989] 1 Lloyd's Rep. 69, 93-94 (Q.B.D.) (Phillips, J.); *March Cabaret Club & Casino Ltd. v. London Assurance*, [1975] 1 Lloyd's Rep. 169, 175-77 (Q.B.D.). *Cf.* *Norwich Union Ins. Ltd. v. Meisels* [2006] EWHC 2811 (Q.B.) (taking a narrower view of materiality).

⁹⁷ *Strive Shipping Corp.* [2002] EWHC 203. *See infra* note 193.

⁹⁸ [2006] EWCA (Civ.) 378.

⁹⁹ *Id.*

¹⁰⁰ *Id.*

¹⁰¹ *Id.*

Avoidance for non-disclosure is a drastic remedy. It enables the insurer to disclaim liability after, and not before, he has discovered that the risk turns out to be a bad one; it leaves the insured without the protection which he thought he had contracted and paid for...I do consider there should be some restraint in the operation of the doctrine. Avoidance for honest non-disclosure should be confined to plain cases.¹⁰²

However, the weight of the case law and the force of the 1906 Marine Insurance Act inevitably present considerable hurdles to judicial intervention. Nevertheless, the subject of non-disclosure and the insurers' remedy has not escaped the attention of law reform agencies.

A. LEGISLATIVE REFORM: A FALSE DAWN

In 1978, the Law Commission was given the opportunity to review non-disclosure.¹⁰³ This was to be carried out in the light of a proposed EEC Directive on the Co-ordination of Legislative, Statutory and Administrative Provisions relating to Insurance Contracts, the object of which was to harmonise the law in the Community.¹⁰⁴ Of particular concern to the U.K. was the recommendation that the proportionality principle should be adopted. Under French law, for example, an insurer is obliged to pay the proportion of the claim which the actual premium paid bears to the premium which would have been payable if the material facts had been disclosed.¹⁰⁵ In this way, any additional risk and the loss attributable to that additional risk is, in effect, borne by the insured. A more complex set of provisions was adopted in the proposed EEC Directive. Article 3.3(c) of the Proposed Directive dealt with the insurer's right in respect of innocent non-disclosure.¹⁰⁶ It adopted the principle of proportionality only where non-disclosure is due to fault (short of fraud) on the part of the insured.¹⁰⁷

¹⁰² [1997] C.L.C. 129, 132-33.

¹⁰³ See THE LAW COMMISSION, *supra* note 15, § 1.1.

¹⁰⁴ *Id.* § 1.2. In fact, the Directive did not materialise.

¹⁰⁵ CODE DES ASSURANCES, art. L113-9.

¹⁰⁶ See THE LAW COMMISSION, *supra* note 15, § 4.2.

¹⁰⁷ *Id.* § 4.3.

In a case where the non-disclosure is not due to fault, the insurer would remain liable for any loss.¹⁰⁸

The Law Commission concluded that proportionality was unworkable in England and Wales.¹⁰⁹ It stated that the principle gives no guidance as to how the insured's entitlement is to be computed if the insurer would have either declined the risk, imposed additional terms on the insured, narrowed the risk via exclusion clauses or imposed or increased "an excess."¹¹⁰ Further, the Commission found that proportionality provides no solution where knowledge of the undisclosed facts would have led the insurer to decline the risk altogether.¹¹¹ Whatever the merits of this strident dismissal, such hurdles have not prevented the Ombudsman importing the principle into the range of his remedies.¹¹² That aside, the Law Commission did endorse the views expressed by way of *obiter* in *Lambert* to the effect that the law should be changed.¹¹³ It found that the insured's duty of disclosure can give rise to grave injustice and there was, notwithstanding the protestations of the insurance industry, "a formidable case for reform."¹¹⁴ In essence, the Commission recommended a substantially revised duty of disclosure that, had it been implemented, would have resulted in shifting the focus away from the "prudent insurer" as the determinative test of materiality.¹¹⁵ It proposed a modified duty of disclosure for both consumers and businesses whereby an insured would be required to disclose those facts that a reasonable person in the position of the applicant would disclose.¹¹⁶ However, an insured's individual personal characteristics would not be taken into account.¹¹⁷

Despite early optimism that legislative reform would follow the Law Commission's recommendations, this soon petered out. While there was an initial flurry of activity by the DTI (now DBIS), the impetus for

¹⁰⁸ *Id.*

¹⁰⁹ *Id.*

¹¹⁰ *Id.*

¹¹¹ *Id.* § 4.5-4.6.

¹¹² The Ombudsman has adopted proportionality for cases of unintentional non-disclosure and misrepresentation. See the Ombudsman Report for 1989, para 2.16-7) and the Annual Report for 1994, para 2.10. See *infra* text accompanying note 121.

¹¹³ THE LAW COMMISSION, *supra* note 15, § 4.44.

¹¹⁴ See *id.* § 3.23.

¹¹⁵ *Id.* § 6.2.3.

¹¹⁶ *Id.* § 4.47

¹¹⁷ *Id.*

reform ground to a halt no doubt because of the intense lobbying by the insurance industry. Some six years after the Report was laid before Parliament, the Secretary of State for Trade and Industry, Mr. Channon, said, by way of a written response to a question asking for a progress report on the reforms, that he was convinced that self-regulation, through the Statements of Practice issued by the Association of British Insurers, would meet the need of protecting private insureds from the harsher aspects of the disclosure duty.¹¹⁸

¹¹⁸ Paul Channon, Secretary of State, Written Answers (Commons) of 21 February 1986, Insurance Contracts (Feb. 21, 1986) (transcript available at http://hansard.millbanksystems.com/written_answers/1986/feb/21/insurance-contracts). The Secretary of State said,

The insurers have informed me that they are willing to strengthen the non-life and long-term statements of insurance practice on certain aspects proposed by the Department. These concern the limitation of the duty of disclosure, warranties, disputes procedures and, in the case of the long-term statements, the payment of interest on life insurance claims. The statements apply to insurance taken out by private consumers...These changes are in the right direction. I am well aware of the arguments, advanced amongst others by the representatives of consumers, in favour of legislation on non-disclosure and breach of warranty. But I consider that on balance the case for legislation is out-weighted by the advantages of self-regulation so long as this is effective. I look to all insurers, whether or not they belong to the Association of British Insurers which has promulgated the statements, to observe both their spirit and their letter. In the light of the insurers' undertakings I do not consider there is any need for the moment to proceed with earlier proposals for a change in the law....

The Statements of Practice, first issued in 1977 (revised in 1986), covered General and Long Term Insurance. The General Statement came to an end in January 2005 when it was incorporated into the Code of Business Conduct Rules (ICOB) by the Financial Services Authority. The ICOB adopts the language of the ABI's Statements. In essence, rule 7.3.6 provides that except where there is evidence of fraud, an insurer should not avoid a claim by a retail customer on the ground of non-disclosure of a fact material to the risk that the customer could not reasonably be expected to have disclosed. It is also noteworthy that in the field of motor insurance the right of insurers to avoid liability to a third party is substantially restricted by section 152 of the Road Traffic Act 1988. The Long

The move towards self-regulation was reinforced by the industry establishing and financing the Insurance Ombudsman in 1981. Rather than following the common law rules relating to avoidance, the Ombudsman seeks to reach a decision that he considers to be a “fair and reasonable” solution to a dispute.¹¹⁹ In reaching this objective, the criteria taken into account include whether the non-disclosure was deliberate or innocent.¹²⁰ He has also sought to mitigate the draconian consequences of “inadvertent” non-disclosure by, for example, requiring insurers to pay a proportion of a claim that the premium actually fixed bears to the premium that would have been charged had the fact been disclosed.¹²¹ The question remains, however, whether the general law should adopt the same approach, and if so, should the position of the consumer be separated out from that of the commercial insured.

Subsequent investigations have arrived at conclusions very similar to those of the Law Commission. For example, in 1997 the National Consumer Council embarked on a thorough review of the disclosure duty among other areas of insurance law. Its report, *Insurance Law Reform: The Consumer Case for a Review of Insurance Law*, written by Professor Birds, recommended that the consumer-insured’s duty of disclosure should be restricted to facts within his or her knowledge which *either* he or she knows to be relevant to the insurer’s decision *or* which a reasonable person

Term Statement remains in place. Long term insurance is governed by COB which has no equivalent provision to ICOB 7.3.6.

¹¹⁹ See, e.g., INSURANCE OMBUDSMAN, ANNUAL REVIEW 1992/93, ¶¶ 6.48-6.55 (1993). The Insurance Ombudsman is now part of the Financial Ombudsman Service, regulated by Part XVI of the Financial Markets and Services Act, 2000, c. 8, §§ 225-34. The Association of British Insurers Statements will continue to inform the approach of the Ombudsman. See Financial Ombudsman Service, *Non-Disclosure in Insurance Cases*, OMBUDSMAN NEWS, May-June 2005, at 8, available at <http://www.financial-ombudsman.org.uk/publications/ombudsman-news/46/46.pdf>.

¹²⁰ Financial Ombudsman Service, *supra* note 119, at 8.

¹²¹ See INSURANCE OMBUDSMAN, ANNUAL REVIEW 1988/89, ¶ 2.17. See also Financial Ombudsman Service, *Insurance Complaints Involving Non-Disclosure*, OMBUDSMAN NEWS, Apr. 2003, at 9, available at <http://www.financial-ombudsman.org.uk/publications/ombudsman-news/27/27.pdf>. See also JOHN LOWRY & PHILIP RAWLINGS, *INSURANCE LAW: CASES AND MATERIALS* 129-230 (Hart Publishing 2004). See also P.M. North, *Law Reform: Processes and Problems*, 101 L.Q.R. 338, 349 (July 1985); John Birds, *The Reform of Insurance Law*, 1982 J. BUS. L. 449, 450, 454 (Nov. 1982).

in the circumstances could be expected to know to be relevant.¹²² More recently, in January 2001, the British Insurance Law Association established a sub-committee to examine contentious areas of insurance law and to make recommendations to the Law Commission “as to the desirability of drafting a new Insurance Contracts Act.”¹²³ Its report, *Insurance Contract Law Reform*, published in September 2002, also endorsed the reasonable insured test, for example, whether a reasonable insured would have considered the undisclosed matter to be material to a prudent insurer.¹²⁴ As will be seen, the most recent report by the English and Scottish Law Commissions broadly follows this proposal.¹²⁵

Although the legislature has not responded to these calls for reform, the issue has not escaped the attention of the courts. Recent case law suggests that there is a distinct shift in the judicial focus and that the attention of the judges is being channelled along several lines of investigation. For example, particular attention is being directed towards the requirement of inducement as a determinant of non-disclosure together with a wider-vised approach being adopted towards the role of the insurer during the disclosure process.¹²⁶ As commented above, this may be seen as adding content to the insurers’ duty of good faith and in this regard, attention is now also being directed towards the exercise of the remedy of avoidance.

IV. JUDICIAL INTERVENTION: REDRESSING THE BALANCE

The opportunity for an authoritative review of the insured’s duty of disclosure came before the House of Lords just over ten years ago in *Pan Atlantic Insurance Co. Ltd. v. Pine Top Insurance Co.*¹²⁷ The defendant

¹²² John Birds *Insurance Law Reform: the Consumer Case for a Review of Insurance Law*, National Consumer Council, 1997.

¹²³ *Id.*

¹²⁴ British Insurance Law Association, *Insurance Contract Law Reform* (Sept. 2002).

¹²⁵ See THE LAW COMMISSION, *supra* note 17.

¹²⁶ [1995] 1 AC 501.

¹²⁷ *Pan Atl. Ins. Co. Ltd. v. Pine Top Ins. Co. Ltd.*, [1995] 1 A.C. 501, 505. Both parties were insolvent by the time the case reached the appellate courts. *Id.* However, it proceeded as a friendly action because of the perceived loss of business being suffered by the UK, and London in particular, as a result of the overly insurer-friendly approach being adopted in relation to non-disclosure. *Id.* It was hoped that the House of Lords would redress the balance. *Id.*

reinsurers had written excess of loss policies for three years.¹²⁸ Their defence to a claim arising out of losses suffered in the third year was based on the inadvertent failure to disclose the extent of losses occurring in the first two years.¹²⁹ The issues were first, should materiality be measured by reference to whether its ‘influence’ on the prudent insurer’s judgment was ‘decisive’, or should some lesser degree of impact be sufficient?¹³⁰ Second, where there has been non-disclosure of a material fact, must it induce the actual insurer to enter into the contract?¹³¹

With respect to the first issue Lord Mustill, with whom Lords Goff and Slynn concurred, could see no good reason for departing from the principle which had guided insurance law for more than 200 years.¹³² Lord Mustill stated that disclosure was not limited to matters which would have caused the prudent insurer to decline the risk or increase the premium but rather the insured’s duty to disclose “all matters which would have been taken into account by the underwriter when assessing the risk . . . which he was consenting to assume.”¹³³ On the question of statutory interpretation, the majority view was that since Parliament had left the word “influence” in section 18(2) unadorned by phrases such as “decisively” or “conclusively,” it must bear its ordinary meaning.¹³⁴ His Lordship stated that “. . . this expression clearly denotes an effect on the thought processes of the insurer in weighing up the risk, quite different from words which might have been used but were not, such as ‘influencing the insurer to take the risk.’”¹³⁵ The majority decision therefore was to reject the “decisive influence” test. The position remains that a circumstance is material and must be disclosed even though the prudent insurer, had he known of the fact, would have insured the risk on the same terms.¹³⁶

¹²⁸ *Id.* at 519.

¹²⁹ *Id.* at 520.

¹³⁰ *Id.* at 516-17.

¹³¹ *Id.* at 517-18.

¹³² *Id.* at 536.

¹³³ *Pan Atl. Ins. Co. Ltd.*, [1995] 1 A.C. at 538. Lord Mustill thus rejected a test based upon the decisive influence of the non-disclosed/misrepresented fact: “I can see no room within [the principle] for a more lenient test expressed solely by reference to the decisive effect which the circumstance would have on the mind of the prudent underwriter.” *Id.* at 536.

¹³⁴ *Id.* at 531.

¹³⁵ *Id.* See also *id.* at 517 (speech of Lord Goff).

¹³⁶ It is noteworthy that Lord Lloyd, in a powerful dissent, agreed with the appellants’ submission that there should be a twofold test under which the insurer

In relation to the second issue, however, the House of Lords unanimously held that in *Pan Atlantic Insurance Company v. Pine Top Limited*, the non-disclosure of a material fact, as with misrepresentation, must induce the particular insurer to enter into the contract.¹³⁷ In reaching this conclusion, their Lordships were clearly influenced by the argument that the 1906 Act codified the common law, and given that inducement was a requirement under the general law which provides for rescission of a contract, the Act must be taken as having the same effect.¹³⁸ In language that resonates with that of Lord Mansfield's in so far as it traverses the terrain of misrepresentation and non-disclosure (and in so doing aligns the requirement of inducement with both vitiating factors), Lord Mustill stated that:

I conclude that there is to be implied in the Act of 1906 a qualification that a material misrepresentation will not entitle the underwriter to avoid the policy unless the misrepresentation induced the making of the contract, using "induced" in the sense in which it is used in the general law of contract.¹³⁹

must show that a prudent insurer, if aware of the undisclosed fact, would either have declined the risk or charged a higher premium and that the actual insurer would have declined the risk or required a higher premium. *Id.* at 554. *See also* John Birds & Norma J. Hird, *Misrepresentation and Non-Disclosure in Insurance Law - Identical Twins or Separate Issues*, 59 M.L.R. 285, 285 (1996).

¹³⁷ *Pan Atl. Ins. Co.*, [1995] 1 A.C. at 551. In essence, the House of Lords were injecting into the law on non-disclosure a requirement of causation analogous to the "but for" test familiar to tort lawyers. *See id.* at 551 (Lord Mustill's reference to causative effect). In his reasoning, Lord Mustill gave prominence to the decision of Kerr J. in *Berger v. Pollock*, [1973] 2 Lloyd's Rep. 442, in which the judge stated the principles in a way that suggested that the insurer could avoid the policy only if he had in fact been influenced by the non-disclosure. *Id.* at 463.

¹³⁸ *Pan Atl. Ins. Co.*, [1995] 1 A.C. at 549.

¹³⁹ *Id.* Where there is a material misrepresentation, there is a rebuttable presumption of inducement. *See* *Redgrave v. Hurd*, (1881) 20 Ch.D. 1, 21; *Smith v. Chadwick*, (1884) 9 App. Cas. 187, 196. Lord Mustill went on to add that, "As a matter of common sense however even where the underwriter is shown to have been careless in other respects the assured will have an uphill task in persuading the court that the withholding or misstatement of circumstances satisfying the test of materiality has made no difference." *Pan Atl. Ins. Co.*, [1995] 1 A.C. at 551. *See also* *Svenska Handelsbanken v. Sun Alliance & London Ins. plc*, [1996] 1 Lloyd's Rep. 519.

Lord Goff, concurring, thought that the need to show inducement on the part of the actual insurer addresses the criticisms directed against the harshness of the duty.¹⁴⁰ He reasoned that it was the absence of this requirement that prompted the call for the test of materiality to “be hardened into the decisive influence test.”¹⁴¹ However, this concession must be measured against the view of Lord Mustill, vigorously opposed by Lord Lloyd, that there should be a presumption of inducement.¹⁴² In essence once objective materiality is established a presumption that the actual insurer was induced triggers. On the facts, the House of Lords held that the non-disclosed losses were so obviously material that inducement could be inferred.¹⁴³

The presumption of inducement has had a chequered reception in the case law following *Pan Atlantic*.¹⁴⁴ From the perspective of the insured, modern case law has sought to preserve the benefit of the requirement by limiting the scope of its presumption to exceptional cases only. For example, in *Marc Rich & Co. A.G. v. Portman*, Longmore J. suggested that unless there was good reason for the underwriter not to give evidence, the presumption would simply not arise.¹⁴⁵ The judge stressed

¹⁴⁰ *Pan Atl. Ins. Co.*, [1995] 1 A.C. at 516-18.

¹⁴¹ *Id.* at 518.

¹⁴² *Id.* at 542, 571.

¹⁴³ *Id.* at 562.

¹⁴⁴ The only cases in which the presumption of inducement has been applied are those involving market subscriptions when one member of the following market has been unavailable to give evidence of his own state of mind. *See, e.g.*, *Talbot Underwriting Ltd. v. Nausch, Hogan & Murray Inc.* [2006] EWCA Civ. 889, [2006] 2 Lloyd’s Rep. 195; *Toomey v. Banco Vitalicio de Espana SA de Seguros y Reaseguros* [2003] EWHC 1102, [2004] Lloyd’s Rep. I.R. 354; *St. Paul Fire & Marine Co. (U.K.) Ltd. v. McConnell Dowell Constructors Ltd.*, [1996] 1 All E.R. 96, [1995] 2 Lloyd’s Rep. 116. *See also* *Sirius Int’l Ins. Group Corp. v. Oriental Assurance Corp.*, [1999] 1 All E.R. (Comm.) 699, [1999] Lloyd’s Rep. I.R. 343; *Ins. Corp. of the Channel Islands v. Royal Hotel Ltd.*, [1998] Lloyd’s Rep. I.R. 151.

¹⁴⁵ [1996] 1 Lloyd’s Rep. 430. *See also* *Sirius Int’l Ins. Group Corp.*, [1999] 1 All E.R. (Comm.) 699, where, in relation to misrepresentation, Longmore J. also stressed that it is for the insurer to prove inducement. The judge did, however, recognize that the onus of proof is difficult to discharge. *Id.* In his Pat Saxton Memorial Lecture, “An Insurance Contracts Act for a New Century”, delivered on 5 March 2001 to the British Insurance Law Association, Longmore J gave death as an example of a good reason for failing to give evidence. Sir Andrew Longmore, Pat Saxton Memorial Lecture at the British Insurance Law Association: An

that in cases where the court is in doubt, the defence of non-disclosure should fail because “[a]t the end of the day it is for the insurer to prove that the non-disclosure did induce the writing of the risk...”¹⁴⁶ Further, in *Assicurazioni Generali SpA v. Arab Ins. Group*, the Court of Appeal took the view that although the non-disclosed (or misrepresented) fact need not be the sole inducement operating on the insurer, it must cause the actual insurer to enter into the contract.¹⁴⁷ Significantly, the majority of the court followed earlier decisions to the effect that the insurer must give evidence as to his state of mind.¹⁴⁸ This, therefore, gives the insured the opportunity to cross-examine the insurer with a view to demonstrating that he was not induced by the non-disclosed fact but would have entered into the contract on the same terms had there had been full disclosure of all material facts.¹⁴⁹ If the underwriter fails to give evidence, without "good reason", inducement will not be made out.¹⁵⁰ Clarke L.J. summarised the position as follows:

1. In order to be entitled to avoid a contract of insurance or reinsurance, an insurer... must prove on the balance of probabilities that he was induced to enter into the contract by a material non-disclosure or by a material misrepresentation.
2. There is no presumption of law that an insurer... is induced to enter in the contract by a material non-disclosure or misrepresentation.
3. The facts may, however, be such that it is to be inferred that the particular insurer... was so induced even in the absence from evidence from him.

Insurance Contracts Act for a New Century? (Mar. 5, 2001), available at http://www.bila.org.uk/lecture_scripts/lectsaxt.asp.

¹⁴⁶ *Marc Rich & Co.*, [1996] 1 Lloyd's Rep. at 442.

¹⁴⁷ [2002] EWCA Civ. 1642, [2003] 1 W.L.R. 577. Following the decision in *St. Paul's Fire & Marine Ins. Co. v. McConnell Dowell Constructors Ltd.*, [1995] 1 All E.R. 96, 104.

¹⁴⁸ *Assicurazioni Generali SpA* [2002] EWCA Civ. 1642.

¹⁴⁹ *Id.*

¹⁵⁰ *Marc Rich & Co. AG*, 1 Lloyd's Rep. at 442.

4. In order to prove inducement the insurer or reinsurer must show that the non-disclosure or misrepresentation was an effective cause of his entering into the contract on the terms on which he did. He must therefore show at least that, but for the relevant non-disclosure or misrepresentation, he would not have entered into the contract on those terms. On the other hand, he does not have to show that it was the sole effective cause of his doing so.¹⁵¹

Insurers must, therefore, establish that the non-disclosed or misrepresented fact was an effective cause, although not necessarily the only cause, of their agreement to underwrite the risk. The link between materiality and inducement has thus been severed. Where the underwriter does give evidence, he will need to demonstrate a causal link between the presentation of the risk and its acceptance.

Further inroads have recently been made into the notion of utmost good faith. This has been done by first adopting a narrow, insured friendly approach towards the requirement of inducement, one incident of which has led to the courts to re-examine the defence of waiver; and second, by aligning the exercise of the remedy of avoidance for non-disclosure with the insurers' duty of good faith. In tandem with this process, the courts have also been examining the role of insurers as recipients of information during the disclosure process. The effect is that the burden imposed by the duty is being recalibrated so as to strike some balance between the respective obligations of the parties especially at the stage when the risk is being presented for underwriting. A further and significant recent development relates to the vexed question of whether the insured's duty of good faith continues after the insurance contract has been concluded so that it again triggers when a claim is made under the policy.

A. REFINING INDUCEMENT

The emphasis now being placed on the need to demonstrate inducement has been bolstered by the radical step taken by the court in *Drake Insurance plc v. Provident Insurance plc*, to the effect that, in deciding whether the non-disclosed fact had induced the insurer to enter the contract, it is necessary to examine what would have happened had full

¹⁵¹ [2003] 1 W.L.R. 577.

disclosure been made.¹⁵² In this case, the insurer sought to avoid the motor policy on the basis that it would have charged a higher premium had a speeding conviction been disclosed.¹⁵³ The insured had disclosed an earlier fault accident, but failed to disclose before the present policy was concluded that it had been reclassified as a no-fault accident.¹⁵⁴ The majority of the Court of Appeal held that, even if the conviction had been disclosed, information would have come to light that the earlier accident had not been the insured's fault and this would have resulted in the proposal being accepted at a normal rate of premium.¹⁵⁵

Rix L.J., delivering the leading judgment,¹⁵⁶ stated that the issue is not what actually happened, but what would have happened had the speeding conviction been declared. To prove inducement, the insurer, Provident, would need to show that a higher premium would have resulted.¹⁵⁷ This it could not do because it was common ground that it would not have increased the premium if the earlier accident had been no-fault: "So the question resolves itself into this: if the conviction had been mentioned, would the question of the status of the accident have been discussed? It seems to me to be very likely that it would have been..."¹⁵⁸

It is noteworthy that Rix L.J. went on to express the view that he could see no reason in principle why an insured should not be able to rely on facts which would have been material in his favour had they been disclosed.¹⁵⁹ This, after all, is the case with insurers and the logic is, therefore, compelling. Further, this reasoning marks a clear departure from the view expressed by Mance L.J. in *Brotherton v. Aseguradora Colseguros SA*,¹⁶⁰ to the effect that an insured is not entitled to prove what the true position was at the time the contract was concluded as a means of proving that a particular fact was immaterial.¹⁶¹

¹⁵² [2003] EWCA Civ. 1834; [2004] Q.B. 601.

¹⁵³ *Id.*

¹⁵⁴ *Id.*

¹⁵⁵ *Id.*

¹⁵⁶ *Id.* Clarke L.J. agreeing and Pill L.J. dissenting.

¹⁵⁷ *Id.*

¹⁵⁸ *Drake Ins.* [2003] EWCA Civ. 1834.

¹⁵⁹ *Id.*

¹⁶⁰ [2003] EWCA Civ. 705. See also Malcolm Clarke, *Non-disclosure and Avoidance: Lies, Damned Lies, and "Intelligence"* [2004] L.M.C.L.Q. 1.

¹⁶¹ *Brotherton* [2003] EWCA Civ. 705 (rejecting the view of Colman J. in *Strive Shipping Corp. v. Hellenic Mut. War Risks Ass'n (Bermuda) Ltd., (The Grecia Express)*, [2002] EWHC 203, 213, that "an insured is, if necessary, entitled

Overall, the approach taken towards inducement by Rix L.J. has the effect of the court placing itself in the position of the underwriter. The result is that the court is prepared to reopen the negotiations between the parties, certainly in relation to disclosure, and to speculate on their likely responses.¹⁶² In effect, the court is assessing what the underwriter's most likely course of action would have been with full disclosure. The onus is on the underwriter to prove that it would not have accepted the risk either at all or on the premium actually charged. Given this proactive position being taken towards the issue, there seems no reason in principle why the judges should not also be able to apply the proportionality doctrine, long harnessed by the Ombudsman, rather than continue with the all or nothing approach of avoidance. The approach adopted by Rix L.J. towards the determination of the particular risk in question thus renders the Law Commission's reasoning, in its 1980 report rejecting the proportionality doctrine as "unworkable," less than compelling.

B. THE PRESENTATION OF THE RISK: A PRO-ACTIVE ROLE FOR INSURERS?

This renewed focus on the requirement of inducement can be seen as part of the overarching anxiety that the presentation of the risk should be fair. Here, the mutuality of the good faith duty has come to the fore and the judges have been directing their attention towards the content of the insurers' obligation. It will be recalled that in *Carter v. Boehm*, Lord Mansfield laid particular emphasis on both the need for a fair presentation of risk and the limits of the insured's disclosure duty.¹⁶³ In this respect, he excluded from the realms of the duty those facts which the insurer "waives being informed of" together with facts the insurer is presumed to know.¹⁶⁴

to litigate the issue of the truth or falsity of known but undisclosed intelligence, in order to argue that, if it is shown to be incorrect, the insurer would be acting in bad faith or unconscionably in avoiding.")

¹⁶² See R. Merkin, 16 *Insurance Law Monthly* (2004). See also *Bonner v. Cox Dedicated Corporate Member* [2004] EWHC 2963 (Comm).

¹⁶³ *Carter v. Boehm*, (1766) 97 Eng. Rep. 1162, 1164-65.

¹⁶⁴ *Id.* at 1165. Lord Mansfield explained that the insured need not disclose

... what the underwriter knows ...what way soever he came to the knowledge. The insured need not mention what the underwriter ought to know; what he takes upon himself the knowledge of; or what he waives being informed of. The underwriter needs

It will be recalled that Lord Mansfield returned to the role expected of underwriters during the disclosure process in *Noble v. Kennoway*, in which he held that the insurer was under a duty to inform himself of the practices of the trade he insures.¹⁶⁵ Further, in *Court v. Martineau*, he was prepared to draw the inference that the insurer had waived the disclosure of certain facts by the large premium he charged for underwriting the risk in question.¹⁶⁶

Opportunities to consider the insurers' duty of utmost good faith have been rare in modern times. However, in *Banque Keyser Ullman S.A. v. Skandia (U.K.) Insurance (C.A.)*, Slade L.J. said that the insurers' duty of disclosure should

... extend to disclosing facts known to him which are material either to the nature of the risk sought to be covered or the recoverability of a claim under the policy which a prudent insured would take into account in deciding whether to place the risk for which he seeks cover with that insurer.¹⁶⁷

The House of Lords approved Slade L.J.'s reasoning in this respect. The only remedy available to the insured where the insurer is in breach of duty is avoidance *ab initio*. In practice, this affords little or no benefit to insureds. An insurer's breach will come to light when the loss has been suffered – a time when an insured will want full recovery rather than a return of the premium.

More recently, however, the issue has come to the fore in the context of determining the insurer's role during the disclosure process

not be told what lessens the risque agreed and understood to be run by the express terms of the policy. He needs not to be told general topics of speculation: as for instance - the under-writer is bound to know every cause which may occasion natural perils....

Id. This particular element of the judgment was later codified, virtually verbatim, in the Marine Insurance Act, 1906, 6 Edw. 7, c. 41, § 18(3) (Eng.).

¹⁶⁵ (1780) 99 Eng. Rep. 326. *See also* *Mayne v. Walter*, (1782) 99 Eng. Rep. 548.

¹⁶⁶ (1782) 99 Eng. Rep. 591. *See also* *Drake Ins. plc v. Provident Ins. plc*, [2004] Q.B. 601 (reasoning of Rix L.J.).

¹⁶⁷ [1990] 1 Q.B. 665 (A.C.) at 772. *See also* *Aldrich v. Norwich Union Life Ins. Co. Ltd.*, [1999] 2 All E.R. (Comm) 707.

where the insured has raised the defence of waiver. Of significance in this respect are the views expressed by the Court of Appeal in *WISE Ltd. v. Grupo Nacional Provincial SA*.¹⁶⁸ The issue arose in the context of commercial insurance. The defendant, Mexican insurer GNP, appealed to the Court of Appeal against a decision of Simon J. that the claimants-reinsurers WISE were entitled to avoid a reinsurance contract on the basis that the presence of high-value Rolex watches in the insured consignment of goods was not disclosed.¹⁶⁹ This occurred as a result of a translation error in which the watches were described as clocks.¹⁷⁰ It was held, by a majority, that GNP was entitled to recover.¹⁷¹ Although it was unanimously held that WISE had been induced by the presentation of the risk, Rix and Peter Gibson L.J.J. held that the reinsurers had affirmed the policy, notwithstanding the breach of the duty of disclosure, by giving notice of its cancellation.¹⁷² Such notice was inconsistent with any claim to avoid the policy *ab initio*.¹⁷³ Both judges took the view that the trial judge had overlooked a vital email which showed that WISE were unequivocal in cancelling the policy.¹⁷⁴

With respect to the issue of waiver, the parties agreed that the law was accurately set out in *MacGillivray on Insurance Law*,¹⁷⁵ which, citing *CTI v. Oceanus Mutual Underwriting Association (Bermuda) Ltd.*,¹⁷⁶ states:

The assured must perform his duty of disclosure properly by making a fair presentation of the risk proposed for insurance. If the insurers thereby receive information from the assured or his agent which, taken into conjunction with other facts known to them or which they are presumed to know, would naturally prompt a reasonably careful insurer to make further inquiries, then, if they omit to make the appropriate check or inquiry, assuming it can be made

¹⁶⁸ [2004] EWCA Civ. 962.

¹⁶⁹ *Id.*

¹⁷⁰ *Id.*

¹⁷¹ *Id.*

¹⁷² *Id.*

¹⁷³ *Id.*

¹⁷⁴ *WISE Ltd.* [2004] EWCA Civ. 962. Longmore L.J. dissented on the basis that the judge's findings of fact could not be reversed. *Id.*

¹⁷⁵ NICHOLAS LEIGH-JONES ET AL., *MACGILLIVRAY ON INSURANCE LAW* (10th ed. 2003).

¹⁷⁶ [1984] 1 Lloyd's Rep. 476.

simply, they will be held to have waived disclosure of the material fact which that inquiry would necessarily have revealed. Waiver is not established by showing merely that the insurers were aware of the possibility of the existence of other material facts; they must be put fairly on inquiry about them.¹⁷⁷

Longmore L.J., with whom Peter Gibson L.J. agreed, took the view that since the carriage of Rolex watches was a material fact which was not disclosed, the presentation of the risk was unfair.¹⁷⁸ The issue was, therefore, whether the facts that were disclosed would prompt a reasonably careful insurer to enquire whether watches were included in the shipment. As Longmore LJ explained, the issue came down to whether or not the insurer was “put on inquiry by the disclosure of facts which would raise in the mind of the reasonable insurer at least the suspicion that there were other circumstances which would or might vitiate the presentation.”¹⁷⁹ On the facts he held that there was nothing in the presentation of the risk that could be said to have raised the suspicion that Rolex watches were to be included in the consignment.¹⁸⁰

Of particular interest for present purposes is Rix L.J.’s dissenting judgment on this issue. In finding that there had been waiver, he placed particular emphasis on the mutuality of the duty of utmost good faith, and stated that the only relevant question was whether the presentation was fair.¹⁸¹ This could not be judged in isolation, although an obviously unfair presentation would rarely leave room for waiver to operate.¹⁸² The insurers’ reaction and the issue of possible waiver had to be taken into account.¹⁸³ The question is not whether an “unfair” presentation had been waived but whether, taking both sides into account, the presentation was unfair or, alternatively, it would be unfair of the insurers to avoid the contract on a ground on which they were put on inquiry and should have satisfied themselves by making appropriate enquiries.¹⁸⁴

¹⁷⁷ LEIGH-JONES ET AL., *supra* note 175, ¶¶ 17-83.

¹⁷⁸ *WISE Ltd.* [2004] EWCA Civ. 962.

¹⁷⁹ *Id.*

¹⁸⁰ *Id.*

¹⁸¹ *Id.*

¹⁸² *Id.*

¹⁸³ *Id.*

¹⁸⁴ *WISE Ltd.* [2004] EWCA Civ. 962.

Ultimately, it seems, the question is: Has the insurer been put fairly on inquiry about the existence of other material facts, which such inquiry would necessarily have revealed? The test has to be applied by reference to a reasonably careful insurer rather than the actual insurer, and not merely by reference to what such an insurer is told in the assured's actual presentation but also by reference to what he knows or ought to know, *ie.* his s 18(3)(b) [of the 1906 Act] knowledge... Overriding all, however, is the notion of fairness, and that applies mutually to both parties, even if the presentation starts with the would-be assured.¹⁸⁵

Rix L.J. therefore concluded that a reasonably careful insurer would have been fairly put on inquiry given what he knew from GNP's presentation and his general, presumed knowledge.¹⁸⁶ The question as to what types of clocks were being transported was one that should have been asked by the reinsurers. He went on to state that:

If the question is instead the overriding question: Is the ultimate assessment of GNP's presentation that it is unfair, or would it be unfair to allow the reinsurers a remedy of avoidance in such a case? I would answer that the presentation was fair, and that it would be unfair to allow reinsurers to take advantage of an error of translation in a case where, on the evidence, an exclusion of watches would seem to have been the obvious solution.¹⁸⁷

For Rix L.J., the duty of utmost good faith and, more particularly, its content insofar as it applies to insurers or reinsurers, requires them to play a pro-active role in the disclosure process rather than relying solely upon the insured's presentation.¹⁸⁸ In this respect, his approach resonates with that taken by Lord Mansfield in *Noble v. Kennoway*,¹⁸⁹ and *Court v.*

¹⁸⁵ *Id.*

¹⁸⁶ *Id.*

¹⁸⁷ *Id.*

¹⁸⁸ See Merkin, *supra* note 162.

¹⁸⁹ (1780) 99 Eng. Rep. 326.

Martineau.¹⁹⁰ Further, the clear implication seems to be that given that the insurers or reinsurers draft the policy terms; there is more than adequate opportunity for them to take the necessary steps to protect themselves in relation to the risks to be underwritten.

C. ALIGNING THE INSURER'S REMEDY OF AVOIDANCE WITH THE DUTY OF UTMOST GOOD FAITH

We have seen that the consequence of non-disclosure is to render the insurance contract voidable, thereby entitling the insurer to avoid it *ab initio*.¹⁹¹ Any premium paid is returnable to the insured except in cases of fraud (unless the policy otherwise provides).¹⁹² Not surprisingly, as with the requirements of inducement and waiver, the conditions governing the exercise of the avoidance remedy have also attracted considerable judicial attention of late. Again, the views expressed have not been entirely consistent. In *The Grecia Express*, Colman J. suggested that the right to avoid is conditional upon the insurer acting with "duty of the utmost good faith."¹⁹³ He also reasoned, as commented above, that an insured is entitled to litigate the issue of the truth or falsity of material circumstances in order to argue that, if it is shown to be incorrect, the insurer would be acting in bad faith or unconscionably in avoiding the policy.¹⁹⁴ In a similar vein, in *Manifest Shipping Co. Ltd. v. Uni-Polaris Shipping Co. Ltd.*,¹⁹⁵ Lord Hobhouse, delivering the leading speech, identifies the need for some balance to be struck between the parties in the post-contract situation and suggests, as did Lord Mansfield in *Carter v. Boehm*, that the courts should guard against the danger of the good faith duty being turned into an

¹⁹⁰ (1782) 99 Eng. Rep. 591.

¹⁹¹ *Abram Steamship Co. v. Westville Shipping Co.*, [1923] A.C. 773, 781. *See also* *Glasgow Assurance Corp. v. Symondson & Co.*, (1911) 16 Com. Cas. 109, 121 (Scrutton J., suggesting that the only remedy available for non-disclosure is avoidance of the contract).

¹⁹² *PARK*, *supra* note 18, at 218. The Marine Insurance Act, 1906, 6 Edw. 7, c. 41, § 83(3)(a) (Eng.) provides, "Where the policy is void, or is avoided by the insurer, as from the commencement of the risk, the premium is returnable, provided that there has been no fraud or illegality on the part of the assured . . ."

¹⁹³ *Strive Shipping Corp. v. Hellenic Mut. War Risks Ass'n (Bermuda) Ltd. (The Grecia Express)* [2002] EWHC 203.

¹⁹⁴ *Id.*

¹⁹⁵ [2003] 1 A.C. 469, 496-97.

instrument permitting unconscionable behaviour on the part of underwriters. He wrote:

The courts have consistently set their face against allowing the assured's duty of good faith to be used by the insurer as an instrument for enabling the insurer himself to act in bad faith. An inevitable consequence in the post-contract situation is that the remedy of avoidance of the contract is in practical terms wholly one-sided. It is a remedy of value to the insurer and, if the defendants' argument is accepted, of disproportionate benefit to him; it enables him to escape retrospectively the liability to indemnify which he has previously and (on this hypothesis) validly undertaken.¹⁹⁶

Against this, a rather more rigid view was taken towards the exercise of the remedy in *Brotherton v. Aseguradora Colseguros SA (No. 2)*.¹⁹⁷ Mance L.J. explained that the right to avoid is a self-help remedy that is exercised without the court's authorisation.¹⁹⁸ He stated that avoidance for non-disclosure is to be treated in the same way as rescission for misrepresentation under the general law of contract, which is "by act of the innocent party operating independently of the court."¹⁹⁹ In short, the court at trial cannot reverse a valid avoidance. This affirms the orthodoxy that holds that an insurer has an unfettered discretion to avoid the contract in cases where there has been a breach of the duty of disclosure even where the facts relied on, which in this case concerned allegations going as to moral hazard, turn out to be unfounded.

Notwithstanding this strict stance, the courts have continued to subject the conditions governing the right of avoidance to scrutiny and have suggested that the good faith duty triggers whenever underwriters seek to exercise the remedy. In his far reaching judgment delivered in *Drake Insurance plc v. Provident Insurance plc*, Rix L.J. observed that the doctrine of good faith should be capable of limiting the insurer's right to

¹⁹⁶ *Id.* at 497.

¹⁹⁷ [2003] EWCA Civ. 705.

¹⁹⁸ *Id.*

¹⁹⁹ *Id.* Note the criticism of this reasoning by Malcolm Clarke, *Rescission: Inducement and Good Faith*, CAMBRIDGE L.J. 286, 287 (2004). See generally Peter Macdonald Eggers, *Remedies for the Failure to Observe the Utmost Good Faith*, L.M.C.L.Q. 249, 262-71 (2003).

avoid in circumstances where that remedy, “which has been described in recent years as draconian,” would operate unfairly.²⁰⁰ He went on to note that in recent years, there has been a realisation that in certain respects English insurance law has developed too stringently.²⁰¹ Citing *Pan Atlantic*, Rix L.J. stated that leading modern cases show that the courts are prepared to introduce safeguards and flexibility.²⁰² Importantly, he said that it would not be in good faith to avoid a policy without first allowing the insured an opportunity to address the reason for the avoidance.²⁰³ He concluded by stressing that not all insurance contracts are made by those engaged in commerce and the widespread nature of consumer insurance presented new problems.²⁰⁴ “It may be necessary to give wider effect to the doctrine of good faith and recognize that its impact may demand that ultimately regard must be had to a concept of proportionality implicit in fair dealing.”²⁰⁵

Turning to the mutuality of the duty of utmost good faith, the Court of Appeal sought to refine the insurer’s duty further. Rix and Clarke L.J.J. took the view that if the insurer had actual knowledge or blind-eye knowledge of the fact that the accident was “no-fault,” it would have been a matter of bad faith had the insurer avoided the policy.²⁰⁶ Rix L.J. left the question open whether “something less than such knowledge would have been enough to qualify an unrestricted right to avoid.”²⁰⁷ Pill L.J., however, discusses blind-eye knowledge and points out that there must be a suspicion that relevant facts exist and a deliberate decision not to make an

²⁰⁰ [2003] EWCA Civ. 1834, [2004] Q.B. 601, 628. For criticism of the court’s finding that the insurer’s right of avoidance was subject to good faith, see Neil Campbell, *Good Faith: Lessons from Insurance Law*, 11 N.Z. BUS. L.Q. 479 (2005). It was, of course, inevitable that the Court of Appeal in *North Star Shipping Ltd. v. Sphere Drake Ins. plc*, [2006] EWCA Civ. 378, did not permit the amendment to the notice of appeal and, therefore, did not have the opportunity to comment on this aspect of *Drake*.

²⁰¹ *Drake Ins.* [2003] EWCA Civ. 1834.

²⁰² *Id.*

²⁰³ *Id.*, [2004] Q.B. at 628, 630.

²⁰⁴ *Id.* at 629.

²⁰⁵ *Id.*

²⁰⁶ *Id.*

²⁰⁷ *Drake Ins.*, [2004] Q.B. at 630. Clarke L.J., denied the existence of “a general principle that insurers must always give the insured an opportunity to address the reason why they are considering avoidance.” *Id.* at 642.

enquiry.²⁰⁸ He goes on to state, “failure to make any enquiry of the insured before taking the drastic step of avoiding the policy was in my judgment a breach by the insurer of the duty of good faith.”²⁰⁹ While he concluded that they did not have blind-eye knowledge, nevertheless, he took the view that they were under a duty of good faith to inform the insured of their intention to avoid the policy and to give him an opportunity to update them with respect to the accident.²¹⁰

It is noteworthy that the issue of the (re)insurer’s good faith duty has recently been considered in relation to express terms contained in the insurance policy. In *Gan Insurance Co. Ltd. v. Tai Ping Insurance Co. Ltd.*, it was held that claims co-operation clauses are subject to a rationality test which owes its origins to the insurers’ duty of good faith.²¹¹ Although there was no implied term that approval of a settlement could not be unreasonably withheld, the right to withhold approval was not unqualified.²¹² It must be exercised in good faith.²¹³ Thus, (re)insurers are under a duty of good faith in exercising their rights under a claims co-operation clause, and must not, therefore, arbitrarily refuse to approve a settlement.

IV. REINING IN THE NOTION OF THE INSURED’S POST-CONTRACTUAL GOOD FAITH DUTY

The move away from the position permitting an unfettered right of avoidance must also be viewed against the wider landscape in which the nature and scope of the insured’s post-contractual duty of good faith has similarly been tested by the courts. Until recently, the judicial consensus was that the insured’s duty of good faith revived in appropriate circumstances during the currency of the contract.²¹⁴ In this regard the

²⁰⁸ *Id.* at 649.

²⁰⁹ *Id.*

²¹⁰ *Id.*

²¹¹ [2001] EWCA Civ. 1047 (Mance L.J.). *See also* *Eagle Star Ins. Co. Ltd. v. Cresswell* [2004] EWCA Civ. 602.

²¹² *Gan Ins. Co. Ltd.* [2001] EWCA 1047.

²¹³ *Id.*

²¹⁴ *Manifest Shipping Co. Ltd. v. Uni-Polaris Shipping Co. Ltd.* [2003] 1 A.C. 469 (Leggatt L.J.) (approving the judgment of Rix J. in *Royal Boskalis Westminster NV v. Mountain*, [1999] Q.B. 674, in which the judge cited the following passage from MALCOLM A. CLARKE, *THE LAW OF INSURANCE CONTRACTS* 708 (2d ed. 1994):

position was stated Mathew L.J. in *Boulton v. Houlder Bros. & Co.*, that it “is an essential condition of the policy of insurance that the underwriters shall be treated with good faith, not merely in reference to the inception of the risk, but in the steps taken to carry out the contract.”²¹⁵ The underlying rationale for this view was explained by Hoffmann L.J. in *Orakpo v. Barclays Ins. Services Co. Ltd.*²¹⁶

I do not see why the duty of good faith on the part of the assured should expire when the contract has been made. The reasons for requiring good faith continue to exist. Just as the nature of the risk will usually be within the peculiar knowledge of the insured, so will the circumstances of the casualty; it will rarely be within the knowledge of the insurance company. I think that the insurance company should be able to trust the assured to put forward a claim in good faith.²¹⁷

Sir Roger Parker agreed with Hoffmann L.J..²¹⁸ There the Court of Appeal held that a claim which is fraudulent entitles the insurer to avoid the contract *ab initio* irrespective of whether there is a term in the policy to that effect.²¹⁹ However, Staughton L.J. differed.²²⁰ While he thought this

As regards insurance contracts, the duty of good faith continues throughout the contractual relationship at a level appropriate to the moment. In particular, the duty of disclosure, most prominent prior to contract formation, revives whenever the insured has an express or implied duty to supply information to enable the insurer to make a decision. Hence, it applies if cover is extended or renewed. It also applies when the insured claims insurance money: he must make ‘full disclosure of the circumstances of the case’ [citing *Shepherd v. Chewter*, (1808) 1 Camp. 274, 275 (Lord Ellenborough)].

²¹⁵ [1904] 1 KB 784, 791-92. See also *Shepherd v. Chewter*, (1808) 170 Eng. Rep. 955, 957.

²¹⁶ [1994] C.L.C. 373 (A.C.).

²¹⁷ *Id.* at 383.

²¹⁸ *Id.* at 384.

²¹⁹ See generally *Id.*

²²⁰ *Id.*

should certainly be the case where the policy so provided,²²¹ he was not convinced this should necessarily be the case in the absence of such term:

I do not know of any other corner of the law where the plaintiff who has made a fraudulent claim is deprived even of that which he is lawfully entitled to... True, there is distinguished support for such a doctrine... But we were not told of any authority which binds us to teach that conclusion.²²²

It is settled that if the insured makes a fraudulent claim, he or she will not be able to recover.²²³ The consequence is that the insured will forfeit all rights under the policy.²²⁴ However, the question whether the policy can be avoided *ab initio* so that the insurer can recover any payments made with respect to an earlier loss, or whether the insurer should be restricted to recovering only from the date of the fraudulent claim, has received inconsistent responses by the courts.²²⁵ For example, in *Black King Shipping Corp. v. Massie (The Litsion Pride)*, it was held that a fraudulent claim could amount to breach of section 17 of the 1906 Act, thereby entitling the insurer to avoid the contract *ab initio*.²²⁶ However, the courts have recently been retreating from this position by placing limits on the insureds' post-contractual good faith duty. In *Orakpo*, the majority of the Court of Appeal was of the view that where an insured's claim is fraudulent to a "substantial extent," it must fail.²²⁷ The meaning of "substantial" was considered by the Court of Appeal in *Galloway v.*

²²¹ Albeit, subject to the Unfair Terms in Consumer Contracts Regulations.

²²² *Orakpo*, [1994] C.L.C. at 382-83 (citing, *inter alia*, *Britton v. Royal Ins. Co.*, (1866) 176 Eng. Rep. 843; *Black King Shipping Corp. v. Massie (The Litsion Pride)*, [1985] 1 Lloyd's Rep. 437).

²²³ *Britton v. Royal Ins. Co.*, (1866) 176 Eng. Rep. 843, 844.

²²⁴ *Id.*

²²⁵ See, e.g., *Gore Mut. Ins. Co. v. Bifford*, [1987] 45 D.L.R. (Ltd.) 763; *Reid & Co., Ltd. v. Employers' Accident & Livestock Ins. Co. Ltd.*, (1899) 1 F. 1031; *The Litsion Pride*, 1 Lloyd's Rep. 437. See also MALCOLM A. CLARKE, *THE LAW OF INSURANCE CONTRACTS*, *supra* note 214, at 859; Malcolm A. Clarke, *Lies, Damned Lies, and Insurance Claims: the Elements and Effect of Fraud*, [2000] N.Z.L.R. 233, 251.

²²⁶ *The Litsion Pride*, 1 Lloyd's Rep. at 438.

²²⁷ *Orakpo*, [1994] C.L.C. at 385 (Hoffman L.J., Sir Roger Parker). This is merely the application of the *de minimis* rule.

*Guardian Royal Exchange (U.K.) Ltd.*²²⁸ The claimant's premises were burgled and he claimed under a home contents policy some £16,133.94 (the probable true value of the loss) and an additional £2,000 for a computer.²²⁹ In fact, there had been no loss of a computer and a receipt which the claimant produced as evidence of purchase was a forgery.²³⁰ Further, when completing the proposal form for this insurance some five months prior to the claim, he had failed to disclose a conviction for obtaining property by deception.²³¹ Lord Woolf M.R., stressing that the policy of the law must be to deter the making of fraudulent claims, stated that the phrase "substantial:"

is to be understood as indicating that, if there is some immaterial non-disclosure, then of course, even though that material non-disclosure was fraudulent dire consequences do not follow in relation to the claim as a whole; but if the fraud is material, it does have the effect that it taints the whole.²³²

For Lord Woolf M.R., the whole of the claim must be looked at in order to determine whether the fraud is material.²³³ On the facts of the case, the claim for £2,000 amounted to some 10 per cent of the whole.²³⁴ This was an amount that was thought substantial and it therefore tainted the whole claim.²³⁵

²²⁸ [1999] Lloyd's Rep. I.R. 209, 213.

²²⁹ *Id.* at 210.

²³⁰ *Id.*

²³¹ *Id.*

²³² *Id.* at 213.

²³³ *Id.*

²³⁴ *Galloway*, [1999] Lloyd's Rep. I.R. at 213-14.

²³⁵ *Id.* Millett L.J., however, disagreed with this reasoning. He said that the determination of whether or not a claim is "substantially" fraudulent should not be tested by reference to the proportion it bears to the entire claim. *Id.* at 214. To do so "would lead to the absurd conclusion that the greater the genuine loss, the larger the fraudulent claim which may be made at the same time without penalty." *Id.* In Millett L.J.'s view, the size of the genuine claim should not be taken into account. *Id.* All that matters is that the insured is in breach of the duty of good faith which leaves him without cover. As a matter of policy, he added that he would not support any dilution of the insured's duty of good faith. *Id.*

The need for certainty was finally addressed by the House of Lords decision in *Manifest Shipping Co. Ltd. v. Uni-Polaris Ins. Co. Ltd. (The Star Sea)*.²³⁶ While the trial judge had doubted the independent application of utmost good faith to the claims process, the Court of Appeal, took the view that the duty of good faith binds both the insured and the insurer when a claim is made.²³⁷ Leggatt L.J. observed that “[i]t is less clear from the cases whether there is a duty to disclose co-extensive with that which exists before the contract of insurance is entered into, as opposed to a rather different obligation to make full disclosure of the circumstances of the claim. But that distinction matters not.”²³⁸ Leggatt L.J. went on to state that that the insured’s duty of good faith requires that the claim should not be made fraudulently and that the duty “is coincident with the term to be implied by law, as forming part of a contract of insurance, that where fraud is proved in the making of a claim the insurer is discharged from all liability.”²³⁹ In conclusion, the judge stressed that given the draconian remedy available to insurers where a claim is made fraudulently, there should be no enlargement of the insured’s duty so as to encompass claims made “culpably.”²⁴⁰

²³⁶ [2001] UKHL 1, [2003] 1 A.C. 469, 508-09. See also Sir Andrew Longmore, *Good Faith and Breach of Warranty: Are We Moving Forwards or Backwards?*, 2004 L.M.C.L.Q. 158, 166-71.

²³⁷ *Manifest Shipping Co. Ltd. v. Uni-Polaris Ins. Co. Ltd.*, [1995] 1 Lloyd’s Rep. 651, 667; *aff’d*, [1997] 1 Lloyd’s Rep. 360. The judge held that even if it did operate there had to be at the very least recklessness by the insured and that the duty came to an end once legal proceedings had been commenced as after that date false statements were to be dealt with as part of the court’s processes rather than as part of the claim. *Id.*

²³⁸ *Manifest Shipping Co.*, [1997] 1 Lloyd’s Rep. at 371.

²³⁹ *Id.* See also *Orakpo v. Barclays Ins. Servs.*, [1999] C.L.C. 373, 383 (Hoffmann L.J. stated “[a]ny fraud in making the claim goes to the root of the contract and entitles the insurer to be discharged.”) As has been seen in *Galloway*, the Court of Appeal held that the absence of an express condition providing that where there was a fraudulent claim the policy would be void made no difference for the duty of good faith continued long after the policy was effected and applied to the claims process. *Galloway*, [1999] Lloyd’s Rep. I. R. at 211.

²⁴⁰ *Manifest Shipping Co.*, [1997] 1 Lloyd’s Rep. at 372. See also *Diggens v. Sun Alliance and London Ins. plc*, [1994] C.L.C. 1146 (the duty is not broken by an innocent or negligent non-disclosure).

The House of Lords, doubting the reasoning of Hirst J. in *The Litsion Pride*,²⁴¹ accepted that the duty of good faith continued to apply after the conclusion of the insurance contract but held that the claim of fraud had not been proved. As seen above, Lord Hobhouse, noting that the right to avoid under the Marine Insurance Act 1906, section 17 entitles the insurer to rescind the contract *ab initio*, thought that were this remedy to apply where the breach of duty occurs post-contractually, the effect would be effectively penal.²⁴² In his reasoning in this regard, Lord Hobhouse could find no authority to support the notion that the duty of utmost good faith declared by section 17 continued to bind the insured post-contractually:

[The] authorities show that there is a clear distinction to be made between the pre-contract duty of disclosure and any duty of disclosure which may exist after the contract has been made. It is not right to reason, . . . from the existence of an extensive duty pre-contract positively to disclose all material facts to the conclusion that post-contract there is a similarly extensive obligation to disclose all facts which the insurer has an interest in knowing and which might affect his conduct.²⁴³

With respect to the majority view in *Orakpo*, Lord Hobhouse observed that the decision “cannot be treated as fully authoritative in view of the contractual analysis there adopted” with respect to the duty of good faith.²⁴⁴ His Lordship, stressing that the duty of utmost good faith applies only up until the conclusion of the contract, noted that a duty to disclose information can nevertheless arise later, during the currency of the policy, as a result of an express or implied term.²⁴⁵

Recently the issue again arose in *K/S Merc-Scandia XXXXII v. Lloyd's Underwriters (The Mercandian Continent)*.²⁴⁶ The insured

²⁴¹ *Black King Shipping Corp. v. Massie (The Litsion Pride)*, [1985] 1 Lloyd's Rep. 437, 437. See also Howard N. Bennett, *Mapping the Doctrine of Utmost Good Faith in Insurance Contract Law*, 1999 L.M.C.L.Q. 165.

²⁴² *Manifest Shipping Co. Ltd. v. Uni-Polaris Ins. Co. Ltd.* [2001] UKHL 1, [2003] 1 A.C. 469, 494.

²⁴³ *Id.*, [2003] 1 A.C. at 496-97.

²⁴⁴ *Id.* at 501.

²⁴⁵ *Id.* at 495.

²⁴⁶ [2001] EWCA Civ. 1275, [2001] 2 Lloyd's Rep. 563.

submitted a forged letter to his liability insurers to assist them in defending a claim that had been brought against the insured by a third party.²⁴⁷ The purpose was to show that the insured had not entered into a contract with the claimant third party conferring exclusive jurisdiction on the English courts.²⁴⁸ The letter was found to be immaterial and the insurers were therefore held liable.²⁴⁹ The Court of Appeal, aligning the duty of disclosure during the claims process with its pre-contract counterpart, took the view that the non-disclosed or misrepresented fact must be material and it must induce the insurer to pay the claim.²⁵⁰ With respect to the remedy available to the innocent party, Longmore L.J. explained that the right to avoid the contract with retrospective effect is only exercisable in circumstances where the innocent party would, in any event, be entitled to terminate the contract for breach.²⁵¹ He went on to state that:

[T]he giving of information, pursuant to an express or implied obligation to do so in the contract of insurance, is an occasion when good faith should be exercised. Since, . . . the giving of information is essentially an obligation stemming from contract, the remedy for the insured fraudulently misinforming the insurer must be commensurate with the insurer's remedies for breach of contract. The insurer will not, therefore, be able to avoid the contract of insurance with retrospective effect unless he can show that the fraud was relevant to his ultimate liability under the policy and was such as would entitle him to terminate the insurance contract.²⁵²

Not surprisingly, the issue continued to be litigated and the Court of Appeal was soon afforded another opportunity to settle the point, at least with some measure of clarity. In *Agapitos v. Agnew (The Aegeon)*, the question which Mance L.J. focused upon was whether a genuine claim could become fraudulent because it was made fraudulently and whether, in

²⁴⁷ *Id.*, [2001] 2 Lloyd's Rep. at 567.

²⁴⁸ *Id.* at 566-67.

²⁴⁹ *Id.* at 576.

²⁵⁰ *Id.* at 573.

²⁵¹ *Id.* at 575.

²⁵² *The Mercandian Continent*, [2001] 2 Lloyd's Rep. at 576.

consequence, the duty of utmost good faith was broken.²⁵³ Holding that the duty did not apply to fraudulent claims so that the policy could not be avoided *ab initio*, Mance L.J. went on to state the position with respect to fraudulent devices.²⁵⁴ He thought that an acceptable solution would be to "treat the use of a fraudulent device as a sub-species of making a fraudulent claim" and to treat as relevant for this purpose

any lie, directly related to the claim to which the fraudulent device relates, which is intended to improve the insured's prospects of obtaining a settlement or winning the case, and which would, if believed, tend, objectively, prior to any final determination at trial of the parties' rights, to yield a not insignificant improvement in the insured's prospects - whether they be prospects of obtaining a settlement, or a better settlement, or of winning at trial.²⁵⁵

The insurer is therefore discharged from liability in respect of such a claim.²⁵⁶ Concluding the point, it was held that the common law rules governing the making of a fraudulent claim (including the use of fraudulent device) fell outside the scope of section 17 of the 1906 Act.²⁵⁷ Further, the

²⁵³ [2003] Q.B. 556. See also, H. Y. Yeo, *Post-contractual good faith – change in judicial attitude?*, 66 M.L.R. 425 (2003).

²⁵⁴ *Agapitos*, [2003] Q.B. at 574-75.

²⁵⁵ *Id.* at 575. In *Stemson v. AMP Gen. Ins. (NZ) Ltd.* [2006] UKPC 30, the Privy Council endorsed this approach.

²⁵⁶ The reasoning of Mance L.J. has recently been applied by Simon J. in *Eagle Star Ins. Co. Ltd. v. Games Video Co. SA (The Game Boy)*, [2004] EWHC 15. See also *Interpart Comercio e Gestao SA v. Lexington Ins. Co.*, [2004] Lloyd's Rep I.R. 690.

²⁵⁷ See also *Goshawk Dedicated Ltd. v. Tyser & Co.* [2006] EWCA Civ. 54, [2006] 1 All E.R. (Comm.) 501, which held that any notion that the insured's duty of good faith continues post-contractually cannot be divorced from the terms of the policy. The way in which such a continuing duty can arise is by implying a term into the contract, on the basis that it is necessary for business efficacy, which requires the insured to provide information in appropriate circumstances. It therefore follows that there is no independent post-contractual good faith duty. All post-contract issues are to be determined according to the terms of the policy and in this respect, the decision appears to accord with Staughton L.J.'s minority view in *Orakpo v. Barclays Ins. Serv. 's*, [1994] C.L.C. 373 (A.C.). See also *Bonner v. Cox Dedicated Corporate Member Ltd.* [2004] EWHC 2963 (Comm.) (Morison J.). Further, following *Friends Provident Life & Pensions Ltd. v. Sirius Int'l. Ins.*

Court of Appeal also went on to hold that once litigation between the insurers and the insured has commenced, the consequences of making a fraudulent claim or promoting a claim with fraudulent devices are superseded by the procedural rules governing civil litigation.²⁵⁸

Mance L.J. was again given the opportunity to revisit the issue in *AXA General Insurance Ltd. v. Gottlieb*.²⁵⁹ The issue was whether under the common law rule relating to fraudulent claims, an insurer could recover interim payments made prior to any fraud in respect of genuine losses incurred on the claim to which the subsequent fraud related.²⁶⁰ The judge rejected the submission of the insureds' counsel to the effect that where a genuine right to indemnity has both arisen and been subject of a payment made prior to any fraud committed in respect of the same claim, there can be no conceptual basis for requiring the insured to repay the sums received.²⁶¹ Mance L.J. stated that:

If a later fraud forfeits a genuine claim which has already accrued but not been paid, the obvious conceptual basis is that the *whole* claim is forfeit... If the whole claim is forfeit, then the fact that sums have been advanced towards it is of itself no answer to their recovery.²⁶²

The effect of counsel's argument would be to result in the anomaly that forfeiture of the whole claim should be restricted to the whole of the outstanding claim only; in other words, to any part that remains unpaid as of the date of the fraud. Mance L.J. explained that the rationale of the rule relating to fraudulent claims is that an insured should not expect that,

Co. [2005] EWCA Civ. 601, it will probably never be the case that breach of such a term is repudiatory of the policy as a whole. Waller L.J. expressed the unconvincing view that perhaps a series of breaches might be repudiatory so that, in effect, the notion of a continuing duty is dead and buried. *Id.*

²⁵⁸ See also *Manifest Shipping Co. Ltd. v. Uni-Polaris Shipping Co. Ltd.*, [2001] UKHL 1, 481. *But see* *Eagle Star Ins. Co. Ltd.* [2004] EWHC 15. Simon J. explained that this could give rise to anomalous consequences: "After litigation has commenced an insured may advance false documentation and lie without the drastic consequences which follow if the deployment of false documentation and lies are less well timed." *Id.*

²⁵⁹ [2005] EWCA Civ. 112 (Keene and Pill L.J.J., concurring).

²⁶⁰ *Id.*

²⁶¹ *Id.*

²⁶² *Id.*

should the fraud fail, he or she will lose nothing.²⁶³ The court should not, therefore, undermine the prophylactic policy of the common law rule by holding that forfeiture should not apply to a part of a claim that is otherwise honest.²⁶⁴ Accordingly, it was held that the effect of the common law is to forfeit the whole of the fraudulent claim so that the consideration for any interim payments made on that claim fails.²⁶⁵ Such sums are thus recoverable by the insurers irrespective of whether they were paid prior to the fraud.²⁶⁶

The issue of fraudulent claims again came to the fore in *Danepoint Ltd. v. Allied Underwriting Insurance Ltd.*, in which a block of some thirteen flats was damaged by a fire.²⁶⁷ The insured lodged a number of exaggerated claims together with a fraudulent claim relating to loss of rent.²⁶⁸ Coulson J. subjected the authorities to thorough review. He concluded that the duty of utmost good faith declared by section 17 of the 1906 Act does not trigger during the claims process.²⁶⁹ An insurer cannot, therefore, avoid the policy *ab initio* on the ground of fraud.²⁷⁰ Where all or part of the claim is fraudulent, or where fraudulent devices are enlisted to promote a genuine claim, the insured will not be permitted to recover in respect of any part of the claim.²⁷¹ Mere exaggeration will not, in itself, suffice to substantiate an allegation of fraud.²⁷² But if the exaggeration is wilful, or is allied to misrepresentation or concealment, it will, in the judge's view, probably be held to be fraudulent. In this regard, an exaggeration is more excusable where the value of the particular claim or head of loss in question is unclear or is a matter of opinion.²⁷³

²⁶³ *Id.*

²⁶⁴ *Id.*

²⁶⁵ *AXA Gen. Ins. Ltd.* [2005] EWCA Civ. 112.

²⁶⁶ *Id.*

²⁶⁷ [2005] EWHC 2318 (TCC). *See also* J. Lowry and P. Rawlings, *Fraudulent claims: framing the appropriate remedy*, [2006] J. BUS. L. 339. *Cf.* *Stemson v. AMP Gen. Ins. (NZ) Ltd.*, [2006] UKPC 30.

²⁶⁸ *Danepoint Ltd.* [2005] EWHC 2318.

²⁶⁹ *Id.*

²⁷⁰ *Id.*

²⁷¹ *Id.* at 432.

²⁷² *Id.* at 438.

²⁷³ *See, e.g., Orakpo v. Barclays Ins. Serv.*, [1994] C.L.C. 373 (A.C.).

V. THE FUTURE

The reasoning expressed in the modern cases demonstrates a significant shift in the way the courts approach the good faith duty. The process of recalibrating the insured's pre-contractual duty of good faith seen in the case law over the last decade or so is not being done in isolation from other aspects of the insurance contract for the courts are also adding content to the duty of good faith which the insurer owes to the insured at the time the risk is presented, at the time when the remedy of avoidance is exercised and at the time when insurers assert the benefit of claims clauses. With respect to consumer insurance there is a considerable body of Ombudsman jurisprudence to be added to the burgeoning case law. The result is that we now have two parallel regimes governing insurance contracts: one relating to commercial insurance and one relating to consumer insurance. In terms of coherence, this is not satisfactory. However, the Scottish and English Law Commissions current re-examination of insurance offers the potential for a thorough overhaul of the law, and in this regard it will be recalled that non-disclosure is identified as a key issue. It was to be hoped that the exercise would seek to assimilate the developments seen in the modern decisions into a single scheme for both consumer and commercial insurance. Admittedly, for other types of contracts the legislature has seen fit to distinguish between consumer and commercial transactions,²⁷⁴ but, as is pointed out by Professor Clarke, for insurance the distinction necessarily results in the adoption of arbitrary tests based on turnover.²⁷⁵ In any case, English judges have shown some reluctance in recognising such a dichotomy. For example, in *Cook v. Financial Insurance Co.*,²⁷⁶ Lord Lloyd,²⁷⁷ considering the effect of a term contained in a policy for disability insurance effected by a self-employed builder, took the view that it "must be construed in the sense in which it would have been reasonably understood by him as the consumer...."²⁷⁸

In framing a suitable model for the duty of good faith, an appropriate balance needs to be struck between the economic costs of reform and the benefits, including social benefits, which a more balanced regime will bring. Such considerations must also be weighed against the

²⁷⁴ For example, in the realms of sales law and credit transactions.

²⁷⁵ Clarke, *supra* note 199, at 288. See *supra* text accompanying note 9.

²⁷⁶ [1998] 1 W.L.R. 1765.

²⁷⁷ Both Lord Steyn and Lord Hope agreed.

²⁷⁸ *Cook*, [1998] 1 W.L.R. at 1768.

objective of improving the competitiveness of the insurance market.²⁷⁹ This is not to under-estimate the difficulties of framing a solution. Any such reform needs to avoid excessive interference with commercial practices and avoid introducing uncertainty into the law.²⁸⁰ Over twenty years ago, the Australian Law Reform Commission (hereafter, the ALRC) published its wide-ranging report on insurance law,²⁸¹ which unlike the 1980 report of its English counterpart, led to statutory reform by way of the Australian Insurance Contracts Act 1984 (hereafter, ICA 1984).²⁸² The ALRC recommended a new test for the determination of the insured's duty of disclosure, namely that the duty should be tested by what the insured knew, or what a reasonable person in the insured's circumstances would have known, to be relevant to the assessment of the risk.²⁸³ The ALRC considered that this formulation of the duty was more consistent with the limits of the insured's duty to exercise utmost good faith.²⁸⁴ It also thought that the formulation would achieve a fairer balance between insured and insurer than would the more objective test recommended by the English Law Commission in 1980.²⁸⁵

In reaching its conclusions, the ALRC took into account a number of factors which offer important lessons for the investigation now underway in the UK. It found that fairness to the insured can only be achieved if both insurers and the law which regulates the insurance relationship are sensitive to the literacy, knowledge, experience, and

²⁷⁹ This objective informed the deliberations and recommendations put forward by the Australian Law Reform Commission. AUSTRALIAN LAW REFORM COMMISSION, INSURANCE CONTRACTS, Rep. No. 20, at xxi (1982).

²⁸⁰ See report of the NEW ZEALAND LAW COMMISSION, *supra* note 6, ¶ 10.

²⁸¹ See AUSTRALIAN LAW REFORM COMMISSION, *supra* note 279. The Report is regarded as authoritative in the interpretation of the Insurance Contracts Act, 1984 (Austl.). See *Ferrcom Pty. Ltd. v. Commercial Union Assurance Co. of Austl. Ltd.*, (1993) 176 C.L.R. 332, 340. By virtue of §§ 13-14 of the 1984 Act, utmost good faith is an implied term that applies to both parties to the contract. Thus, breach of the duty is a breach of contract giving rise to damages or to an estoppel and not to avoidance *ab initio*.

²⁸² Most of the 1984 Act's provisions came into operation on 1 January 1986.

²⁸³ See AUSTRALIAN LAW REFORM COMMISSION, *supra* note 279, ¶ 24, at xxix. See also ICA 1984, *supra* note 281, §§ 21 and 21A (as amended).

²⁸⁴ AUSTRALIAN LAW REFORM COMMISSION, *supra* note 279, ¶ 328, at 202.

²⁸⁵ *Id.* ¶ 182, at 110.

cultural background of insureds.²⁸⁶ More particularly, it was emphasised that the law should recognise the modern conditions in which insurance is marketed.²⁸⁷ Nowadays, insurance contracts are concluded with a minimum of formality and so, subject only to the principle of good faith, insurers should take individual members of the relevant market as they find them. The ALRC found that the existing duty of disclosure imposes obligations which many prospective insureds, acting in the utmost good faith, are unable to discharge.²⁸⁸ Indeed, in the current market place marketing methods are adopted which increase the risk of non-disclosure, and where intermediaries are not involved, there is no one to bring to the insured's attention the breadth of the disclosure obligation. For reasons of cost and competition, proposal forms are often kept to a minimum, especially so where direct marketing of insurance products is used whereby policies are purchased by means of computer-based communications systems. Taken in the round, these developments increase the risk of innocent non-disclosure. A modern regime should therefore take account not only of the various subjective factors affecting insureds, but also of the diverse methods enlisted by insurers to transact with their prospective customers. A test of disclosure based on the twin attributes of the actual insured together with the reasonable insured strikes the optimum balance in maintaining a single test, albeit dual-limbed, for both consumer and

²⁸⁶ Perhaps surprisingly, these factors were specifically excluded by the Law Commission's final report in 1980. That said, it should be noted that the ICA 1984 differs from the ALRC's formulation of the duty. Section 21(1)(b) refers to matters that "a *reasonable person* in the circumstances" could be expected to know (emphasis added). Notwithstanding the pure objectivity of this statutory formulation, it has received positive endorsement beyond the shores of Australia. See, e.g., INSURANCE LAW REFORM: THE CONSUMER CASE FOR A REVIEW OF INSURANCE LAW, *supra* note 15; BRITISH INSURANCE LAW ASSOCIATION, REFORM OF INSURANCE CONTRACT LAW (2006); and more recently it has received broad support from the English and Scottish Law Commissions, *supra* note 17. It has also been welcomed by the New Zealand courts. For example, in *State Insurance v. McHale*, [1992] 2 N.Z.L.R. 399, 415, Richardson and Hardie Boys J.J. concluded that: "[t]he law in New Zealand as to materiality and the duty of disclosure is not satisfactory. It can lead to uncertainty and injustice.... The test of the reasonable assured has much to commend it. The Australian legislation adopting that test ... could well be followed in this country." See also the remarks of Cooke P, [1992] 2 N.Z.L.R. at 404. See also *Quinby Enter. Ltd. v. Gen. Accident Ltd.*, [1995] 1 N.Z.L.R. 736, 740.

²⁸⁷ AUSTRALIAN LAW REFORM COMMISSION, *supra* note 279, ¶ 183, at 111.

²⁸⁸ *Id.*

commercial insurance.²⁸⁹ Although insurers often have exclusive recourse to data relevant to particular types of risks, they do not possess superior knowledge in relation to factors peculiar to the particular risk sought to be insured for this generally lies within the province of the insured. Therefore, an insured under this test would be required to prove the existence of any circumstances which he or she relies on to reduce the scope of the duty of disclosure.

As seen above, a particular feature of the modern English case law is the emphasis now being placed on the way in which insurers seek to exercise the remedy of avoidance. It will be recalled that while insurers are entitled to avoid the contract *ab initio*, the judges have expressed considerable unease over the draconian consequences suffered by insureds. In *Drake Insurance*, Rix L.J. addressed the issue in wider terms than most in calling for regard to be had to the concept of proportionality.²⁹⁰ As noted above, the Law Commission's 1980 report expressly rejected the proportionality doctrine on the basis that it was unworkable.²⁹¹ It also rejected a "nexus test" whereby the insurers would be required to demonstrate that the undisclosed fact is in some way connected to the loss.²⁹² In reaching this conclusion the Law Commission reasoned that:

all considerations relating to non-disclosure must focus on the moment when a proposal for insurance is put forward and either accepted on certain terms or rejected, in either event by reference to what the insurer judges to be the quality of the risk. The technique - one might almost say the art - of good underwriting is to judge all the factors affecting an offered risk at this moment, when the underwriter must then and there assess its quality on the basis of his experience, as though he were considering the

²⁸⁹ Such a subjective/objective form is now the accepted test for determining the appropriate standard of care for directors. *See, e.g.*, *Norman & Anor. v. Theodore Goddard & Ors.*, [1992] B.C.C. 14. This appears in the statutory statement of directors' duties contained in the Companies Act, 2006, c.46, § 174 (Eng.).

²⁹⁰ [2003] EWCA Civ. 1834, [2004] Q.B. at 628, 629.

²⁹¹ For the Law Commission's reasoning in this regard, *see supra* text accompanying note 15.

²⁹² *See* THE LAW COMMISSION, *supra* note 15, ¶¶ 4.91-4.97.

overall impression given by a “still photograph” of the risk at this point.²⁹³

While the ALRC had some sympathy for the misgivings expressed by the Law Commission over the difficulties of proof in relation to proportionality and causation, it did not think these were insurmountable. The ALRC saw no reason why in most cases insurers would not be able to establish, “whether from rating guides, from its instructions to its agents or staff or from its prior conduct, the nature and extent of the loss which it had suffered.”²⁹⁴ While conceding that it would sometimes be difficult to establish how it would have reacted to additional moral, as distinct from statistical, risks the ALRC concluded, in a robust statement of principle, that:

difficulties of proof cannot be avoided if a proper balance is to be reached between the interests of the insurer and those of the insured. It is quite plainly contrary to the true principle of *uberrima fides* to impose on the insured a burden which far exceeds the harm which he has done. The insurer should not be entitled to any redress which exceeds the loss which it has in fact suffered. That is the basic principle which lies behind the law of damages, both in contract and in tort.²⁹⁵

It therefore recommended that the nature and extent of the insurer’s redress should depend on the nature and extent of the loss which it has suffered as a result of the insured’s conduct and that it should no longer be entitled to avoid a contract, and a heavy claim under that contract, merely because it has suffered a small loss as a result of non-disclosure.²⁹⁶ This certainly

²⁹³ *Id.* ¶ 4.94.

²⁹⁴ AUSTRALIAN LAW REFORM COMMISSION, *supra* note 279, ¶ 194.

²⁹⁵ *Id.*

²⁹⁶ *Id.* See Insurance Contracts Act, 1984, §§ 28–31 (Austl.) (as amended). See *Advance (NSW) Ins. Agencies Pty. Ltd. v. Matthews*, (1987) 4 ANZ Ins. Cas. 60–813 (Young J.) (New South Wales Sup. Ct.); *cf. Lindsay v. CIC Ins. Ltd.*, (1989) 16 N.S.W.L.R. 673; *Ferrcom Pty. Ltd. v. Commercial Union Assurance Co. of Austl. Ltd.*, (1989) 5 ANZ Ins. Cas. 60-907 (Giles J.) (New South Wales Sup. Ct.). See also *Alexander Stenhouse Ltd. v. Austcan Inv. Pty. Ltd.*, (1993) 112 A.L.R. 353 (Mason C.J., Deane, Dawson, Toohey & McHugh J.J.). Section 31 of the ICA provides that a court may disregard an avoidance and order the insurers to

aligns the position in insurance with fundamental principles underpinning the law of damages both in contract and tort. More particularly, it also reflects the approach taken in the general law of contract towards restricting the right to rescind a contract for innocent misrepresentation.²⁹⁷ Accordingly, the ALRC took the view that the right of insurers to avoid a contract from its inception should be abolished except for cases of fraudulent non-disclosure on the basis that avoidance *ab initio* was wholly disproportionate to the harm caused by an insured's non-fraudulent non-disclosure.²⁹⁸ Rather, the insurer should be able to cancel the contract prospectively and be entitled to adjust a claim to take into account the loss actually suffered by it as a result of the insured's breach of the disclosure duty.²⁹⁹ As to the question of assessing damages, the ALRC favoured the approach taken by the common law in claims for misrepresentation whereby damages for a breach of duty would depend upon what the insurer would have done had it known of the true facts.³⁰⁰ Any other remedies available to the insurer would depend on the response it would have made if it had known of the undisclosed material facts. For example, if it would have declined the risk outright, the insurer's loss is equivalent to the amount claimed against it. If it would have accepted the risk albeit at a higher premium, its loss is the difference between the actual and notional premiums. If it would have accepted the risk but on different terms, whether at the same premium or not, its loss is the difference between its liabilities under the actual and notional contracts.

Not surprisingly, modern English decisions such as *Drake Insurance* and Rix L.J.'s dissent in *WISE Ltd.* illustrate the anxiety of the modern judges to address unfair dealings, certainly with respect to the

pay some or all of a claim consistently with what is just and equitable in the circumstances. It should be noted that the case law on sections 28-31 of the 1984 Act is confusing. The courts have developed their own principles on the assessment of damages.

²⁹⁷ See Misrepresentation Act, 1967, c. 7, § 2(2) (Eng.). It should be noted, however, that the Australian courts have frequently questioned whether the analogy between damages for breach of contract and damages for misrepresentation/non-disclosure is strictly correct.

²⁹⁸ AUSTRALIAN LAW REFORM COMMISSION, *supra* note 279, § 194.

²⁹⁹ *Id.* The Australian Law Reform Commission recommended reforms along similar lines for the law relating to misrepresentation.

³⁰⁰ *Id.*

exercise of avoidance by insurers.³⁰¹ From the pre-contractual standpoint, we have also seen the courts aligning insurance law with general contract law in terms of the assimilation of non-disclosure with misrepresentation whereby the insurers must have been induced by the non-disclosed fact.³⁰² This process, an objective of which appears to be a rebalancing of the respective rights and obligations of the parties, is not restricted to the modern case law. It should be viewed in tandem with the regulatory approach now being adopted by the UK's Financial Services Authority through its Insurance: Conduct of Business (ICOB). The ICOB, while less prescriptive in relation to commercial policies than it is for consumer contracts, nevertheless requires insurers to treat commercial customers "fairly" and not to unreasonably reject claims.³⁰³

From the European perspective it is of interest to view these developments against the background of the Directive on Unfair Commercial Practices which was approved on 11 May 2005.³⁰⁴ The Directive is aimed at providing a uniform and comprehensive standard for prohibiting unfair commercial practices. Although it is aimed at the position of consumers, including those affected by unfair commercial practices of insurers, there seems no reason in principle why the objective of the Directive (taken with the emphasis placed on 'fairness' in the ICOB) together with the current judicial thinking on the content of the duty of good faith borne by both parties, should not underpin the Law Commissions' current review of insurance law. Indeed, in this respect, a degree of optimism is warranted. In September 2006, the English and Scottish Law Commissions published an Issues Paper on Misrepresentation and Non-Disclosure which was intended to promote discussion and

³⁰¹ *Drake Ins. plc v. Provident Ins. plc* [2003] EWCA Civ. 1834, [2004] Q.B. 601; *WISE Ltd. v. Grupo Nacional Provincial SA* [2004] EWCA Civ. 962.

³⁰² Whether this is particularly novel is another question. Certainly, in *Carter v. Boehm*, (1766) 97 Eng. Rep. 1162, 1165, Lord Mansfield did not draw any sharp distinction between them. As Professor Clarke has observed, "If I describe the shandy that I have just bought you as lemonade, is that non-disclosure of part, the beer, or misrepresentation of the whole?" Clarke, *supra* note 199, at 288.

³⁰³ Insurance: Conduct of Business Sourcebook, Financial Services Authority Handbook, Release No. 094, § 8.1.2 (Oct. 2009).

³⁰⁴ Unfair Commercial Practices Directive, Council Directive 2005/29, 2005 O.J. (L 149) 22 (EC). See Hugh Collins, *The Unfair Commercial Practices Directive*, 1(4) EUR. REV. CONT. L. 417 (2005); and Hugh Collins, *Harmonisation by Example: European Laws against Unfair Commercial Practices* 17 M.L.R. 89 (2010).

feedback.³⁰⁵ The current system is criticised as “incoherent and flawed” on the basis that insurers can avoid policies inappropriately; that consumers are “deprived of a genuine choice between the FOS and the courts;” and that it “requires the FOS to exercise undue discretion.”³⁰⁶

The initial recommendation was that the duty of disclosure in consumer insurance should be abolished. This proposal survived the various consultation exercises carried out by the Law Commissions and now forms the central plank of their proposals and draft Bill which was published in December 2009.³⁰⁷ Clause 2 of the Bill replaces the consumer-insured’s duty of disclosure with the duty “to take reasonable care not to make a misrepresentation.” This therefore removes the consumer’s duty to volunteer information to the insurer. Instead, consumers will be required to answer insurers’ questions honestly and to take reasonable care that their replies are accurate and complete. If consumers do, however, provide insurers with information which was not asked for, they must also do so honestly and carefully. The thinking here is that abolition of the disclosure duty would force insurers to ask the right questions in proposal forms. The draft Bill does not require the insurer to ask specific questions. However, clause 3(2) provides that in assessing the reasonableness of the consumer’s answer to a question, the court (or ombudsman) will take account “how clear, and how specific, the insurer’s questions were.” Clause 10, amongst other things, goes on to prevent insurers from contracting out of the provisions of the Bill. Thus, a policy term, or a term in any other contract, is rendered void to the extent that it would put the consumer in a worse position than under the draft Bill.

The prudent underwriter test is thus replaced with a reasonable insured test. Schedule 1 of the draft Bill goes on to lay down the insurers’ remedies for misrepresentation. The applicable remedy should depend on the insured’s state of mind. Where a consumer acts honestly and reasonably the insurer will be required to pay the claim. In cases of fraud (termed a “deliberate or reckless” misrepresentation), the insurer will be entitled to refuse to pay the claim. In such a case the insurer will need to prove on the balance of probabilities that the consumer knew (a) that the statement was untrue or misleading, or did not care whether it was or not, and (b) that the matter was relevant to the insurer, or did not care whether it

³⁰⁵ See The Law Commission, *supra* note 17.

³⁰⁶ *Id.* § 5.24.

³⁰⁷ CONSUMER INSURANCE LAW: PRE-CONTRACT DISCLOSURE AND REPRESENTATION, *supra* note 17.

was or not. However, with respect to negligent misrepresentation, a scheme of apportionment will apply in that both parties should be put into the position they would have been in had the insurer known the true facts. For example, if the insurers would have charged a higher premium, the insured will be able to recover that proportion of his loss which corresponds to the proportion of the proper premium actually paid.³⁰⁸ For innocent misrepresentation, i.e. where the insured had reasonable grounds for believing the truth of what is stated, the insurer will have no remedy.

The Law Commissions thus draw a distinction between consumer and non-consumer insurance contracts. The received wisdom is that businesses require less protection in part because businesses use expert brokers, but also because the market for commercial insurance is competitive and businesses can generally negotiate with insurers in a way not available to consumers. Overall, it is provisionally recommended that the duty of disclosure should continue to apply to non-consumer insurance but subject to a “reasonable insured” test for materiality. This would also apply to misrepresentation.³⁰⁹ It is also proposed that the same remedies as recommended for consumer policies should be available for fraudulent, negligent and innocent misrepresentation, although a range of questions concerning negligent misrepresentation are put forward for discussion. A policy statement on pre-contract disclosure and misrepresentation in business insurance is expected to be issued in 2010.

It would be churlish not to welcome the Law Commissions’ proposals for reforming insurance contracts. More particularly, a positive feature is that a number of their proposals resonate with those originally put forward by the ALRC in 1982. However, perpetrating the distinction between consumer and business insureds is open to question.³¹⁰ Maintaining the division between the two does little to further the objective

³⁰⁸ For example, if an insurer only charged a premium of £1,000 but should have charged £1,500, the consumer will receive two thirds of the claim. See the Explanatory Notes to the draft Bill, A.17(3), CONSUMER INSURANCE LAW: PRE-CONTRACT DISCLOSURE AND REPRESENTATION, *supra* note 17.

³⁰⁹ Such rules will be compulsory for business policies. See Law Commission, *Insurance Contract Law, Warranties* (Nov. 2006), available at http://www.lawcom.gov.uk/docs/Insurance_Contract_Law_Issues_Paper_2.pdf. See also, the Law Commissions’ Consultation Paper No 182; Discussion Paper No 134, (July, 2007), *Insurance Contract Law: Misrepresentation, Non-Disclosure and Breach of Warranty by the Insured*.

³¹⁰ Regrettably, the Australian ICA 1984, § 21A also draws the distinction between private and business insureds.

of constructing a coherent regime for insurance. On a more positive note, the recommendations relating to remedies are to be welcomed especially in so far as the Law Commissions now embrace a form of apportionment. Other issues still remain to be addressed, not least the question of whether or not the good faith duty continues to operate post-contractually. As seen above, the case law has at last settled the point and no doubt the Law Commissions will assimilate this position in their final proposals. Their work on this is expected to begin in 2010.³¹¹ Finally, in relation to claims there remains a significant question which the authorities have thus far failed to address. Will a fraudulent claim bring the insurance contract to an end so that insurers can refuse to pay a legitimate claim that is made subsequent to that which is held to be fraudulent? Although the issue has so far evaded judicial determination, the English and Scottish Law Commissions now have the opportunity to grasp the nettle.

³¹¹ See http://www.lawcom.gov.uk/insurance_contract.htm (accessed 11 January 2010).

RISK DATA IN INSURANCE INTERPRETATION

Michelle Boardman*

Insurance companies use facts about past risks—actuarial data—in essential ways. The data are at the crux of creating the product and clearing regulatory hurdles to selling the product. Given the centrality of the data in the drafting, pricing, and legitimizing of insurance policies, it is peculiar that courts, insurers, and policyholders tend to ignore it when the time comes to interpret and apply a policy in court. This article imagines the shape its use would take and considers casual empirics on why it is not used more now.

There are three ways for actuarial data to advance interpretation and construction. The first is in proving or disproving insurer good faith. Actuarial data can show an insurer's bona fides—countering the universal underlying assumption of the swindling insurer. Comparing money taken in (the premium calculation) with money paid out (the risks covered) can confirm or deny a bait and switch scheme. Second, the data can prove an otherwise abstract claim of actuarial purpose, providing the context that resolves a nascent ambiguity. This is important because a finding of ambiguity is four-fifths of a finding that the policyholder wins. Third, actuarial data can reveal insurer intent—not simply a lack of bad faith but a particular intent. Recognizing this intent brings some surprising benefits to both consumers and insurers.

In insurance, courts are often engaged in a project that is both more than and less than interpretation. Courts are engaged in regulation of the insurance policy directly, dictating the clauses insurers can and cannot enforce. Actuarial function provides the court intent on regulating the insurance field with the policy implications of a particular ruling. The data

* Assistant Professor of Law, George Mason University School of Law. Brown University, B.A. 1994, University of Chicago Law School, J.D. 1998. Thanks to Tom Baker, Lloyd Cohen, Kimberly Moore, Tony Sebok, James Siewert, Nicholas Quinn Rosenkranz, Alice Wellford, Hill Wellford, and participants in the Vanderbilt University Law School faculty workshop. Thank you to Adina Horvath and Cody Williams for excellent research support.

will allow a court engaging in regulation to consider policyholders other than the one momentarily before it.

Insurance companies use facts about past risks—actuarial data—in essential ways. The data are at the crux of creating the product (what risks to cover, at what cost) and clearing regulatory hurdles to selling the product (proving that they are charging enough but not too much).¹ Given the centrality of the data in the drafting, pricing, and legitimizing of insurance policies, it is peculiar that courts, insurers, and policyholders tend to ignore it when the time comes to interpret and apply a policy in court. This article imagines the shape its use would take and considers why it is not used more than sparingly now.

Behind each insurance contract there lies a city of statistics. Here you will find answers to when a risk becomes a loss, how many people will lose each year, and how much will be lost. Following the hundreds of intertwining streets will lead you to the insurer's ultimate question: how much one would have to be paid to take on the risk of all the losses together.

These city streets are paved with data: raw data of yesterday's losses and calculations of tomorrow's. (Those craving a more elaborate description of actuarial data can dash ahead to section I.) From these data, insurers decide which risks to insure and which to omit. An insurance policy can be thought of as the insurer's attempt to explain this, to explain in words what the equation includes and excludes.² But just as describing a dance is not the dance itself, the contract language is not the underlying truth of what risks insurers have included and excluded from their premium and expected risk calculations.

The preliminary question is why courts engaged in interpretation would care about this hidden city. Its secrets may seem no more relevant to

¹ See 1 LEE R. RUSS & THOMAS F. SEGALLA, *COUCH ON INS.* §§ 2:27, 2:29, 2:31 (3d ed. 2009). See generally, Susan Randall, *Freedom of Contract in Insurance*, 14 CONN. INS. L. J. 107, 124-34 (2007).

² To whom the insurer offers its explanation is not an easy question. From a contractual standpoint, the policyholder should be the audience. For my explanation of why insurers might be more interested in communicating with the courts, see Michelle Boardman, *Contra Proferentem: The Allure of Ambiguous Boilerplate*, 104 MICH. L. REV. 1105 (2006).

the act of interpretation than the fact that the policy was issued on a Tuesday. In insurance, however, courts are often engaged in a project that is both more than and less than interpretation. Courts are engaged in regulation of the insurance policy directly, dictating the clauses insurers can and cannot enforce. Actuarial data can directly address the construction and judicial regulation of insurance contracts.

This article offers three main ways in which actuarial data can advance interpretation and construction. The first is in proving or disproving insurer good faith, an area of heightened relevance in insurance because insurers have a near-fiduciary duty toward policyholders. Actuarial data can show an insurer's bona fides—countering the universal underlying assumption of the swindling insurer. Is an insurer arguing for a particular reading of a provision, not because the insurer “means it” in some sense, but because the insurer seeks to avoid paying for any loss, at any time, under any theory? At times, actuarial data can answer this question. Comparing money taken in (the premium calculation) with money paid out (the risks covered) can confirm or deny a bait and switch. Ultimately, the threat of this comparison will shrink the number of genuine swindlers by bringing the con to light.

Second, the data use contains within it its own purpose. It can *prove* an otherwise abstract claim of actuarial purpose, such as avoiding synchronized losses across many people or avoiding moral hazard.³ Actuarial purpose supports a “reasonable reading” of a clause, perhaps providing the context that resolves a nascent ambiguity. This is important because a finding of ambiguity is four-fifths of a finding that the policyholder wins.⁴ Actuarial purpose also provides the court intent on regulating the insurance field with the policy implications of a particular ruling. Courts are accustomed to looking beyond (or over and around) the language of insurance policies to determine not just their written meaning, but the meaning the court is willing to enforce. Kenneth Abraham's central

³ Black's Law Dictionary defines “moral hazard” as a “hazard that has its inception in mental attitudes,” such as the “risk that an insured will destroy property or allow it to be destroyed (usually by burning) in order to collect the insurance proceeds is a moral hazard.” BLACK'S LAW DICTIONARY 786 (9th ed. 2009). In modern insurance discussions, however, it is used to refer to the danger that a person may take less care in avoiding a hazard, knowing that insurance will cover part of the loss. *See generally*, Tom Baker, *On the Genealogy of Moral Hazard*, 75 TEX. L. REV. 237 (1996).

⁴ *See* TOM BAKER, INSURANCE LAW AND POLICY: CASES AND MATERIALS 466 (2003).

“A Theory of Insurance Policy Interpretation,” explores how several of the fundamental principles of insurance law “direct that under specific circumstances the meaning of even clear policy language must be disregarded, not interpreted.”⁵ Actuarial data will allow a court engaging in regulation to consider policyholders other than the one before momentarily before it.⁶

Third, actuarial data can reveal insurer intent—not simply a lack of bad faith but a particular intent. Courts do not ordinarily treasure the secret meaning a seller harbors in his heart when making an offer to a buyer. But left with complex contract language and an amorphous buyer’s intent that does not mirror the seller’s, judges are hard pressed to turn away guidance.

Intent has a few specific uses. First, hewing more closely to insurer intent should decrease the cost of insurance. Uncertainty about how courts will rule increases “contract risk,” which increases premiums.⁷ Surveys “illustrate that uncertainty about losses and ambiguity about probability lead to higher prices.”⁸ Second, knowing the intent behind a

⁵ Kenneth S. Abraham, *A Theory of Insurance Policy Interpretation*, 95 MICH. L. REV. 531, 532 (1996) (Some of the doctrines “have nothing to do with ‘interpretation’ as it is normally understood.”).

⁶ For a modern defense of (limited) judicial regulation, see Daniel Schwarcz, *A Products Liability Theory for the Judicial Regulation of Insurance Policies*, 48 WM. & MARY L. REV. 1389 (2007). Schwarcz does not advocate the introduction of specific actuarial data but his framework requires courts to consider whether the policy causes “insurance harm” and to “ask whether the insurer has any legitimate *underwriting purpose* for not insuring against the *specific loss* that befell the insured.” *Id.* at 1448 (emphasis added).

⁷ See Tom Baker, *Insuring Liability Risks*, 29 GENEVA PAPERS ON RISK & INS. 128, 139-40 (2004).

⁸ Howard Kunreuther, Robin Hogarth & Jacqueline Meszaros, *Insurer Ambiguity and Market Failure*, 7 J. OF RISK AND UNCERTAINTY 71, 79 (1993). The 1993 Kunreuther et al. studies of insurer ambiguity presented the insurance actors with a set probability on which “all experts agree” and contrasted that with an ambiguous probability, defined as a “wide disagreement about the estimate of p [the probability of loss] and a high degree of uncertainty among the experts.” *Id.* at 72. See also Howard Kunreuther et al., *Ambiguity and Underwriter Decision Processes*, 26 J. ECON. BEHAV. & ORG. 337, 342-44 (1995). Instead of a wide range in the probability that a loss will occur, here the issue is a wide range that a policy will be read to cover the loss. The surveys also found some evidence that the risk premium could be double-charged: “To the extent that primary underwriters do not recognize that the prices of actuaries may *already* include

clause makes it possible to assess whether the clause could have been written more clearly. Where the intent, while legitimate, is too complex to be conveyed well to consumers, courts must decide whether to effectuate the intent or forbid any similar clause.

Courts adopt a regulatory view toward the health of the insurance industry and the interests of those policyholders not before the court. If the purpose of the clause is insurer solvency, for example, a regulatory court may prefer to protect solvency over literal language interpretation or other values. Courts most commonly regulate by mandating coverage and forbidding exclusions to coverage. But regulating the substance of insurance clauses without access to the actuarial function of those clauses is looking left and leaping right. Insurer purpose and intent do not need to control the outcome of a court's decision to improve the outcome of its regulation.

Three parties are in this game—policyholders, insurers, and courts. Each has incentives, the pursuit of which creates externalities. Each has motives that can be described generously or with distrust; there is benefit to doing both. There is a fourth major player, of course, in the public, but the public's interest is ubiquitous. The public's needs are partially represented by courts (although their power to pursue policy aims is cabined), and partially represented by policyholders, in that most individuals are policyholders. But a policyholder after a loss may pursue his individual compensation over a healthy insurance market and insurers can represent the interest of the many policyholders against the few (or the future policyholders against the demands of the present).⁹

adjustments for ambiguity and uncertainty, they may recommend a premium that reflects their concerns with these factors." Kunreuther, Hogarth & Meszaros at 75 (emphasis added).

⁹ "First, insurance looks at groups, at the socialization of risk through standard contracts sold to large numbers of similarly situated persons who face an uncertain risk. What is good for the group, as a whole, in face of uncertainty, may not be what is good for any individual when sued." Kent D. Syverud, *What Professional Responsibility Scholars Should Know About Insurance*, 4 CONN. INS. L.J. 17, 19 (1998). "The needs of the many outweigh the needs of the few or the one." Spock's dying words, with help from Captain Kirk. STAR TREK II: THE WRATH OF KHAN (Paramount Pictures 1982).

In addition to being policyholders, the public's interest in insurance can hardly be exaggerated. (The public's intellectual interest in insurance can, alas, readily be exaggerated.) We are all potential tort victims, to be affected by a tortfeasor's insurance coverage, and we are all safer or less safe as insurers create incentives to prevent their insureds from harming us as consumers. We are also taxpayers, whose contribution to the public fisc may increase if government pays what insurance does not. Who would benefit from the introduction of actuarial data in court? The conflicting needs of the public will come into play in the normative question but the incentives of the three direct players will help answer the descriptive question first.

Section one is optional reading; it presents a primer on actuarial data and premium calculation. The core of the article lies in section two. Section three analyzes the limited ways courts already use actuarial data, in and outside of insurance. The concluding section presents some casual empirics on why insurers—those with direct access to the data—do not bring it to court.

The difficulty for both insurers and policyholders is that unless the courts are willing to adopt a rule that benefits one over the other, whether increased use of actuarial data will harm or hurt individual players is a factual question that can only be answered with certainty if the experiment takes place. This article invites the experiment.

I. ACTUARIAL DATA AND PREMIUM CALCULATION (OPTIONAL READING)

Two basic types of data can be submitted to support actuarial claims. First, the underlying statistical data used to price a category of risk, which would show what was being "counted" in a particular risk. This is the actuarial data proper. The second type is the *application* of the actuarial data to evaluation of a particular policyholder or risk, which is performed by an underwriter. Either would need to be tied to the premium charged the policyholder (or the decision to offer the coverage at all) and the language in the policy. For example, the additional premium charged for a hurricane endorsement would be linked to the data under the hurricane risk, the language of the endorsement, and highlighted by the fact that those without the endorsement pay less.

Actuaries and underwriters create and analyze the relevant data. An actuary is a "statistician who determines the present effects of future

contingent events, esp[ecially], one who calculates insurance and pension rates on the basis of empirically based tables.”¹⁰ “Underwriter” and “underwriting” have their etymology in Lloyd’s Coffeehouse, where those willing to take part in the insuring of a ship’s outgoing cargo would write their names under a description of the ship, journey, and goods.¹¹ Today, underwriters use the data compiled and created by actuaries to evaluate a specific risk seeking insurance.

Taking the simple case of life insurance, “the actuary develops the mathematical models to be used to analyze the data, and the underwriter applies the findings of the” medical professionals and actuaries “into the underwriting decision made for a single proposed insured.”¹² Doctors and medical researchers report the initial facts of death by cause and age. Actuaries collect and interpret that data, creating tables into which people of various ages with various traits and medical histories fall. Underwriters then apply those tables, with judgment, to an individual applying for life insurance and recommend a certain place on the table or recommend rejection. The price that corresponds to a place on the table will depend on the insurer—their costs, profit expectations, and more.

How much to charge for an insurance policy is not simply a calculation of the expected risk of loss. The premium charged includes the actuarial premium and the non-actuarial premium.¹³ Underwriters calculate the actuarial portion for a particular risk by applying “statistical data and judgment. Probably in no case is either the sole basis for a rate; every conceivable combination of the two is found.”¹⁴

¹⁰ BLACK’S LAW DICTIONARY 41 (9th ed. 2009).

¹¹ Hence the individual members of the Lloyds insurance market are called Names. See PETER L. BERNSTEIN, *AGAINST THE GODS: THE REMARKABLE STORY OF RISK* 90-91 (1998).

¹² F. Daniel Perkins, *Can “Sound Actuarial Principles” Be Found in Life Insurance Underwriting?*, 38 *TORT TRIAL & INS. PRAC. L.J.* 125, 138 (2002).

¹³ The actuarial portion may include “feature rating,” where “[d]ata collected over the years and intuitive hunches by insurers suggest which features are correlated with loss rates.” “Experience rating,” which is coupled with feature rating, “uses the loss experience of the insured during one period to help set the premiums charged” in the next. KENNETH S. ABRAHAM, *DISTRIBUTING RISK: INSURANCE, LEGAL THEORY, AND PUBLIC POLICY* 72 (1986).

¹⁴ Ralph H. Blanchard, *The Basis of Premium Rates*, *J. OF AM. INS.* (Feb. 1928), reprinted in RALPH H. BLANCHARD, *RISK AND INSURANCE AND OTHER PAPERS* 159, 161 (1965). What was true in 1928 is true today.

The non-actuarial premium covers costs and profits, broadly. It may be calculated as a percentage of the actuarial premium, as a lump sum, or some combination of the two.¹⁵ This bifurcation does not affect the applicability of actuarial data in court; the question remains whether the insurer has made a calculation of the relevant loss or made a calculation that clearly excludes the loss.

Imagine a 1% chance that your house will burn down in the next year, leading to a \$100,000 insured loss. The expected loss is \$1,000 = (\$100,000)(.01). For any one policyholder we can assume the insurer will not need \$1000, but either \$0 or \$100,000. An insurer worthy of the name will have spread the risk across a large pool of similarly situated policyholders and will need to be able to pay out \$1000 per policyholder by the end of the policy year. The risk being carried is therefore a \$1000 risk.

A premium calculation will (usually) include this computation, but what insurers want to determine is how much money it would take to meet the expected cost of the risk, allowing for administrative costs and profit. Even without cost and profit, however, the answer is not \$1000. The time value of money should allow an insurer to collect less than \$1000 in January and have \$1000 or more in December.¹⁶ Fluctuation in interest rates and the insurer's access to investments must be predicted in order to calculate how much needs to be collected from the policyholder in January.¹⁷

In short, insurers would not want to report, and courts would not want to hear, the entire premium calculation for a policy. Nonetheless, the data for a particular clause or the actuarial function behind the structuring of a policy could be explained without excessive fuss.¹⁸

¹⁵ See NEIL A. DOHERTY, INSURANCE PRICING AND LOSS PREVENTION 16-21 (1976).

¹⁶ Come the start of 2009, some are laughing at this statement.

¹⁷ The process gets more complex, of course. On the surface, the timeline for loss payments in the Commercial General Liability context is often years if not decades after the collection of the original premium. This puts a strain on even the most careful predictions of interest rates, taxes, loss amounts, and reserves for payment.

¹⁸ This is not to say there are not questions the data will not answer. There are many. See *infra* text accompanying notes 73-74.

II. USING ACTUARIAL DATA IN INTERPRETATION

Opinions in insurance cases, read collectively, have been compared to “a chapter out of *Alice in Wonderland*.”¹⁹ In some ways this is true; the policyholder does not have to read the policy, yet it is still the basis of the contract. The policy language gets a lay reading except when it doesn’t. Worse, the individual words get a lay reading where there is no lay meaning to the string of the words put together. The parties have no mutual intent in this contract of adhesion but their individual unshared intent might define the contract. Indeed, the policyholder’s “reasonable expectation” of the policy might define coverage even if the expectations come from nowhere and would be dashed by a reading of the policy. But the policy is incomprehensible and thus does not have to be read.²⁰ “It takes all the running you can do, to keep in the same place.”²¹

Still, there is reason in insurance policy interpretation and construction.²² Actuarial data may inject additional reason or at least cut off some of the whims of the Red Queen²³. There are holes in the existing paradigm of insurance policy interpretation into which the data fit. Three main avenues for the useful introduction of actuarial data are explored here: faith, purpose, and intent.

First, the data can show insurer faith—good faith or bad faith and potentially fraud. In particular, it could offer evidence instead of conjecture in the otherwise presumed bait and switch scheme where the insurer attempts to sell big and deliver small. The value here is to prove (or

¹⁹ EMERIC FISCHER, PETER NASH SWISHER, & JEFFREY W. STEMPEL, *PRINCIPLES OF INSURANCE LAW* 90 (rev. 3rd ed. 2006).

²⁰ See, e.g., *Prudential Ins. Co. v. Lamme*, 425 P.2d 346, 347 (Nev. 1967) (policies are “complex instrument[s], unilaterally prepared [by the insurer], and seldom understood by the assured.”).

²¹ So says The Red Queen. LEWIS CARROLL, *THE ANNOTATED ALICE: ALICE’S ADVENTURES IN WONDERLAND AND THROUGH THE LOOKING-GLASS* 210 (New American Library ed. 1960).

²² “The distinction [between interpretation and construction] is, for the most part, not dwelled upon by the courts, with the result that it is difficult to tell which process is being employed.” JOHN D. CALAMARI & JOSEPH M. PERILLO, *THE LAW OF CONTRACTS* 165 (3rd ed. 1987). The distinction will be partially honored here because the two approaches call for distinct uses of actuarial data.

²³ “Alice . . . explained, as well as she could, that she had lost her way. ‘I don’t know what you mean by *your* way,’ said the Queen: ‘all the ways about here belong to *me*.’” CARROLL, *supra* note 21, at 206.

disprove) consistent behavior, consistent from drafting the policy to adjusting the claim and, finally, to taking a litigation position. Inconsistent behavior, either sloppy or greedy, should decline as insurers realize courts now have insight into their inner workings.

Second, actuarial data can reveal, astoundingly, the actuarial *purpose* of the structure of coverage or the actuarial pressure behind an exclusion. Under several doctrines, courts reconstruct, misread, or refuse to enforce a clause because the court can either discern no meaning, no “reasonable reading,” or can discern only an illusory or devious meaning. In these cases, the court may be blameless, left adrift by poor insurer counsel, but the outcome is no less mistaken.²⁴ A court with a good handle on the purpose of a clause may still reject it, of course, but often the court will find a purpose worth protecting. Between an insurer and a lone policyholder in court, the lone policyholder’s needs cry out sharply. Actuarial purpose can show where those needs are misaligned with the needs of other policyholders.

Third, the data can show substantive insurer intent, not just a consistent but a specific intent. It might appear from a casual review of insurance cases that insurer intent has no relevance to courts but this is not so. It can, should, and (sometimes) does matter to courts, regulators, and policyholders that an insurer has a specific intent for taking a particular action. Where the courts have little to go on, insurer intent remains, even disfavored, one piece of the puzzle.

This value of this piece is explored below to address the question whether “tangible property” includes information stored on computers.²⁵ Before the computer age, commercial general liability policies covered damage to tangible property but not intangible property. What should courts have done in the first days of damage to electronically stored information? It was an indeterminate question that could have had one clear input: insurers had not calculated or charged for the risk.

Insurer intent and actuarial purpose can blend together at the edges but they are distinct. To start, the proof may differ. To show actual intent, in lieu of ex post excuses, courts could require an insurer to provide its own underwriting data—proof that the insurer did in fact apply the actuarial data

²⁴ Of course, it is foolish to expect the sins of inadequate counsel to be remedied by asking that same counsel to present actuarial data. Unless the insurer, the policyholder, or the court invokes the data, the lawyer who fails to explain the function of a clause will also fail to introduce the data behind it.

²⁵ See *infra* notes 71-74 and accompanying text.

available. In some circumstances, an insurer will not be able to prove use. In others, an insurer will not want to submit its own underwriting data, thereby forgoing the intent argument. Both could still submit general data on the actuarial purpose of a clause, thus giving the clause another plausible, legitimate meaning from which the court can choose.

Actuarial purpose at times will be more appropriate than intent. In the wake of September 11, insurers collectively stated that they would not seek to exclude coverage under their various war exclusions.²⁶ If the attacks had been more widespread and even more destructive, insurers might not have taken that position. But an insurer would have a hard time proving that its specific *intent* in drafting or including the war exclusion was to avoid losses from a large scale attack by a nongovernmental organization with whom the United States was not at war, declared or undeclared. Where intent would fail, however, actuarial purpose might convince a court that the social purpose of war exclusions—insurer solvency—applied as equally to large scale terrorist attacks as to conventional war.

Finally, those courts that see no slot in the interpretative equation for the input of insurer intent may still be open to the purpose of a clause. Courts adopt a regulatory view toward the health of the insurance industry and the interests of those policyholders not before the court. If the purpose of the clause is insurer solvency, for example, a regulatory court may prefer to protect solvency over literal language interpretation or other values.

But to judicially regulate an insurance clause without access to the actuarial function of the clause is to read a map while ignoring the road. It only makes sense for a court to create coverage if it can accurately identify desired coverage that insurers have failed to perceive. In litigation, a court can be confident that the policyholder *before* the court wants coverage—of flood damage from Katrina, for example. And so do other similarly situated policyholders, at least *ex post*. But non-losing current and future policyholders of the state might prefer a ruling that does not cause insurers

²⁶ Under the ISO war exclusion, the insurer “will not pay for loss or damage caused directly or indirectly by . . . (1) war, including undeclared or civil war; (2) warlike action by a military force, including action in hindering or defending against an actual or expected attack, by any government, sovereign or other authority using military personnel or other agents; or (3) insurrection, rebellion, revolution, usurped power, or action taken by governmental authority in hindering or defending any of these.” Insurance Services Office, Cause of Loss Special Form CP1030, *available at* <http://www.vwcos.com/documents/forms/CP1030-0402.pdf>.

to flee or prices to rise.²⁷ The court faces a tradeoff, although it may not know or care to know.

Let us assume that insurers fail to provide the best coverage the market will bear. Policyholders as a group want and are willing to pay for a particular coverage, call it the happy clause. Should not judges simply award that coverage to the policyholder seeking redress in court? Courts could read the happy clause into policies, perhaps under the reasonable expectations doctrine because policyholders reasonably expect it.²⁸ After all, if judges regularly award happy coverage, in theory insurers will soon sell what they are being forced to provide. Sure, the premium will rise to reflect the cost of the new coverage but recall that in this example policyholders as a group are willing to pay.

Still, policyholders in general may ask the court not to do them any favors. The specific preference of a given policyholder after a loss can easily be at odds with the preference of policyholders generally. Moreover, policyholders' preferences are not simply cumulative. Case by case, policyholders might approve of each of one thousand subtypes of coverage at various prices. There are therefore one thousand "units" of coverage that Every Policyholder would be willing to buy at its market price. However, the Policyholder is not willing to buy all thousand because he has a limit on how much he is willing to devote to insurance, say \$1000 a year.

If he can rank the units of coverage, his ideal policy will include unit 1 to unit 100, or wherever the premium reaches \$1000. If a third party (the courts, legislature, or regulators) instead requires that unit 20 be expanded or that unit 500 be added, he either will have to let unit 100 go or pay more than his maximum price. Of course, the policyholder cannot go to the insurer in response to a court case and request the substitution; the insurer will decide for him. If the insurer removes nothing, but adds in the court's mandated coverage, the insurer will eventually charge more.

Whatever insurers are selling today, we can be sure that Every Policyholder is not in fact receiving his perfect policy at his perfect price. The question is whether the terms that courts "add" are likely to make the policy better or worse. To the extent courts are regulating insurance—mandating coverage because policyholders deserve it or intractably expect

²⁷ On fleeing insurers, see Daniel Hays, *State Farm to Leave Florida Homeowners Insurance Market*, NAT'L UNDERWRITER PROP. & CASUALTY, Feb. 2, 2009, at 7. See also Richard A. Epstein, *Exit Rights and Insurance Regulation: From Federalism to Takings*, 7 GEO. MASON L. REV. 293 (1999).

²⁸ See *infra* notes 32-34 and accompanying text.

it—that regulatory decision is poorly made if the court does not consider (or makes erroneous assumptions about) the actuarial facts behind coverage.

A. INSURANCE INTERPRETATION WITHOUT ACTUARIAL DATA

A full discussion of how a judge ought to approach the interpretation of an insurance policy is beyond the scope of this article. However, we can loosely assume that a judge will approach the policy as a contract and consider, in order: text, intent, purpose, and public policy.²⁹

Differences across states and a general lack of a clear order for interpretive inputs makes a linear statement about the principles of insurance interpretation mostly fictional.³⁰

The components can be identified, however, even as the order and interaction of the components vary. The components fall into (at least) three categories: basic rules, overrides, and penchants. These lists are illustrative, not comprehensive. Do not read the lists horizontally; public policy is a general override, for example, not one specific to plain language.

<u>Basic rules</u>	<u>Overrides</u>	<u>Penchants</u>
Plain language	Public policy	Pro-coverage
Mutual intent	Bad faith	Ambiguity Probable
Purpose	Industry health	Pay tort victim
<i>Contra proferentem</i>		Avoid moral hazard
Reasonable expectations		Find adhesion

²⁹ The text will be weighed in context with intent and purpose potentially providing part of the context. It may be better to consider the policy's purpose before the specific intent of the parties, as evidence of intent will be extrinsic and potentially less reliable than the self-evident purpose of a policy. On the other hand, undue confidence in the "self-evident" purpose of particular clauses is one of the ripest areas for the introduction of actuarial data.

³⁰ Some of the best works on insurance interpretation are: ABRAHAM, *supra* note 13, at 101; Abraham, *supra* note 5; James M. Fischer, *Why Are Insurance Contracts Subject to Special Rules of Interpretation?: Text Versus Context*, 24 ARIZ. ST. L.J. 995 (1992); Michael B. Rappaport, *The Ambiguity Rule and Insurance Law: Why Insurance Contracts Should Not be Construed Against the Drafter*, 30 GA. L. REV. 171 (1995); Peter Nash Swisher, *Judicial Interpretations of Insurance Contract Disputes: Toward a Realistic Middle Ground Approach*, 57 OHIO ST. L.J. 543 (1996).

Most of these concepts are familiar from ordinary contract interpretation. Two basic rules bear elaboration. As early as 1923, the Supreme Court applied *contra proferentem* in the insurance context: “The rule is settled that in case of ambiguity that construction of the policy will be adopted which is most favorable to the insured.”³¹ The rule takes on subtleties³² but this definition will do here.

Next, the strong form of the reasonable expectations doctrine gives as “the rule that the reasonable expectations of the insured should be honored even if those expectations are unambiguously contradicted by fine-print provisions in the policy.”³³ After first being recognized by Robert Keeton in 1970, the doctrine has enjoyed a sharp upswing followed by a gradual relaxing of its application.³⁴ Many courts trend toward the doctrine’s weak form, ruling that a policyholder’s intent cannot be reasonable if an ordinary reading of the policy would have corrected it.

For one view of how these pieces work together, consider the “textual approach” and “modern contract theory” models proposed by the authors of “Insurance Coverage Litigation.”³⁵

The textual model starts with
 (1) the plain and ordinary meaning of the language,
 which if (2) patently ambiguous,
 triggers (3) *contra proferentem*,
 and/or (4) reasonable expectations,
 and/or (5) the admission of extrinsic evidence.³⁶

Note that this model sensibly does not attempt to claim a hierarchy for the treatments of ambiguity.

If the language is not ambiguous on its face, the textual court may admit extrinsic evidence to determine (a) a latent ambiguity; (b) the intent

³¹ Mut. Life Ins. Co. of N.Y. v. Hurni Packing Co., 263 U.S. 167, 174 (1923).

³² See Abraham, *supra* note 5, at 533.

³³ Abraham, *supra* note 5, at 532 (citing ROBERT E. KEETON & ALAN I. WIDISS, INSURANCE LAW §§ 6.1(a)-(b), at 614-21 (Practitioner’s ed. 1988)).

³⁴ See *infra* Section II.C. Keeton “introduced” the doctrine in Robert E. Keeton, *Insurance Law at Variance with Policy Provisions*, 83 HARV. L. REV. 961 (1970).

³⁵ ANDERSON ET AL., INSURANCE COVERAGE LITIGATION § 2.01 (2nd ed. 2009 & Supp. 2010).

³⁶ *Id.*

of the drafters; (c) any accepted usage in the industry; (d) if the policyholder's interpretation is reasonable in light of the insurer's; and (e) insurer bad faith.³⁷ According to Anderson et al., if after this investigation the court finds the language ambiguous, it will construe it in favor of coverage.³⁸ This is accurate as long as one understands that a finding of ambiguity after the total analysis is a legal conclusion that is the same as finding that the court will construe in favor of coverage.

A formalist judge will start with the policy text and veer off-document with reluctance. Actuarial data could be viewed as extrinsic evidence but it can also be viewed as a contextual fact about the purpose of the document. Purpose, text, and logic will help resolve a budding ambiguity or surface tension between two clauses.

As with non-insurance contracts, the "modern contract theory" model does not completely supplant this textual approach. The tools used are the same, but with a greater focus on reasonable expectations and the intended use and purpose of the policy.³⁹ On the one hand, the contextual, functionalist approach should be less open to actuarial data because it rejects the idea of subjective mutual intent in the standardized adhesive contract.⁴⁰ On the other hand, it also rejects a full obligation on the part of the policyholder to read the policy; as the policy language loses its use in determining the meaning of the contract, other sources, including insurer intent and actuarial purpose, become more useful. Moreover, the actuarial purpose discussed below may be of greatest use to those courts taking a regulatory approach to policy language. Regulating courts are more prone to be in the "modern" interpretative camp.

That said, all judges behave somewhat curiously in the insurance realm.⁴¹ A functionalist or realist judge views a policy through its purpose but will focus on the public purpose; this judge will be more open to evidence about actuarial purpose than to an insurer's specific intent. He

³⁷ *Id.*

³⁸ *Id.* §2.02

³⁹ *Id.* § 2.01.

⁴⁰ See, e.g., E. Allan Farnsworth, "Meaning" in *the Law of Contracts*, 76 YALE L.J. 939, 943 (1967).

⁴¹ Associate Justice of the Supreme Court, Sonia Sotomayor provides one example of unexpected behavior in insurance cases. On the lower bench, then Judge Sotomayor ruled "consistently, across the board in favor of insurers." Andrea Ortega-Wells, *Supreme Court Nominee Sotomayor Shows Record of Favoring Insurers*, INSURANCE JOURNAL, June 1, 2009, <http://www.insurancejournal.com/news/national/2009/06/01/101001.htm>.

may also be more inclined to construct than interpret a policy, reforming inequitable clauses or refusing to enforce “bait and switch” clauses.

Both functionalist and formalist judges will be open to evidence of insurer faith, good or bad. Formalist judges will “tend to focus less on the meaning of particular words and more on the organization of the policy, its history, and insuring intent and purpose—tactics that are more normally employed by functionalist judges.”⁴² Actuarial data address all of this.

B. CONSISTENT INTENT

To generalize, the bad or good faith of an insurer is the backdrop to many interpretive questions. To over generalize, courts tend to assume insurer bad faith. Although faith does not always fit neatly into the interpretation decision tree, actuarial data might make the most strides in demonstrating an insurer’s bona fides. A leading treatise on insurance coverage litigation states the common view that “the insurance coverage that was promised at the time of purchase often disappears down the road when a policyholder submits a claim.”⁴³ Sorting insurers whom this accurately describes from the rest will help the policyholders in court, those not in court, and honest insurers.

The ability to show consistent intent is more important in insurance than elsewhere. While the disappointed party to a contract may be able to claim bad faith, insurers have fiduciary-like responsibilities toward their

⁴² FISCHER ET AL., *supra* note 19, at 95 n.10.

⁴³ ANDERSON ET AL., *supra* note 35 at xx. *See also* Jarrett v. E.L. Harper & Son, 235 S.E.2d 362, 366 (W.Va. 1977). The West Virginia Supreme Court wrote: Insurance is different from any other business. If a man goes into a butcher shop, asks for two pounds of ground meat, and tenders \$2.89 in payment, he will expect his meat to be forthcoming from the grinder. Imagine the scene were the customer to ask for his meat, and be answered that the butcher has no intention to deliver the same. ‘Where is my meat?’ the customer would reply, possibly in other than dulcet tones. ‘I won’t give you any meat,’ replies the butcher firmly. ‘Then give me back my \$2.89 and I shall go elsewhere,’ says the customer. ‘I won’t give you the \$2.89 either,’ replies the butcher, ‘for you must bring a law suit to get it from me.’ Sock! Pow! Blam! And much property damage of a different sort. *Id.*

policyholders.⁴⁴ A few courts go further, describing the insurer's obligation as fully fiduciary.⁴⁵ In general, however, an insurer has some level of duty to its policyholders, one that may "preclude the insurer from taking positions on the meaning of a contract term that would be available in an ordinary commercial setting."⁴⁶ Even in those jurisdictions where insurance policies are interpreted using ordinary contract principles, there is at least one "proviso", that "the contract raises quasi-fiduciary obligations owed by the insurer to the insured [...] [and a]s a result, the insurer has a common-law duty 'not to unreasonably withhold payment of benefits it is obligated to make under the insurance contract.'"⁴⁷

The question of bad faith can go beyond contract interpretation. Numerous states allow a claim for bad faith breach, including for the denial of a claim, as an independent tort worthy of noneconomic and punitive damages.⁴⁸ This can be true where the law does not allow the same claim to sound in tort for other contracts because "[i]nsurance is different. Once an insured files a claim, the insurer has a strong incentive to conserve its financial resources balanced against the effect on its reputation."⁴⁹ If a jurisdiction does not allow the tort claim in insurance, policyholders may still be able to seek additional damages under a bad faith contract claim not permitted outside of the insurance context.⁵⁰

⁴⁴ See JEFFREY W. STEMPEL, LAW OF INSURANCE CONTRACT DISPUTES, § 10.01 (2d ed. 1999 & Supp. 2005); see also William T. Barker, Paul E.B. Glad, and Steven M. Levy, *Is an Insurer a Fiduciary to Its Insureds?*, 25 TORT & INS. L.J. 1 (1989).

⁴⁵ See, e.g., *Tank v. State Farm Fire & Cas. Co.*, 715 P.2d 1133, 1137 (Wash. 1986); *Short v. Dairyland Ins. Co.*, 334 N.W.2d 384, 387 (Minn. 1983).

⁴⁶ FISCHER ET AL., *supra* note 19, at 93. Fischer et al. describe the "tension of sorts between [an insurer's] duty to its shareholders or other investors (to make money) and duty to its policyholders (to pay money)." *Id.* One could add the tension with the duty to other policyholders to pay other money or money in the future.

⁴⁷ *State Farm Mut. Auto. Ins. Co. v. Kastner*, 77 P.3d 1256, 1259-60 (Colo. 2003) (quoting *Farmers Group, Inc. v. Williams*, 805 P.2d 419, 423 (Colo. 1991)).

⁴⁸ FISCHER ET AL., *supra* note 19, at 88.

⁴⁹ *E.I. DePont de Nemours & Co. v. Pressman*, 679 A.2d 436, 447 (Del. 1996).

⁵⁰ See, e.g., *Tibbs v. Great Am. Ins. Co.*, 755 F.2d 1370, 1375 (9th Cir. 1985) ("[T]here is sufficient evidence to support a finding that Great American refused to defend Tibbs in bad faith and is guilty of oppression, fraud, or malice," triggering punitive damages).

In the right case, evidence of “faith” can be readily found in actuarial data. Take an insurer’s premium calculation based on risks A + B but not risk C. If the policyholder claims or the court suspects that the insurer sold C and is now trying to avoid paying for it, actuarial data can show whether or not C was indeed “sold”. More concretely, if a premium calculation were based on wind damage + hurricane damage but not flood damage (and not flood damage during a hurricane), the insurer should be able to prove this to the court.

This next example assumes the insurer can prove a premium calculation based on risks A + B but not risk C; this assumption sets up the question whether courts can ever find meaning in an insurer’s proof of a particular actuarial calculation. First, imagine a duplicitous insurer.⁵¹ This first insurer will be the Tricky Insurer, to be compared to the Breaching Insurer down the road. The Tricky Insurer intends to mislead consumers into believing they are purchasing A + B + C but when loss C occurs, it does not intend to provide coverage. For some percentage of policyholders, when the C claim is denied, the policyholder will take no action and the insurer will have what they want: purchases based on the illusion of C without having to pay out C losses.

But some (smaller) percentage of policyholders will sue for coverage C. Assume one half wins and one half loses. The insurer has to pay the cost of defense against all of these policyholders. It also must pay coverage C for the winning half. Even if the percentage of policyholders who sue is ten percent—a high number—the insurer has collected a premium based on A + B but not C and now has to pay litigation costs and coverage C for some percentage of policyholders.⁵² The insurer puts itself in this situation:

Insurer pretends to cover: A + B + C
 Insurer takes in premium charges: A + B
 Insurer plans to cover: A + B
 Insurer required to pay: A + B + (.05)(C) + litigation costs of C

⁵¹ Most people report being able to do this with ease.

⁵² Some speculate that insurers care not for litigation costs because the time value of the money being retained while the litigation progresses pays for itself, but this cannot be true in small cases. See Alan O. Sykes, “Bad Faith” Breach of Contract by First-Party Insurers, 25 J. LEGAL STUD. 405, 413 (1996).

This arrangement should be profitable for few insurers, far fewer than courts assume can profit from the pretense of coverage.

The sole potential benefit to the tricky type of duplicitous insurer is in an increase in clients. It hasn't increased its rates, after all. To be a profitable ruse, the increase in sales from pretending to provide coverage C would have to outweigh the costs of litigating *and* the cost of covering some percentage of C risks. Recall, no premium was charged for C.

How does the insurer lure in new clients using risk C? Given that most policyholders do not fully read or understand their policies, the creation of the expectation of coverage C would either have to come from (a) fraudulent advertising or (b) taking advantage of a pre-existing expectation of coverage. Fraudulent advertising is a dicey proposition; it will invite the scrutiny of state insurance commissioners and the attention of class action lawyers. On an individual basis, an ad that promises a particular type of coverage will solidly support the consumer's claim of "reasonable expectations" in court. The percentages of those who sue on C and win will rise, leading to a rise in the percentage of consumers who sue. This may explain why most insurance ads make vague promises of "good neighbors" with "good hands" in the "company you keep."⁵³

In the second, more likely scenario, the insurer does not advertise but relies on expectations to ensure sales. But an insurer will be hard-pressed to increase sales based on the expectations of coverage C because the policyholder's *pre-existing* expectations of coverage are not insurer specific—the policyholder will assume *all policies* include coverage for risk C. Thus, a Tricky Insurer who does not charge more but only hopes to increase sales based on a universal assumption is not so tricky. In other words, this scenario seems an unlikely one for swindling insurers, but this is an empirical question courts can investigate on their own using actuarial data.

A more likely model for the swindling insurer is one who calculates and charges a premium based on risks A + B + C but, when C loss occurs, does not intend to provide coverage. (This differs from the insurer above because the premium actually includes risk C here.)

⁵³ State Farm Mutual Automobile Insurance Company: "Like a good neighbor, State Farm is there."® Allstate Insurance Company: "You're in Good Hands with Allstate."® New York Life Insurance Company: "New York Life. The Company You Keep."®

Insurer pretends to cover: A + B + C
 Insurer takes in premium charges: A + B + C
 Insurer plans to cover: A + B

The benefit here is obvious: collect money for C and keep it. This Breaching Insurer may be similarly happy where policyholders either fail to claim the coverage or do not sue when denied. If the pretense of providing for C is in the ether, not an ad or the policy language, the insurer can hope that many will not sue after a loss. But if courts look to the actuarial data when the policyholder does sue, the court should be able to readily grant the C coverage paid for (assuming the policyholder has suffered C).

This will be most straightforward where risk C falls within the general grant of coverage, say to a small guesthouse near the insured home, but is removed by an exclusion, such as if the guesthouse is used as an office part-time.⁵⁴ Many homeowners policies “cover other structures on the ‘residence premises’ set apart from the dwelling by clear space,” but not “[o]ther structures from which any ‘business’ is conducted,” where “business” is defined as “[a] trade, profession or occupation engaged in on a full-time, part-time or occasional basis.”⁵⁵ If the premium charged takes into account the loss data for all “other structures” but fails to exclude those losses where “business is conducted,” the consumer has been charged for and deserves compensation.

Of course, the set-up need not be so formalized or consistently applied. Many people and judges seem to believe that insurers randomly deny payment for losses that are covered, for which the policyholder has paid a premium. Actuarial data should be able to give a clearer picture of how often this happens. If courts regularly examine actuarial data, as time goes by insurers should become less and less willing to charge for coverage and then deny the claim.

In addition to tricks and breach, there is a middle way. An insurer may have created an expectation of coverage from past sales; all prior policies did cover risk C, at price P. When a new day dawns, the insurer

⁵⁴ INSURANCE SERVICES OFFICE, INC., STANDARD HOMEOWNERS POLICY, HOMEOWNERS 3 – SPECIAL FORM HO 00 03 10 00 1, 3 (1999), *quoting* p. 3, Section I – Property Coverages, B.1. & B.2.c. and p. 1, Definitions, B.3.a, *available at* <http://www.mypolicyforms.com/ho3/default2.aspx>. The definition of “business” contains other provisions not relevant here.

⁵⁵ *Id.* at 1, 3.

continues to charge P (or even slightly less than P) but removes risk C from the policy language and stops compensating for C loss.⁵⁶ Behold the Tricky Breaching Insurer. The marketing of the policy is unchanged, so applicants assume C coverage continues to exist. If the coverage is sold at slightly less than P, it appears to the unsophisticated buyer to be a great bargain.

What will the actuarial data show here? The analysis parallels that of the Breaching Insurer if the premium fully includes the risk of C. If the premium is the original P, discounted slightly, the analysis may be identical. It depends on how the discount is taken. If the actuarial data continues to include risk C but a slight discount is taken from the total premium calculation at the end, the court still easily sees that C has been charged for but denied. Breach. If the C component of the premium is itself reduced by some percentage, as long as the basis is still C, breach is found again.

This is not to say that the calculation cannot be too convoluted to prove breach. Indeed, if actuarial data becomes widely used in court, some swindling insurers can be counted upon to re-master the premium calculations to increase convolution. On the other hand, honest insurers who currently operate under a cloud of doubt may work to ensure their calculations show their honesty to its best advantage.

Let us return for a moment to methods of misleading the consumer at the front end. We can imagine either that the insurer actively attempts to provide the illusion of coverage C, as in the first example, or that the insurer simply takes advantage of a pre-existing consumer misconception about that type of coverage. It would be easy for insurers to identify the many misconceptions policyholders have about their insurance coverage. A 2007 phone survey of policyholders found that 71% of Americans with homeowners insurance believed they had full coverage to rebuild after a natural disaster or fire and 72% believed their personal belongings were covered at the cost of replacement.⁵⁷ Their actual coverage had caps below

⁵⁶ This is a common marketing move in the grocery aisle. You may sporadically find your cereal box is slightly thinner or has fewer flakes, although the price per box does not decrease. See, e.g., Jessica Dickler, *The Incredible Shrinking Cereal Box: The packaging may look the same but the amount inside has gone down, that's how companies try to pass on food inflation*, CNNMONEY.COM, Sept. 10, 2008, http://money.cnn.com/2008/09/09/pf/food_downsizing/index.htm.

⁵⁷ Press Release, Metlife Auto & Home, Insurance Surprises: Survey Finds Many Americans Dramatically Overestimate the Level of Insurance Protection

full coverage and would cover the depreciated value of personal items, well below replacement cost.

Do insurers take advantage of mistakes like this? Oren Bar-Gill and Richard Epstein recently asked the more general question “do sophisticated sellers respond strategically to consumer misperception? In particular, do sellers design their products, contracts, and pricing schemes in response to consumer misperception?”⁵⁸ They consider the answer in a debate between behavioral and neoclassical law and economics; Bar-Gill emphasizes that the answer is ultimately empirical.⁵⁹ Contrary to courts’ assumptions, both conclude that it “is probably correct” that “mistakes about a standardized product are not sustainable.”⁶⁰

Insurance may be the exception. Under a reasonable conception of standardization, homeowners insurance policies are a standardized product and consumer mistakes certainly persist. Whether insurers are strategically taking advantage of these mistakes can be answered by actuarial data. And while Bar-Gill is correct that the question is ultimately empirical, the analysis of premium calculation presented here can provide some initial answers even before the actuarial data is reviewed by courts.

The point here is that if the insurer attempts to swindle policyholders by the second breaching method—charging for A+B+C but not covering C—actuarial data should reveal the C charge and confirm that policyholders are due C coverage. The assumption here is not that there are only honest insurers, but that actuarial data can either keep insurers honest or at least allow courts to identify the dishonest ones.⁶¹

They Have (July 10, 2007) (available at <http://www.metlife.com/about/press-room/us-press-releases/2007/index.html?compID=518>).

⁵⁸ Oren Bar-Gill, *The Behavioral Economics of Consumer Contracts*, 92 MINN. L. REV. 749, 749 (2008); Richard A. Epstein, *The Neoclassical Economics of Consumer Contracts*, 92 MINN. L. REV. 803, 807 (2008).

⁵⁹ Bar-Gill, *supra* note 58, at 751.

⁶⁰ *Id.* at 750.

⁶¹ On the flip side, when a policyholder is dishonest and commits fraud in applying for a policy, state law often requires the insurer to prove that the fraud was material; the insurer must be able to show that it would not have issued the policy or would have required different terms in the absence of the policyholder’s misrepresentation or concealment. Actuarial data can inform this question too. *See, e.g., Hill v. Allstate Ins. Co.*, No. 04-CV-0865-REB-CBS, 2006 WL 173693, at *2 (D. Colo. Jan. 24, 2006) (Insurer “charged a premium based on the actuarial risk associated with a one-household [auto] policy,” and “[a]bsent the fraudulent misrepresentations and concealment” of the policyholders “it would not have

C. SPECIFIC INTENT

Policyholder intent has a settled place in insurance interpretation. The traditional form of the reasonable expectations doctrine gives as “the rule that the reasonable expectations of the insured should be honored even if those expectations are unambiguously contradicted by fine-print provisions in the policy.”⁶² Not all courts go so far, however. Many courts are more likely to hold that a policyholder’s intent cannot be reasonable if a reading of the policy would have corrected it.

Courts using the milder version of reasonable expectations recognize the “rule that the policy must receive a reasonable interpretation consistent with the parties’ object and intent.”⁶³ Courts that are interested in insurer intent (and expectations⁶⁴) will consider the intersection of policyholder and insurer intent, but insurers have two ongoing problems in proving intent.

First, given the presumption of poor faith, courts are simply less likely to believe an insurer. An insurer may be taking a litigation position or the insurer’s lawyer may simply be cleverly supporting whatever interpretation supports the insurer’s cause in this case. By contrast, if a policyholder states that it expected or intended a particular coverage, or that it would have expected coverage had it read the policy language, courts are inclined to believe the policyholder if the claim is at all reasonable.

Second, insurers tend to point to the language of the policy for proof of their intent. This is a textbook contractual approach but it severely limits the insurer where the court finds the language confusing and inaccessible to the policyholder. Moreover, it conflates what could be two separate thrusts for the insurer: one about the natural reading of the language and one about the value of its own intent. Both of these problems could be solved with the introduction of actuarial data. The data are a separate and different input for proof of the insurer’s intent. More

issued and renewed the policy as it did.”)

⁶² See Abraham, *supra* note 5, at 532 (citing ROBERT E. KEETON & ALAN I. WIDISS, *INSURANCE LAW* §§ 6.1(a)-(b), at 614-21 (Practitioner’s ed. 1988)).

⁶³ *Brown v. Ind. Ins. Co.*, 184 S.W.3d 528, 540 (Ky. 2005) (quoting *St. Paul Fire & Marine Ins. Co. v. Powell-Walton-Milward, Inc.*, 870 S.W.2d 223, 226-27 (Ky. 1994)).

⁶⁴ “Party expectations are a cousin of party intent.” FISCHER ET AL., *supra* note 19, at 97.

importantly, the data are actual *proof* of intent; insurers would no longer have to rely on their mere claims of intent in the face of skeptical courts.

If courts could be made more comfortable that an insurer's representation of its intent *was* its actual intent, the question becomes what to do with the knowledge. On the day the contract was formed, the insurer had a definite intent as to what the words meant; the (consumer or unsophisticated) policyholder had a general intent—cover my losses if my house is partially or totally destroyed. Karl Llewellyn called a consumer's intent in a standardized contract the “*blanket assent . . . to any not unreasonable or indecent terms the seller may have on its form, which do not alter or eviscerate the reasonable meaning of the dickered terms.*”⁶⁵ In keeping with this, not all reasonable expectations options require the policyholder to have had an actual expectation before the loss occurred; many courts are content to ask what the policyholder would have expected had the question come to mind. Indeed, the policyholder is rarely asked to explain, let alone prove, the source of his expectation.

But what if the policyholder would not have had a particular expectation even had the question come to mind? For example, does your homeowners policy provide coverage if the fence around your yard is damaged?⁶⁶ Does it provide coverage for limited personal effects lost in hotel rooms while traveling? (Yes, to some extent).⁶⁷ If your spouse intentionally burns down your house, are you still covered for your half of the loss? (It depends).⁶⁸ Does the pollution exclusion exclude coverage

⁶⁵ KARL N. LLEWELLYN, *THE COMMON LAW TRADITION—DECIDING APPEALS* 370 (1960) (emphasis added); *see also* Swisher, *supra* note 30, at 570 n.77.

⁶⁶ Your homeowners policy may cover your neighbor's fence—if you negligently set a fire in your backyard. *See* Prather v. Audubon Ins. Co., 488 So. 2d 383, 384-85 (La. Ct. App. 1986). A homeowner was recently paid \$25,000 for damage from Hurricane Katrina to a den and the fencing around his home. Although the plaintiff was not happy with the sum he received, a portion was attributed to the fence. *See* Gustings v. Travelers & Standard Fire Ins. Co., No. 07-4443, 2008 WL 4948837 at *1 (E.D. La. Nov. 18, 2008).

⁶⁷ INSURANCE SERVICES OFFICE, *supra* note 54, at 3 (“We cover personal property owned or used by an ‘insured’ while it is anywhere in the world.”).

⁶⁸ *See* Rachel R. Watkins Schoenig, Note, *Property Insurance and the Innocent Co-Insured: Was It All Pay and No Gain for the Innocent Co-Insured?* 43 *DRAKE L. REV.* 893, 895-96 (1995). If the policy is jointly held between husband and wife, the non-arsonist has suffered an unintentional loss; whether that loss is covered depends on the contract and the state. *See* Randall, *supra* note 1, at 144-45. “Whether the intentional acts of a co-insured will defeat coverage for an

for the injury to people who inhaled fumes from floor cleaner? (Depends on the jurisdiction).⁶⁹

If the policyholder would not have had a particular view, it could not be disappointed in a lack of coverage. If the insurer has a particular intent, one acted upon in premium calculation, it has something to lose from an adverse interpretation. There is always a decent chance that nonetheless a court will apply *contra proferentem*, construing the language in favor of coverage. The undercurrent of *contra proferentem* is protection of the policyholder against a scheming insurer and “encouragement” to the insurer to draft more clearly. In circumstances where neither of these motivations applies, courts that give lip service to valuing the insurer’s intent could pay up, with the aid of actuarial data.⁷⁰

For example, commercial general liability policies provide coverage for liability arising from damage to “tangible property.”⁷¹ Data stored on computers came along, quickly followed by possible liability arising from its destruction. When policyholders first started seeking

innocent co-insured turns on the exclusionary language used in the policy. A policy excluding losses caused by intentional acts of ‘any insured’ or ‘an insured’ creates a joint obligation among co-insureds and bars coverage for both the malefactor and innocent co-insureds. Where the policy uses the words, ‘the insured’, the obligation is several, and the exclusion applies only to the insured who intended the act and caused injury, not an innocent co-insured.” *Id.* See also *N.J. Mfrs. Ins. Co. v. Carney*, No. 3:04-CV-2468, 2006 WL 2092571 at *3-4 (M.D. Pa. July 26, 2006) (holding the intentional act of one insured excludes coverage for the innocent co-insured under the language “an insured” or “any insured”; but a wife’s arson does not stop her husband’s recovery when he is the sole owner).

⁶⁹ Compare *Nav-Its, Inc. v. Selective Ins. Co. of Am.*, 869 A.2d 929 (N.J. 2005) (holding that the pollution exclusion in a Commercial General Liability policy applies only to traditional environmental pollution, not to indoor chemical use) with *Fireman’s Ins. Co. v. Kline & Son Cement Repair, Inc.*, 474 F. Supp. 2d 799 (E.D. Va. 2007) (holding that fumes from floor sealant were “pollution” within the definition of the pollution exclusion).

⁷⁰ Some will say that the two animating factors behind *contra proferentem* always apply. At a minimum, neither is at stake when older policy language is applied to emergent and new risks.

⁷¹ The policyholder is covered for “property damage,” defined as “physical injury to tangible property . . . or loss of use of tangible property that is not physically injured.” INSURANCE SERVICES OFFICE, INC., COMMERCIAL GENERAL LIABILITY COVERAGE FORM CG 00 01 10 01 15 (2000) available at <http://www.certifiedriskmanagers.com/NewISOforms.htm>.

coverage for damage to computer files, courts were at a loss.⁷² Consider how a court could attempt to answer the question.

First, the language is not ambiguous as much as indeterminate. Some courts found ambiguity but most found the language unambiguously supported the judge's individual sense of whether electronic data was tangible property; opinions were widely split, yea or nay. Some courts had a better understanding than others of the physical space (however miniscule) electronic data occupy, and these courts found the data to be tangible property. While reading the words and applying them accurately to a new context has some basis for support, it does require "tangible" to do work it was not selected to do; "tangible" property was meant to set apart intangible property, such as intellectual property. Like threats to intellectual property, threats to electronic data come from different sources. These threats require different risk calculations than threats to tangible property.⁷³

Second, the courts could have turned to the reasonable expectations doctrine. But, in the earlier years, the policyholders either did not have any expectation or would not have expected coverage. (The earlier years were when the question was relevant; insurers eventually addressed the split between tangible property and electronic data with policy options.) The actuarial data from these years would not have included the risk of loss to electronically stored data—an example of the lack of data providing definite information.

Third, those courts finding the language ambiguous could have construed in favor of the policyholder, including electronic data in tangible property. This is not necessarily a good policyholder outcome because the newly found coverage will cost future policyholders. Again, however, many courts were not willing to find the language ambiguous.⁷⁴

⁷² See, e.g., *Computer Corner Inc. v. Fireman's Fund Ins.*, 46 P.3d 1264, 1266 (N.M. Ct. App. 2002) (electronically stored data is "tangible property"); *Am. Online, Inc. v. St. Paul Mercury Ins. Co.*, 207 F. Supp. 2d 459 (E.D. Va. 2002) (electronically stored data not "tangible property").

⁷³ Of course, threats to physical property also threaten the physical storage of electronic data. When fire destroys a computer, it destroys any data stored on that computer but electronic data are subject mainly to electronic threats.

⁷⁴ Try this experiment. Show the policy language on tangible property to several friends and ask each one (a) if the clause covers electronically stored data and (b) if the question is close, i.e. if the language is ambiguous as applied to electronic data. You will find that people vary in their answer to (a) but do not believe (b).

Fourth, courts could have considered the insurer's intent, not as asserted but as shown by the actuarial data. Insurers had not been counting electronic data in the tangible property loss statistics. They were not charging for the risk of loss to electronically stored materials. In other words, one of the parties had a specific intent about "tangible property" and relied on that intent in its initial performance under the contract. Barring a public policy conclusion that insurers should have been on the loss for not addressing electronic data in the policy language at all, the evidence of this specific insurer intent would have provided useful, perhaps definitive, interpretive guidance.

Insurer intent thus could have given courts a tool for solving the "tangible property" dilemma without resorting to individual judge's happenstance first impressions. But there is another benefit to construing language in keeping with insurer intent, at least some of the time. Bringing the interpretation of a clause closer to the coverage intended when the policy was issued decreases what Tom Baker has called the "contract risk".⁷⁵ The contract risk is part of the insurers' risk equation. Baker defines it as "the risk relating to the drafting and interpretation of insurance policies."⁷⁶ Baker partitions the risk liability insurers take on into:

- (1) the baseline risk,
- (2) the developments risk ("relating to developments that change the rate or cost of loss"),
- (3) the contract risk, and
- (4) the financing risk.⁷⁷

As with the other three risks, the higher the contract risk, the more expensive it will be to cover and the more policyholders will have to pay.

Another way to think about contract risk is as a species of ambiguity aversion for which insurers will charge a risk premium. Surveys of actuaries, underwriters, and reinsurance underwriters suggest strong aversion to loss ambiguity. These surveys "illustrate that uncertainty about

⁷⁵ See Baker, *supra* note 7, at 139-40.

⁷⁶ *Id.* at 128, 130. Kenneth Abraham calls this an "uncertainty tax." Kenneth S. Abraham, *The Insurance Effects of Regulation by Litigation*, in REGULATION THROUGH LITIGATION 212, 222-23 (W. Kip Viscusi ed. 2002).

⁷⁷ Baker, *supra* note 7, at 128-30.

losses and ambiguity about probability lead to higher prices.”⁷⁸ The research suggests that the risk premia for ambiguity is inefficient and results in higher prices to policyholders.

There can be good reason for knowingly increasing contract risk. If a clause is avoidably unclear or misleading, for example, an insurer should bear the risk its confusion causes.⁷⁹ Courts want to motivate insurers to decrease the level of contract risk the insurers cause themselves.⁸⁰ Even clear language does not eviscerate contract risk; a clearly written clause may go unenforced because it is against public policy as applied to the particular facts at hand. In other words, the optimum contract risk is not zero. If the risk can be decreased appropriately, however, it will benefit all parties by lowering the price of insurance and decreasing unnecessary litigation.⁸¹

D. ACTUARIAL PURPOSE

A primary and simple reason to show actuarial purpose is to explain why an insurer includes a particular exclusion. Courts often behave as though exclusions are mere traps for unwary policyholders, not a decision by an insurer (in keeping with other insurers) that a particular risk is uninsurable. Actuarial data would be able to demonstrate the legitimate business purpose behind these “natural” exclusions.

Other exclusions often fall into the categories of public policy exclusions or exclusions used by insurers to fight moral hazard. Intentional act exclusions fall into both. These exclusions bar coverage for harm that results from an intentional act of the policyholder, although whether both the act and its result must be intentioned differs among jurisdictions.

In Arizona, for example, there is a wrinkle in the application of intentional act exclusions. The Arizona courts have “abandoned” the *contra proferentem* approach, believing that “a finding of ambiguity is the easy way out since it permits the court to create its own version of the

⁷⁸ Kunreuther, Hogarth & Meszaros, *supra* note 8, at 79. See also sources cited *supra* note 8.

⁷⁹ Inflicting a high contract risk for *unavoidably* complex language is another story. See Michelle Boardman, *Insuring Understanding: The Tested Language Defense*, 95 IOWA L. REV. (forthcoming 2010).

⁸⁰ *But see* Boardman, *supra* note 2, at 1112-17.

⁸¹ As with any increase in litigation certainty, decreasing the contract risk should also increase the resolution of claims out of court.

contract and to find, or fail to find, ambiguity in order to justify an almost predetermined result.”⁸² Ambiguity cannot be willed away, however, so while the Arizona Supreme Court has cautioned against hunting for ambiguity, the courts need an approach to truly ambiguous clauses: the Arizona courts “determine the meaning of the clause . . . by examining *the purpose* of the exclusion [or clause] in question, the public policy considerations involved and the transaction as a whole.”⁸³

In *Transamerica Insurance Group v. Meere*, the policyholder injured a person when acting in self defense.⁸⁴ The policy, like most, had an unambiguous exclusion for intentional acts; the act of self defense was intentional. Should coverage be found to be excluded, giving policyholders some incentive to refrain from self defense? After all, the moral hazard concern behind intentional act exclusions (and their common law parallel) applies weakly, if at all, to self defense.

Consider the approaches a court could take:

1. Apply the clause as written. The meaning is clear.
→ no coverage for policyholder acting in self defense
2. Find the unambiguous clause ambiguous.

⁸² Ohio Cas. Ins. Co. v. Henderson, 939 P.2d 1337, 1339 (Ariz. 1997) (quoting *Transamerica Ins. Group v. Meere*, 694 P.2d 181, 185 (Ariz. 1984)).

⁸³ *Id.* at 1339 (emphasis added). See also Cal. Cas. Ins. v. Am. Family Mut. Ins., 94 P.3d 616, 618 (Ariz. App. 2004).

⁸⁴ *Transamerica Ins. Group v. Meere*, 694 P.2d 181 (Ariz. 1984).

On March 27, 1980, at about 12:30 a.m., Meere and a friend, Leon Ivey, were leaving Lindy’s, a bar in Florence, Arizona. Outside Lindy’s, Meere and Ivey were confronted by several off-duty employees of the Arizona State Prison. Meere alleges that he was quite apprehensive because he had been informed by a captain at Arizona State Prison that a rumor was circulating among the guards that Ivey and Meere, both ex-police officers, were undercover investigators of narcotics flow into the prison. One of the guards, Dennis Pruitt, approached Meere. Meere and Pruitt exchanged words. Pruitt then struck Meere, knocking him to the ground; Meere put up his hands, said ‘I don’t want to fight,’ and was struck again by Pruitt. The two then exchanged blows. The fight ended when Meere knocked Pruitt to the ground and kicked Pruitt as he attempted to get up and come at Meere again. Pruitt lost partial use of an eye as a result of this fight. *Id.* at 183.

- coverage without honor
- litigation increases because policyholders and their counsel conclude there is always an argument for ambiguity

3. Review the actuarial data and actuarial purpose of the clause.

- the actuarial purpose of the clause, which is to avoid creating a moral hazard of lowering a policyholder's disincentive to cause harm, does not apply to true self defense

- the actuarial data would reflect this fact

- policyholder covered, in keeping with the insurer's intent

4. *Meere* court: determine purpose of clause from case law, a treatise, and public policy

- coverage, in keeping with public policy

- route most courts haven't taken

Choices three and four may be equally good outcomes. Unfortunately, a number of courts to look at this issue have taken choices one and two, unnecessarily.⁸⁵ Perhaps these courts felt uncomfortable making public policy openly dispositive⁸⁶ and were unwilling to deny the policyholder at hand coverage while futilely recommending a future change to the state insurance commission.

While the *Meere* court's choice seems sufficient, the approach was not attractive enough to other courts. Use of actuarial data would have provided another avenue, a line of reasoning that could be substantiated, and that was fact-based and provable. Moreover, an insurer aware that a court might consider the actuarial data and purpose of the clause would have been less likely to deny and litigate the claim in the first place.

This case strikes some as silly—"it is not a serious question whether the policyholder should be covered for liability stemming from self defense." But it was (and is) a question in Arizona and elsewhere. "In a majority of cases, courts have held that injury inflicted in self-defense is

⁸⁵ See, e.g., *Lockhart v. Allstate Ins. Co.*, 579 P.2d 1120, 1122-23 (Ariz. Ct. App. 1978); *Auto-Owners Ins. Co. v. Harrington*, 538 N.W.2d 106, 108-10 (Mich. Ct. App. 1995); *Grange Ins. Co. v. Brosseau*, 776 P.2d 123, 126-27 (Wash. 1989).

⁸⁶ The dissent in *Meere* characterized the decision by the majority as one "based on policy to distribute the consequences of the loss on an insurance company." *Meere*, 694 P.2d at 190.

expected or intended under the intentional injury exclusion clause.⁸⁷ Insurers fight these claims. One benefit of even the potential use of actuarial data is that swindling or sloppy insurers should more readily pay claims for risks the insurer has calculated and charged. Note that for this benefit, the policyholders do not need to be able to determine where the insurer is being inconsistent; the insurer's self-knowledge is enough.

Actuarial data should also be able to prove or disprove the existence of a "natural" exclusion. In a typical Hurricane Katrina case, *Tuepker v. State Farm*, the court concluded that a combination of clauses was ambiguous and therefore to be construed against the insurer.⁸⁸ As in most homeowners policies, wind and rain damage were covered and water damage (flood, inundation by water) was excluded. This policy also included a "hurricane deductible endorsement," meaning that the policyholder had paid an additional premium for a type of hurricane coverage (with a deductible).⁸⁹

Finally, the policy excluded losses that were caused in part by a covered cause but would not have occurred without an excluded cause.⁹⁰

⁸⁷ ROBERT H. JERRY, II & DOUGLAS R. RICHMOND, UNDERSTANDING INSURANCE LAW § 63C(c) (4th ed. 2007). *See, e.g.*, *Stout v. Grain Dealers Mut. Ins. Co.*, 201 F. Supp. 647 (M.D.N.C. 1962) (coverage for intentional shooting and unintentional killing of persistent prowler excluded). What is at stake in these cases is the insurer's duty to defend, or pay for the defense, of the policyholder. A policyholder who is found in a civil action to have acted properly in self-defense is not liable to his foe. A policyholder whose claim of self-defense is rejected is liable to his foe but his intentional act and harm clearly falls within the intentional act exclusion. *See also* John Dwight Ingram, *The Expected or Intended Exclusion in Liability Insurance: What About Self-Defense?*, 42 CREIGHTON L. REV. 123 (2009).

⁸⁸ *Tuepker v. State Farm Fire & Cas. Co.*, No. 1:05CV559 LTS-JMR, 2006 WL 1442489 at *4-*5 (S.D. Miss., May 24, 2006).

⁸⁹ *Id.* at *4.

⁹⁰ STATE FARM HOME OWNERS POLICY 24100401-1 (2005), *reprinted in Tuepker* 2006 WL 1442489 at *2.

We do not insure under any coverage for any loss which would not have occurred in the absence of one or more of the following events. We do not insure for such loss regardless of: (a) the cause of the excluded event; or (b) other causes of the excluded event; or (c) whether other causes acted concurrently or in any sequence with the excluded event to produce the loss; or (d) whether the event occurs suddenly or gradually, involves isolated or widespread damage, arises from natural or external forces, or

For these purposes, the policy would therefore exclude wind and rain damage (otherwise covered) if it would not have occurred *but for* the storm surge (excluded water damage). If a house withstood the wind up until the storm surge knocked down a wall, for example, leading wind and rain to damage the inside of the house, all would be excluded. On the other hand, if wind tore a hole in the roof, letting rain in to damage the second story, both would be covered whether or not the first floor was also flooded.

This is a simple explanation of what is admittedly dense policy language. Given that the court had to find some meaning in the language, however, it is a reasonable reading that does not torture or ignore any of the policy language. However, the judge in *Tuepker*—a capable judge who handled many Katrina cases with some skill—found the addition of the hurricane endorsement to the rest of the policy (wind covered, flood excluded) ambiguous.⁹¹ The judge seemed to conclude that since the hurricane coverage must have meant something, and the policy already covered wind and rain damage, it must have meant that the “combined cause” exclusion did not apply during a hurricane.⁹²

This reading may be less reasonable than the one offered above or it may be an improvement; the point is that the actuarial data should have been able to answer the question. The policyholder paid an additional premium for the hurricane endorsement. Which risks did the insurer enter in calculating that premium? Which risks are excluded? In particular, the

occurs as a result of any combination [enumerated excluded causes]. *Id.*

⁹¹ On the other hand, Former Senator Trent Lott’s home was destroyed during Hurricane Katrina, leaving nothing but a slab. State Farm held that a “storm surge,” and not hurricane winds, caused the damage. This meant that Lott’s insurance policy did not cover his damage. See Bob Kemper, *Senator who Lost his Home Sues Insurer*, ATLANTA JOURNAL-CONSTITUTION, Dec. 16, 2005, at C3.

⁹² The court took this reading to avoid a finding that the hurricane coverage was illusory. Insurance policies or clauses that are “illusory”—that would provide coverage under no circumstances—are fraudulent and unenforceable. Policyholders bringing suit against an illusory clause should be able to demonstrate that the insurer knew the clause to be illusory from the insurer’s calculations of expected loss. Conversely, an insurer defending a legitimate clause from attack might be able to provide proof that it expected loss under certain circumstances and planned for it. See, e.g., *Frye v. S. Farm Bureau Cas. Ins.*, 915 So. 2d 486, 491 (Miss. Ct. App. 2005) (Policyholders seeking to prove “phantom” coverage attempted to discover “information relating to the historical make-up and design of the policy, as well as information relating to actuarial composition, loss reserves, claims experience of [the insurer’s] agents, and pure profit.”).

correlated losses from inundation by water, caused by hurricane or any other source, may be too high for insurers to offer solvent coverage. If so, the court's reading of the hurricane endorsement is flawed.

Here, actuarial data could have made a clause sensible to a court; it could have shown not just a lack of fraud or decreased chance of deceit by the insurer but also an understanding of what the insurer was getting at with its language. A circumlocutory or technical clause will not be transformed by data into a thing of beauty, but the court will no longer suspect the insurer of careless randomness. Moreover, the court may see the public policy value behind the enforcement of a "natural" exclusion.

Of course, judges sometimes mandate coverage for public policy reasons and some judges involved in the many Katrina cases clearly saw themselves in that role; thousands of people lost their homes and stingy insurers could be made to pay. "Judges in insurance cases not only make insurance law; sometimes they also make insurance."⁹³ A full explanation of why or how judges award uncontracted-for coverage is beyond this discussion. Courts do it and will continue to do it; the point here is that courts engaging in the regulation of the insurance industry should not do so blindly. They should be armed with a full understanding of the actuarial purpose behind the clauses to be regulated and the likely outcome of any regulation.⁹⁴

⁹³ ABRAHAM, *supra* note 13 at 101. Abraham is right that while "judicial techniques of interpretation frequently create insurance coverage where policies do not provide for it," the total interpretation "practice turns out to be considerably more complicated" than courts simply handing money from wealthy insurers to policyholders. *Id.* at 101-02.

⁹⁴ For example, state regulation can have disastrous results. State Farm is currently withdrawing all homeowners insurance from the state of Florida. The last straw for State Farm was not court action but the denial by the state Insurance Commissioner of State Farm's request to raise rates 47.1 percent. *State Farm Can Go? But Not on Its own Terms*, SUWANNEE DEMOCRAT, Feb. 19, 2009 available at http://www.suwanneedemocrat.com/archivesearch/local_story_044133837.html. State Farm claims that Florida's policies, such as required discounts to its customers "have further reduced needed revenues. During the first three quarters of 2008 (a year with relatively modest catastrophe impact and no major hurricane), State Farm Florida saw its surplus reduced by \$201 million." Hays, *supra* note 27, at 7.

1. Ambiguity

A side note about ambiguity is called for here. More than half of insurance disputes involve some claim of ambiguity.⁹⁵ If a court concludes that a clause (or collection of clauses) is ambiguous, it will be construed against the insurer and in favor of coverage under the doctrine of *contra proferentem*. Actuarial data can help here too, although its role is more amorphous.

In an opinion that is a favorite of insurance textbooks, a court found the phrase “occupied as Janitor’s residence” ambiguous.⁹⁶ The question was whether a floor, a portion of which was used as a massage parlor, was “occupied as a janitor’s residence” because a janitor slept there on occasion. The court attempted to consider the purpose of the clause from the insurer’s perspective but did so by guessing. The court assumed the value of the janitor’s residence was to monitor for fire or trouble, and this may have been so. On the other hand, the court also speculates that the insurer may have wanted the janitor to exclusively occupy the floor, keeping out more dangerous uses.⁹⁷

Actuarial data should have been provided or solicited to prove one of these purposes. The kink here is that the data would not directly have addressed the question of whether the clause was ambiguous; either the clause is susceptible of two plausible interpretations, in context, or it isn’t. Appropriately or no, if one reasonable meaning of a clause is proven and sensible, courts are less likely to find ambiguity. First, while courts often admonish themselves not to “seek” or “create” ambiguity for policy reasons, more factors go into the decision than the sheer ambiguity of the word at hand. Second, once a plausible meaning is available, backed with proof of consistent intent, the mind is less willing to entertain a weaker alternative as proof of legitimate ambiguity.

⁹⁵ See Rappaport, *supra* note 30, at 173. (“The ambiguity rule is probably the most important rule in insurance law.”).

⁹⁶ *Vlastos v. Sumitomo Marine & Fire Ins. Co.*, 707 F.2d 775, 776 (3d Cir. 1983) (warranty as to state of building on day policy issued).

⁹⁷ *Id.* at 779. The court hypothesizes why having a janitor occupying the floor would be in the best interest of the insurance company, and states “[a] full-time resident janitor might also deter prowlers and vandals from entering the building,” among other reasons. *Id.*

III. ACTUARIAL DATA ALREADY IN COURT

This section addresses the concern that courts are ill-equipped to handle actuarial data. Bottom line: courts have no choice. Actuarial data are ubiquitous in modern decision making, too ubiquitous for courts to duck entirely.⁹⁸ The Supreme Court has dealt with the data directly on several occasions, such as evaluating the use of gendered statistics in employee benefits.⁹⁹ Despite controversy, its use is increasing in civil commitment and criminal sentencing.¹⁰⁰ These cases often require expert analysis of the data but, as would be the case in interpretation, they do not require courts to undertake calculations or learn advanced statistics.

Courts accept, on occasion demand, and use actuarial or underwriting data in limited insurance cases now. While these circumstances do not usually involve the interpretation of insurance policies head on, they do indicate that courts have some facility with the data. For example, courts expect insurers to use actuarial or underwriting data to prove materiality in a claim against a policyholder for misrepresentation. A recent California opinion is a typical case:

The insurer's "senior underwriter testified to the misrepresentation's materiality. She explained joint ventures pose

⁹⁸ See Jonathan Simon, *The Ideological Effects of Actuarial Practices*, 22 LAW & SOC'Y. REV. 771 (1988).

Actuarial techniques play a central role in a proliferating set of social practices. They are at the same time a regime of truth, a way of exercising power, and a method of ordering social life. Actuarial practices have not seemed very important nor attracted much interest from social observers in part because they are already so familiar, and in part because they fit so unobtrusively into various substantive projects (e.g., educating, hiring, premium setting) in which they are subordinated as a means to an end. *Id.* at 772.

⁹⁹ See *Los Angeles Dep't of Water & Power v. Manhart*, 435 U.S. 702 (1978) (holding that requiring larger pension contributions from women than men violated Title VII of the 1964 Civil Rights Act, actuarial evidence of longer female life expectancy notwithstanding).

¹⁰⁰ See, e.g., Thomas R. Litwack, *Actuarial Versus Clinical Assessments of Dangerousness*, 7 PSYCHOL. PUB. POL'Y & L. 409 (2001); Christopher Slobogin, *Dangerousness and Expertise*, 133 U. PA. L. REV. 97 (1984); John W. Bagby, *Book Review*, 4 J.L. ECON. & POL'Y 213 (2006) (reviewing BERNARD E. HARCOURT, *AGAINST PREDICTION: PROFILING, POLICING, AND PUNISHING IN AN ACTUARIAL AGE* (2006)).

increased risks, require additional underwriting, and warrant charging ‘an additional premium’ before [the insurer] will cover them. Thus, the misrepresentation is also material because it affected [the insurer’s] evaluation of risk and the amount of the premium charged.”¹⁰¹

In a second example, if a policyholder is turned down for insurance, the policyholder may be able to seek redress for a violation of state law or regulation.¹⁰² The same is true when an existing policy is cancelled or not renewed.¹⁰³ Courts in these cases evaluate the underwriting process but not for the purpose of interpreting policy language.¹⁰⁴ Similarly, insurers can challenge a state insurance commissioner’s rejection of their language or rates, a dispute that will involve the insurer’s premium data.¹⁰⁵

Third, parties can point to actuarial theory or abstract fact without presenting numbers or calculations.¹⁰⁶ For example, insurers can point to

¹⁰¹ *LA Sound USA, Inc. v. St. Paul Fire & Marine Ins. Co.*, 67 Cal. Rptr. 3d 917, 924 (Cal. Ct. App. 2007). *See, e.g.*, *Hill v. Allstate Ins. Co.*, No. 04-CV-0865-REB-CBS, 2006 WL 173693, at *2 (D. Colo. 2006) (Insurer “charged a premium based on the actuarial risk associated with a one-household [auto] policy,” and “[a]bsent the fraudulent misrepresentations and concealment” of the policyholders “it would not have issued and renewed the policy as it did.”). In *L.A. Sound*, the court held that the insurer did not need to produce the specific underwriter who had processed the policyholder’s application. *L.A. Sound*, 67 Cal. Rptr. 3d at 924.

¹⁰² *See* CAL. INS. CODE § 791.10 (2006) (Adverse underwriting decisions; declination, cancellation or nonrenewal of enumerated policies; specific reasons for decision). Discriminatory underwriting was treated under the McBride-Grunsky Insurance Regulatory Act of 1947, but repealed by Prop 103. CAL. INS. CODE § 1861.05 (2009).

¹⁰³ CAL. INS. CODE § 791.10 (2006).

¹⁰⁴ *See, e.g.*, *Gonzales v. Pac. Specialty Ins. Co.*, No. B193051, 2007 WL 2005059 (Cal. Ct. App. July 12, 2007).

¹⁰⁵ *See, e.g.*, *Mass. Auto. Rating & Accident Prevention Bureau v. Comm’r*, 453 N.E.2d 381 (Mass. 1983) (Action by insurers challenging the Massachusetts Insurance Commissioner’s establishment of automobile insurance rates).

¹⁰⁶ As with most forms of evidence or logical support, courts at times include a reference to the actuarial function of a clause or policy where the point is not outcome determinative but merely serves to buttress the court’s conclusion as doubly right. The non-essential support for a legal outcome is not necessarily dictum but nor is it always as well-considered as the evidence upon which the court actively relies. Current judicial use of abstract actuarial principles is more

the fact that a particular type of coverage can be had for an additional premium as proof that, in the absence of that premium, the loss is not covered. Courts may but often do not require the insurer to demonstrate that the policyholder was *aware* of the additional coverage for sale. Thus courts are at times willing to use “inside” insurer information to interpret a clause despite the information asymmetry between the parties. (Focusing on a policyholder’s reasonable expectations, ignoring any insurer role in those expectations, is a more common example of courts using asymmetrical information, but to the policyholder’s benefit.)

Some courts seem to be open to a more sophisticated use of actuarial data. If deeds speak louder than words, however, it should be noted that encouraging language such as this often appears in an opinion that does not *rely* upon the data:

If the primary goal is to fulfill the reasonable expectations of the insured, then there is no need to look at anything beyond the language of the policy itself. If, on the other hand, the primary goal is to give insureds what they pay for, then we should, at the very least, be concerned with the actuarial methods used to arrive at the premium and should look behind the policy language itself.¹⁰⁷

It is an open question whether courts would ever use the same data to see that an insurer gives no more than he *sells*.¹⁰⁸

Fourth, in the specific area of uninsured motorist coverage, some courts have been considering how premiums are calculated in policy interpretation. In these cases courts look at the structure of the premiums charged rather than at specific actuarial calculations. The question arises in

likely to fall into this category because a court need not expend much effort to evaluate these forms of support.

¹⁰⁷ Montano v. Allstate Indem. Co., 92 P.3d 1255, 1261-62 (N.M. 2004). The court went on to “conclude that we need not resolve which rationale to give primary effect.” *Id.* at 1262. Other courts have agreed that the “actuarial methods used to arrive at the premiums [can be] considered to determine whether the insured gets what he pays for.” Rehders v. Allstate Ins. Co., 135 P.3d 237, 248 (N.M. Ct. App. 2006).

¹⁰⁸ In 1934, the Supreme Court expressed a similar sentiment: “While it is highly important that ambiguous clauses should not be permitted to serve as traps for policyholders, it is equally important, to the insured as to the insurer, that the provisions of insurance policies which are clearly and definitely set forth in appropriate language, and *upon which the calculations of the company are based*, should be maintained unimpaired by loose and ill-considered interpretations.” Williams v. Union Cent. Life Ins. Co., 291 U.S. 170, 180 (1934) (emphasis added).

underinsured and uninsured motorist (UM) coverage, where one's own auto policy provides coverage for losses from an accident caused by another driver who has insufficient or no insurance. The premium for UM coverage is usually broken out, so that the driver can see how much is being charged for the UM portion of the policy.

Imagine a judge faced with an auto policy that has a UM coverage limit of \$50,000 and charges a UM premium per car. For example, \$20 per car, so that a one-car policy has a \$20 UM premium and a three-car policy has a \$60 UM premium.¹⁰⁹ If the three-car policy owner gets into an accident with one of the cars, does he have \$50,000 in coverage or \$150,000? Where policy language was ambiguous, courts tended to grant the greater coverage.¹¹⁰

Insurers apparently had not meant to allow this "stacking" of limits; insurers rewrote the policies to charge a unified UM premium. It might appear that this was an insurer sleight of hand. The driver with three cars is financially indifferent to paying \$60 once or \$20 three times. But most courts reviewing the change in premium calculation have concluded that with the new single premium, policyholders have "no reasonable expectation of aggregate coverage."¹¹¹

One court referred to the new single premium as "actuarial and not based on the number of vehicles covered."¹¹² This is twice right and twice wrong: both calculations are actuarial ones and both take account of the number of cars. The new single premium is probably more actuarially accurate (or specific) than taking the premium for one car and multiplying it, but it should still include the increased risk that comes from owning additional cars. Assume a uni-car family with two drivers carries the risk of needing UM coverage, R. Adding a second car to the family might

¹⁰⁹ See *Adkins v. Ky. Nat'l Ins. Co.*, 220 S.W.3d 296, 297 (Ky. Ct. App. 2007).

¹¹⁰ The endorsement page of the policy in *Adkins* read: "The limit of liability shown in the Declarations for each person for Uninsured Motorists Coverage is our maximum limit of liability for all damages ... sustained by any one person in any one accident... This is the most we will pay regardless of the number of ... [v]ehicles shown in the Declarations..." *Adkins*, 220 S.W.3d at 299. See also *Sturdy v. Allied Mut. Ins. Co.*, 457 P.2d 34, 42 (Kan. Sup. Ct. 1969) ("When we pay a double premium we expect double coverage. This is certainly not unreasonable but, to the contrary, is in accord with general principles of indemnity that amounts of premiums are based on amounts of liability.").

¹¹¹ *Adkins*, 220 S.W.3d at 300.

¹¹² *Id.* at 299.

double that risk to 2R because now both drivers can be on the road at the same time, although one would have to see the actual data to know. Adding a third car seems unlikely to raise the risk to 3R, however, even if it is greater than 2R. (The third car might be correlated with being a higher risk rather than directly causing the risk to increase, for example if those who own more than one car per person tend to drive more frequently or hastily.)

Moreover, the reverse implication of the court's description is that the initial per-car premium was not "actuarial." No doubt the initial premium charged for one car did reflect the amount necessary for UM coverage. The shorthand of multiplying that number by additional cars is not as actuarially precise but nor was it pulled from thin air. Insurers presumably found that any more elaborate calculation of the risk per added car was not worth the candle.

This "stacking" question is the one most consistently being answered with reference to actuarial data. In two exceptionally rare circumstances, insurance policies refer directly to actuarial data. First, courts have no choice but to accept actuarial data if the policy is one with a retrospective premium. "A retrospective premium has two components: a basic premium and a conversion loss factor to adjust the premium by consideration of the insured's actual losses during the policy period."¹¹³

Second, in unusual circumstances, actuarial data is incorporated into the contract.¹¹⁴ Even then, courts will not allow themselves to be forced to consider data in a policy; courts have shown their willingness to ignore portions or entire policies that they conclude policyholders could not

¹¹³ *Hartford Fire Ins. Co. v. Terra Ins. Co.*, No. CIV.A.01-5961, 2004 WL 1770298 at *2 (E.D. Pa. Aug. 2, 2004). "An insurance policy with retrospectively-rated premium is sometimes referred to as a form of 'self-insurance' because the policy covers only claims exceeding the maximum premium under the policy." *Id.* See also Douglas R. Richmond, *Issues and Problems in "Other Insurance," Multiple Insurance, and Self-Insurance*, 22 PEPP. L. REV. 1373, 1450 (2002).

¹¹⁴ See, e.g., *Conrad v. Ace Property & Cas. Ins. Co.*, No. CV-05-5117-FVS, 2006 WL 1582376, at *1 (E.D. Wash. 2006) ("The Policy contains the following definition of 'Policy': The agreement between you and us consisting of the accepted application, these provisions, Special Provisions, *actuarial documents*, and the applicable regulations published in 7 C.F.R. chapter IV.") (emphasis added)).

understand. But these examples suggest that judges are able to handle actuarial data.¹¹⁵

As one of the largest insurance markets in the country (and the world), California may be a good place to begin the actuarial data experiment.¹¹⁶ California's courts often provide a prototype of legal evolution for other jurisdictions. "Other states may or may not choose to follow California's example on a particular issue or principle, but they certainly note and examine what California does."¹¹⁷

On the actuarial data front, the California Supreme Court has repeatedly stated that it will look to "the reasonable expectations of the insurer *and* the insured . . . as manifested in the distribution of risks, the *proportionate premiums charged* and the coverage for all risks except those specifically excluded."¹¹⁸ In a recent decision, a court of appeals held:

No rational insurer would wish to undertake such an insuring obligation. It would be literally impossible, *from an actuarial standpoint*, to set appropriate premiums to guard against the risk that an association would enter into multimillion-dollar construction contracts, and then not pay for the construction work. That type of risk would be virtually impossible to underwrite.¹¹⁹

¹¹⁵ These examples show courts have the ability to handle the data on some level but in these last three examples, the data are represented in part in the policy itself. Therefore, these examples are not meant to strongly show that courts are necessarily inclined to use data more broadly.

¹¹⁶ In California, a search to capture every case that includes the word "insurance" and some form of either "underwriting" or "actuarial" resulted in close to 2000 cases. (Westlaw search for "insurance & (underwr! actuar!)".) More specific searches were used as well but one extremely broad search proved useful in locating categories of use.) The majority of these were not relevant to the question of whether or how California courts currently use the data. Many were not insurance cases and many simply had citations to common party names that included the word "underwriters," such as "Universal Underwriters Ins. Co." or "Certain Underwriters of Lloyd's London." Some involve the underwriting of pension plans or other funds that do not involve application of an insurance policy.

¹¹⁷ H. Walter Croskey, *The Doctrine of Reasonable Expectations in California: A Judge's View*, 5 CONN. INS. L.J. 451, 452 n.1 (1998) ("California insurance jurisprudence has considerable influence on that of other jurisdictions.").

¹¹⁸ *Garvey v. State Farm Fire & Cas. Co.*, 770 P.2d 704, 711 (Cal. 1989) (emphasis added).

¹¹⁹ *Oak Park Calabasas Condominium. v. State Farm Fire & Cas. Co.*, 40 Cal. Rptr. 3d 263, 268 (Cal. Ct. App. 2006) (emphasis added). This approach was

Although the California courts do not now entertain the data in the way encouraged here, its use is compatible with the general approach taken because “[t]he goal is to give effect to the reasonable expectations of *both* the insured and the insurer.”¹²⁰

IV. CONCLUSION: RESISTANCE AND REAL WORLD OBSERVATIONS

This article introduces the possibility of improving the interpretation and construction of insurance contracts through actuarial data. As with any innovation, one must ask if there is a barrier beyond simple lack of creativity or inertia to explain its current absence. Each of the main players offers their own potential resistance.

A. INSURERS

Insurers have voted with their briefs, so to speak, in that they have access to data and rarely seek to introduce it in court. As I have argued elsewhere, there is some reason to believe that insurers are resistant to change.¹²¹ Insurers also may assume that courts would not welcome this development or, and only some will find this plausible, it may be that it has simply not occurred to insurers to routinely use this type of data in court.¹²²

Based on casual interviews with insurer counsel, I have collected explanations for why insurers resist the advice of their own counsel that the use of actuarial data would support an important position. In interviewing outside counsel I was most interested in cases where the firm lawyer instigated the idea of using actuarial data and was rebuffed by inside counsel.¹²³

specifically followed in *August Entm't, Inc. v. Philadelphia Indem. Ins Co.*, 52 Cal. Rptr. 3d 908, 914-16 (Cal. Ct. App. 2007).

¹²⁰ *August Entm't, Inc.*, 52 Cal. Rptr. 3d at 913 (emphasis added). As with almost all jurisdictions, the opening move in California is to determine the intent from the language if possible. California courts are not reticent to move on to additional inputs, however.

¹²¹ See Boardman, *supra* note 2, at 1116-17.

¹²² As a lawyer, I would occasionally suggest to in-house insurer counsel that actuarial data be used to prove a particular point. The resistance I encountered rested upon some of the reasons discussed below.

¹²³ In a future project I may ask insurers directly, although I would expect to receive vetted public relations answers.

Most of the reasons given are not general objections but reasons to resist the use of actuarial data in certain cases. The approach appears to be piecemeal rather than a considered policy of always resisting actuarial data. An insurer who has concluded, on net, that the general use of actuarial data is not in their interest may of course object to all policyholder attempts to discover or use it, even in cases where the data could be beneficial.

As expected, sometimes there is no data on point. An outside counsel suggested backing up an insurer's claim about an "additional insured" provision with reference to the data. The insurer response: "There is no data. We sometimes add 'additional insureds' without changing the premium or putting aside additional reserves;" a fact the insurer did not want to advertise.¹²⁴ This does not seem to be an uncommon practice with additional insured provisions.¹²⁵

Similarly, there are circumstances where the data would not be useful because the premium charged does not directly reflect the cost of the risk but instead reflects interest rates, price competition, (past expected but not yet incurred losses from past calculations that turned out to be insufficient), etc. Of course, there are other times when the available actuarial data just does not speak to the question at hand. In these cases insurers may not be resistant to the court reading it, if its production were costless, but they would not seek to introduce it themselves. Given that production is never costless, an insurer may still object to discovery.

Another common response: "We just don't refer to actuarial data in our pleadings and briefs. We never have and nor do our competitors."¹²⁶ A taller objection than inertia: Courts will reject the data but we will be worse off for having offered it; it creates the impression that in order to understand our policies a policyholder would need to be an actuary.

Finally, certain applications of actuarial data in the underwriting process are proprietary. Insurers have moved to quash subpoenas on the grounds that "materials contain[ing] reserving information and actuarial

¹²⁴ As with all of these illustrations, this is a paraphrase of the insurer counsel's response, not a direct quotation.

¹²⁵ "In practice, [additional insured] endorsements that are issued automatically or without charge are usually limited to vicarious liability by express statement." James E. Joseph, *Indemnification and Insurance: The Risk Shifting Tools (Part II)*, PA. B. ASS'N Q. 1, 16 (2009). For a collection of cases in which courts note that no or little additional premium had been charged for additional insured coverage, see Note, *Recognizing the Unique Status of Additional Names Insured*, 53 FORDHAM L. REV. 117, 120 and n.12 (1984).

¹²⁶ Again, this is a paraphrase.

formulas or analysis particular to [that insurer]” are proprietary, in part because in some segments “it is industry standard for each company to develop its own product forms and underwriting systems.”¹²⁷ Similarly, insurers may fear that opening their books will open them to charges of misbehavior, such as price-fixing or redlining by racial data.

B. POLICYHOLDERS

In many cases policyholders benefit from moving away from the “contract context,” and away from what the insurer expected, so that actuarial data might be viewed as a step in the wrong direction. Policyholders have discovered or sought to discover actuarial data in limited cases.¹²⁸ Perhaps the data would open a Pandora’s box that policyholders and their counsel are afraid to open because they are unsure what waits inside.

The small sample of cases shows that policyholders are more likely to seek to discover the data in class actions. The starting assumption is that policyholders seek out the data when they predict or know (from others’ litigation or regulatory history) that it will support their preferred interpretation and seek to exclude it when it will not. Another simplifying but plausible assumption is that policyholders in litigation are pursuing an interpretation of a policy or a clause that provides or extends coverage and the insurer is pursuing an interpretation that excludes or limits coverage.¹²⁹

¹²⁷ *Richter v. Mut. of Omaha Ins. Co.*, No. 06-Misc.-011, No. CV 05-498 ABC (PJWX), 2006 WL 1277906 at *1 (E.D. Wis. May 5, 2006) (assertion of privilege deemed waived).

¹²⁸ *See, e.g., Vos v. Farm Bureau Life Ins. Co.*, 667 N.W.2d 36, 41 (Iowa 2003) (discovery in class action against life insurer for breach of contract and various fraud claims included “actuarial material regarding the policies in question”); *Frye v. S. Farm Bureau Cas. Ins. Co.*, 915 So. 2d 486, 490-91 (Miss. Ct. App. 2005) (denying a policyholder’s motion for discovery of actuarial information); *Robinson v. S. Farm Bureau Cas. Co.*, 915 So. 2d 516, 520 (Miss. Ct. App. 2005) (denying a claim by policyholder plaintiffs that “alleged in their motion for continuance that [the insurer] Southern Farm failed to respond to written discovery concerning the historical makeup and actuarial composition of the disability income coverage.”); *Beller v. William Penn Life Ins. Co. of N.Y.*, 828 N.Y.S.2d 869, 875 (N.Y. Sup. Ct. 2007) (actuarial data about life insurance plan sought and obtained in class action).

¹²⁹ Note that in considering the benign or evil motives of each player, this last does not require a dark view. There are policyholders who would prefer their

Apart from wanting to use specific data that support their view and exclude data that do not, the increased use of actuarial data should help policyholders in general across time. Start with the arguable but contentious assumption that insurers tend to offer the coverage the market demands. This does not mean that insurance coverage is a wonderland for consumers. Some coverage that policyholders would like to purchase is not sold because policyholders are not willing to pay the minimum price insurers demand. Some coverage is not sold because insurers are not willing to insure the type of event. To maintain risk distribution in the risk pool, insurers do not want to sell coverage for correlated large losses, such as flood coverage in a flood plain. For reasons of public policy, moral hazard, and adverse selection, insurers limit other options that individual policyholders desire.

If there are sound actuarial reasons to limit coverage, the limitation is generally to the benefit of insurers *and* policyholders. Courts can make a mistake when they interfere with these types of insurer choices. They do make a mistake when they choose the needs of one plaintiff policyholder over all other policyholders. Actuarial data can improve the decision-making of courts inclined to regulate.

C. COURTS

Are courts willing to entertain actuarial data more regularly? The recitations of canons of insurance interpretation supply conflicting statements. These are sometimes one on top of another, as though the proximity of delivery will blind the reader to the inconsistency. Take a representative summary from the Wisconsin Supreme Court:

The goal of construction is to ascertain *the true intentions of the parties* to an insurance contract. In the case of an insurance contract, the words are to be construed in accordance with the principle that the test is *not what the insurer intended* the words to mean but what a reasonable

insurance to cover a loss but who read the policy to exclude it and who thus do not sue. A particularly rosy view would be required to assume that every policyholder in court firmly believes in the accuracy of their interpretation, but many no doubt do. Likewise, insurers pursuing no or low coverage positions in court is not proof that all insurers deny all claims all the time; if the insurer pays what the policyholder expects, or close to it, there is no need for court.

person in the position of an insured *would have understood* the words to mean.¹³⁰

Translation of the interpretive goal:

- (1) to find the “true,” i.e. subjective, intent of both the policyholder and the insurer
- (2) to *not* find the intent of the insurer
- (3) to find the objective understanding of a reasonable policyholder, which may or may not equate with
 - (a) the intent of *this* policyholder, or
 - (b) the understanding of this policyholder

Given these goals, actuarial data have a role to play. If contradiction accepted, a court is interested in knowing the insurer’s intent, the court would be interested in actuarial intent. If a court is interested neither in the policyholder’s intent (as it may not have one) nor the insurer’s intent, it will pursue goal three.

Under this goal, the court will pursue the policyholder’s original understanding, not the original intent. The goal is to decipher what the policyholder would have understood the words to mean. With this aspiration, often there is no answer to the question or there are several answers; the language is ambiguous. Courts are then open to explanations of:

- (1) why one of several interpretations is reasonable (actuarial purpose),
- (2) what one party thought or did (actuarial intent),¹³¹
- (3) whether the insurer acted in bad faith (consistent intent), to decide how far to construe against the drafter, and
- (4) the future drafting and premium consequences of reading the language a particular way (actuarial purpose).¹³²

¹³⁰ Wood v. Am. Fam. Mut. Ins. Co., 436 N.W.2d 594, 599 (Wis. 1989) (internal citations omitted) (emphasis added) (partially overruled on grounds not relevant here).

¹³¹ Wisconsin courts have taken the position that “the policy should not be rewritten by construction to cover matters not contemplated by the insurer nor paid for by the insured.” Vidmar v. Am. Fam. Mut. Ins. Co., 312 N.W.2d 129, 131 (Wis. 1981).

¹³² “[W]ere insurers not able to enforce reasonable conditions upon their liability, in accordance with actuarial standards and projections, their industry

As with other types of evidence, there will be debate about whether most judges are qualified to evaluate actuarial data. Aspects of actuarial calculation obviously require higher math skills and training. But one does not need to be an actuary to understand the pricing mechanism or to see that a certain risk was included in a calculation and therefore that the risk should be covered. The relevant “pricing principles were in operational use 14 years before the publication of Adam Smith’s *The Wealth of Nations* and considerably before the nineteenth century classical economists had turned their attention to a general theory of prices.”¹³³

Of course, the ability to use data is not sufficient; courts have to be willing to use it. The limited judicial use of the data now shows both ability and willingness to handle the evidence when it is part of set interpretative routine.¹³⁴ Those places where it has arisen in recent years suggest that courts use the data when a framework for its use is obvious or available.

Whether particular players in the insurance interpretation game will support the introduction of actuarial data into the process is an empirical question. To the extent the data can accurately reveal consistent (or inconsistent) pricing and paying behavior, it will supplant what is now mere assumption. This will decrease swindling insurer behavior by increasing the chance of getting caught. It will also free honest insurers from excessive court interference.

Where the data reveal insurer intent, hewing closer to that intent will increase contract certainty, decrease cost, and reduce litigation. With an understanding of actuarial purpose, the many courts engaged in the regulation of insurance contracts can avoid current mistakes that harm insurers and policyholders. Rather than regulate blindly, based on the current needs of the one policyholder before it, proof of actuarial purpose will allow a court take account of all the absence policyholders and the health of the industry.

would be hindered in its ability to serve the important function it does in our society.” *Hartzo v. American Nat. Property. & Cas. Ins. Co.*, 951 So. 2d 1120, 1124 (La. Ct. App. 2006).

¹³³ James C. Hickman & Robert B. Miller, *Insurance Premiums and Decision Analysis*, 37 J. OF RISK AND INS. 567, 568 (1970).

¹³⁴ See *supra* Section III.

THE LAW AND ECONOMICS OF FIRST-PARTY INSURANCE BAD FAITH LIABILITY

Sharon Tennyson *

William J. Warfel **

States differ in the legal avenues available to policyholders to pursue actions against their insurers for bad faith in claims settlement. This article discusses the various approaches to first-party insurance bad faith law that have been taken by the states, and discusses the potential benefits and costs of different approaches. Regimes that are likely to grant large damages awards to aggrieved policyholders provide the greatest deterrent to insurer bad faith; but such regimes may also create incentives for fraudulent insurance claiming and disincentives for rigorous claims investigations by insurers. This article evaluates the empirical relevance of these potential incentive distortions through an analysis of automobile insurance claim settlement data in states with different bad faith regimes. The data show that claim characteristics and claim investigations differ significantly in states which permit tort-based bad faith from those in other states, in ways consistent with the hypothesized effects.

* Sharon Tennyson is Associate Professor in the Department of Policy Analysis and Management at Cornell University. She holds a Ph.D. in economics from Northwestern University with a specialization in industrial organization and regulation. Her professional interest centers on economic and policy analysis of insurance markets, and she has published widely on these topics. Sharon is a noted expert on insurance rate regulation and insurance fraud, and is a frequent speaker on these issues.

** William J. Warfel, Ph.D., CPCU, CLU, is Professor of Insurance and Risk Management at Indiana State University. He received his doctorate from Indiana University in 1990. His research focuses largely on the interface of law and insurance; he has published extensively in the CPCU eJournal, The John Liner Review, Risk Management Magazine, and various legal publications. Also, to date, he has been retained as a testifying and consulting expert witness in about 45 cases; his specialty includes breach of contract, bad faith, and agent/broker liability issues. He can be reached at wwarfel@indstate.edu.

I. INTRODUCTION

The idea that insurers should be penalized for unfair claim settlement practices involving first-party insurance coverage is a relatively recent development in the long history of insurance law. Historically, insurers were not penalized for unfair claim settlement practices involving first-party insurance coverage. Pursuant to the nineteenth-century English common law rule articulated in *Hadley v. Baxendale*,¹ the policyholder was allowed to recover only those damages that were in the contemplation of the parties to the contract at the time the policy was purchased.² This rule meant that damage awards could not exceed the amount specified in the insurance policy, even if the breach of contract was intentional on the part of the insurer.

Beginning in the early 20th century, this rule was modified through enactment of a progression of state statutes including model legislation on unfair trade practices developed by the National Association of Insurance Commissioners (NAIC) in 1959 and amended in 1972.³ The model Unfair Trade Practices Act prohibits specified acts by an insurer when “committed flagrantly and in conscious disregard of” the statute, or “with such frequency as to indicate a general business practice.”⁴ The 1971 model legislation and many of the “statutes originally adopted by the states” were silent as to whether it created a private cause of action.⁵ This silence meant the insured’s only recourse was to file a complaint with the state insurance department.⁶

¹ See *Hadley v. Baxendale*, (1854) 156 Eng. Rep. 145 (Exch. Div.).

² See *id.*

³ 4 NAT’L ASS’N OF INS. COMM’RS, MODEL LAWS, REGULATIONS, AND GUIDELINES: UNFAIR TRADE PRACTICES ACT (2008).

⁴ *Id.* See also Laureen Regan & Paul M. Rettinger, *Private Rights of Action Under State Unfair Claims Settlement Practices Acts: A Review*, J. INS. REG. (1998) (identifies all 14 of the prohibited acts). This model legislation, or some variant of it, has been adopted by all U.S. states. Efforts to expand first-party bad faith liability continue today. In Connecticut, for example, a proposal has been made to delete the requirement that violations occur with such frequency as to indicate a general business practice. If this proposal is adopted, a single violation would be sufficient to constitute bad faith.

⁵ Regan & Rettinger, *supra* note 4, at 298.

⁶ See *Allen v. State Farm Fire & Cas. Co.*, 59 F. Supp. 2d 1217, 1227 (S.D. Ala. 1999) (“[P]laintiffs should follow the procedure for review required by the

In the landmark case *Gruenberg v. Aetna Insurance Company*⁷ the California Supreme Court extended the tort of bad faith to include first-party insurance coverage disputes.⁸ In the wake of this decision, courts and state legislatures across the country began to recognize the right to file private causes of action against insurers alleging unfair claim settlement practices in first-party insurance coverage disputes.⁹ Three different avenues have been taken by state courts and legislatures in recognizing this right to file a private cause of action: tort action based solely on bad faith; contract action with broad definition of damages; and statutes.¹⁰

Tort Action Based Solely on Bad Faith: Tort action based solely on bad faith relies exclusively on breach of the implied covenant of utmost good faith.¹¹ “Policyholders are not required to allege an independent tort such as fraud or intentional infliction of emotional distress in order to recover under the tort laws.”¹² The general rule of damages in tort is that the injured party may recover for all harm or injuries sustained (including legal expenses, and damages for economic loss and mental distress), regardless of whether these damages could have been anticipated.¹³ Punitive damages may be awarded if the conduct giving rise to liability was particularly egregious.¹⁴

Contract Action with Broad Definition of Damages: A contract action with broad definition of damages involves a good faith and bad faith inquiry confined to the realm of contract¹⁵ where damages are broadly defined to include both general damages (i.e., those following naturally from the breach)¹⁶ and consequential, or incidental, damages (i.e., those reasonably within the contemplation of, or reasonably foreseeable by, the

insurance code and first seek relief from the insurance department and the insurance commissioner.”).

⁷ 510 P.2d 1032 (Cal. 1973).

⁸ *See id.*

⁹ *See* SHARON TENNYSON & WILLIAM J. WARFEL, NAT’L ASS’N OF MUT. INS. COS., ISSUE ANALYSIS: FIRST-PARTY INSURANCE BAD FAITH LIABILITY: LAW, THEORY, AND ECONOMIC CONSEQUENCES 3 (Sept. 2008) *available at* www.namic.org/insbriefs/080926BadFaith.pdf.

¹⁰ *Id.* at 3.

¹¹ *Id.*

¹² *Id.*

¹³ RESTATEMENT (SECOND) OF TORTS §§ 901, 903 (1979).

¹⁴ *Id.* § 908.

¹⁵ *See Beck v. Farmers Ins. Exch.*, 701 P.2d 795, 801-02 (Utah 1985).

¹⁶ BLACK’S LAW DICTIONARY 416 (8th ed. 2004).

parties at the time the contract was made).¹⁷ “Consequential damages may reach beyond the strict contract terms and include prejudgment interest and legal expenses, and damages for economic loss and mental distress.”¹⁸ However, “an independent tort such as fraud or intentional infliction of emotional distress must be alleged in order to make a claim for punitive damages.”¹⁹

Statute: The right to file a private cause of action alleging bad faith is based on statute and judicial recognition of an implied, private cause of action under an Unfair Trade Practices Act that includes an unfair claim settlement practices provision.²⁰ Damages may include prejudgment interest and legal expenses, consequential, or incidental, damages for economic loss and mental distress. With a few notable exceptions,²¹ in states that have adopted a statutory approach to first-party insurance bad faith, punitive damages are not permitted.²² Or, if punitive damages are permitted, a cap is placed on such damages and/or the standard of conduct for awarding such damages is very stringent and thus the exposure to punitive damages is minimal.²³

A majority of states that recognize first-party insurance bad faith allow actions under tort law. The tort of bad faith is a unique application of tort law because it applies despite the existence of a contract, and does not require the policyholder to allege a traditional tort such as fraud or intentional infliction of emotional distress in order to recover punitive damages. Relative to other bad faith liability regimes, the tort of bad faith increases both the potential damages and the uncertainty of judgments for insurance companies. Thus, the legal basis for a first-party insurance bad faith allegation determines the realistic potential for a punitive damages

¹⁷ *Id.* at 417.

¹⁸ Tennyson & Warfel, *supra* note 9, at 3.

¹⁹ *Id.*

²⁰ See CONN. GEN. STAT. ANN. §§ 38a-816, 42-110q (West 1995).

²¹ Pennsylvania statutory law recognizes private actions for first-party insurance bad faith, and authorizes punitive damages. 42 PA. CONS. STAT. ANN. § 8371 (West 2006). The state of Washington recently adopted a bad faith statute that allows punitive damages awards under a standard similar to that in many tort-based regimes. WASH. REV. CODE ANN. § 48.30.010 (West 1998).

²² See generally Mark J. Browne, et. al., *The Effect of Bad Faith Laws on First-Party Insurance Claims Decisions*, 33 J. LEGAL STUDIES 355, 355-390 (2004).

²³ *Id.*

award, dramatically altering the “stakes” of first-party insurance bad faith litigation.

In theory, allowing policyholders to easily recover damages in excess of the insurance benefit owed will provide insurers with added incentives to engage in fair and efficient claims settlement practices. Policyholders are provided assurance that an insurer will not unreasonably withhold payment of a rightful policy benefit, or otherwise engage in conduct that is designed to withhold such payment. However, tort liability may also affect the claim settlement process in ways that are not socially beneficial. For example, high and uncertain penalties for insurer bad faith may create potential gains to policyholders from initiating bad faith actions based on questionable or fraudulent claims. High penalties may also reduce the willingness of insurers to vigorously challenge questionable claims.

This article provides a discussion and analysis of first-party insurance bad faith liability. It traces the evolution of first-party insurance bad faith law, and identifies and discusses the various approaches that have been taken by the courts and state legislatures. The economic rationale for allowing bad faith actions in first-party insurance cases is developed, and the potential gains to insurance market participants are considered. This article also considers potential adverse effects of excessive or uncertain first-party bad faith liability for the insurance claim settlement process, in terms of creating incentives for policyholders to file fraudulent claims and creating disincentives for insurers to investigate potentially fraudulent claims. Automobile insurance claims data are analyzed to investigate the empirical importance of these effects.

II. LEGAL PERSPECTIVES

A. LEGAL DEVELOPMENT OF TORT ACTION BASED SOLELY ON BAD FAITH

Among jurisdictions that permit a tort action based solely on bad faith, a large minority have adopted a “negligence” standard for determining whether an insurer has acted in bad faith;²⁴ the most common

²⁴ At least eleven states have embraced the negligence standard for first-party bad faith claims. These states include Alaska (*State Farm Fire & Cas. Co. v. Nicholson*, 777 P.2d 1152 (Alaska 1989)); California (*Gruenberg v. Aetna Ins. Co.*, 510 P.2d 1032 (Cal. 1973)); Connecticut (*Grand Sheet Metal Prods. Co. v.*

standard among tort jurisdictions is an “intentional tort” standard,²⁵ and one jurisdiction has adopted a “quasi-criminal” standard.²⁶ We discuss each of these standards as they relate to first-party bad faith claims.

1. Negligence Standard

The “negligence” standard was first adopted in a third-party liability insurance case in California.²⁷ Courts following this approach in

Protective Mut. Ins. Co., 375 A.2d 428 (Conn. Super. Ct. 1977)); Hawaii (Best Place, Inc. v. Penn Am. Ins. Co., 920 P.2d 334 (Haw. 1996)); Nevada (Hart v. Prudential Prop. & Cas. Ins. Co., 848 F. Supp. 900 (D. Nev. 1994)); North Dakota (Seifert v. Farmers Union Mut. Ins. Co., 497 N.W.2d 694 (N.D. 1993)); Ohio (Hoskins v. Aetna Life Ins. Co., 452 N.E.2d 1315 (Ohio 1983)); Oklahoma (Christian v. Am. Home Assur. Co., 577 P.2d 899 (Okla. 1977)); South Carolina (Nichols v. State Farm Mut. Auto. Ins. Co., 306 S.E.2d 616 (S.C. 1983)); Texas (Arando v. Ins. Co. of N. Am., 748 S.W.2d 210 (Tex. 1988)); and Washington (Griffin v. Allstate Ins. Co., 29 P.3d 777 (Wash. Ct. App. 2001), *review denied*, 45 P.3d 551 (Wash. 2002)).

²⁵ At least fourteen states have embraced the intentional tort standard for first-party bad faith claims. These states include Alabama (Chavers v. Nat'l Sec. Fire & Cas. Co., 405 So. 2d 1 (Ala. 1981)); Colorado (Herod v. Colo. Farm Bureau Mut. Ins. Co., 928 P.2d 834 (Colo. App. 1996)); Idaho (Robinson v. State Farm Mut. Auto. Ins. Co., 45 P.3d 829 (Idaho 2002)); Indiana (Erie Ins. Co. v. Hickman, 622 N.E.2d 515 (Ind. 1993)); Iowa (Dolan v. Aid Ins. Co., 431 N.W.2d 790 (Iowa 1993)); Kentucky (Wittmer v. Jones, 864 S.W.2d 885 (Ky. 1993)); Mississippi (Universal Life Ins. Co. v. Veasley, 610 So. 2d 290 (Miss. 1992)); Nebraska (Braesch v. Union Ins. Co., 464 N.W.2d 769 (Neb. 1991)); New Mexico (Chavez v. Chenoweth, 553 P.2d 703 (N.M. Ct. App. 1976)); Rhode Island (Zarella v. Minn. Mut. Life Ins. Co., 824 A.2d 1249 (R.I. 2003)); South Dakota (Stene v. State Farm Mut. Auto. Ins. Co., 583 N.W.2d 399 (S.D. 1998)); Vermont (Bushey v. Allstate Ins. Co., 670 A.2d 807 (Vt. 1995)); Wisconsin (Anderson v. Cont'l Ins. Co., 271 N.W.2d 368 (Wis. 1978)); and Wyoming (Hulse v. First Am. Title Co., 33 P.3d 122 (Wyo. 2001)).

²⁶ Aetna Cas. & Sur. v. Broadway Arms Corp., 664 S.W.2d 463 (Ark. 1984). See JEFFREY W. STEMPEL, STEMPEL ON INS. CONTRACTS 10-87 (3d ed. 2009) (1994). According to Stempel, “[t]o appreciate the differences across the states concerning insurer bad faith, one must pay attention to the nuances of precedent and doctrine, which tend to be glossed over in any classification....” *Id.* For this reason, legal scholars often disagree on occasion over whether a particular state has adopted the “negligence” standard or the “intentional tort” standard in first-party cases.

²⁷ See *Comunale v. Traders & Gen. Ins. Co.*, 328 P.2d 198 (Cal. 1958).

third-party cases have reasoned that insurers must be held to a stringent standard because of their disproportionate ability to influence the acceptance or rejection of a settlement offer made by a claimant.²⁸ In particular, the standard demands that an insurer consider the insured's interest in addition to its own in deciding whether to accept or reject the settlement offer.²⁹

Claim-handling practices that are arguably unreasonable can extend beyond third-party claims to include first-party claims; therefore, plaintiffs' attorneys soon asserted that first-party insureds also should be permitted to file a tort action based solely on bad faith.³⁰ Insurers countered that breach of contract should be the exclusive cause of action for first-party insurance bad faith actions because the relationship between an insurer and a policyholder in a first-party context differs from that in a third-party context.³¹ In a first-party context, the relationship might lead to a dispute that could be characterized as "adversarial" (i.e., first-party cases involve disputes over the terms of coverage, whether a loss occurred, or the value of the loss).³² The relationship between an insurer and a policyholder in a third-party context could be characterized as "fiduciary" (i.e., the policy agreement transfers from the insured to the insurer the authority to accept or reject on behalf of the insured a settlement offer presented by a claimant).³³

In the landmark *Gruenberg v. Aetna Insurance Company* decision,³⁴ the California Supreme Court rejected an insurer's argument that third-party cases are different from first-party cases, extending the tort of bad faith to include first-party insurance coverage disputes.³⁵ In *Gruenberg*, the policyholder's business was destroyed in a fire.³⁶ The claim representative informed the fire department investigator that the

²⁸ *See id.*

²⁹ *Id.* at 200-201 (citing *Ivy v. Pac. Auto. Ins. Co.*, 320 P.2d 140 (Cal. 1958)).

³⁰ *See Gruenberg*, 510 P.2d at 1032; Margaret Cronin Fisk, *Looking for a New Cause of Action?*, NAT'L L. J., May 19, 1997, at A1.

³¹ *See* Brief for Am. Ins. Ass'n et al. as Amici Curiae Supporting Appellant, *St. Paul Fire & Marine Ins. Co. v. Onvia, Inc.*, 196 P.3d 664 (Wash. 2008) (No. 80359-5), 2009 WL 907292.

³² Tennyson & Warfel, *supra* note 9, at 4.

³³ *See id.*; *see* STEMPER, *supra* note 26, at 10-38.

³⁴ 510 P.2d 1032 (Cal. 1973).

³⁵ *Id.* at 1038.

³⁶ *Id.* at 1034.

policy limit in place was excessive, suggesting that the policyholder intentionally caused the loss.³⁷ Shortly thereafter, the policyholder was charged with arson.³⁸ Based on the advice of criminal defense counsel, the policyholder initially declined to submit to an examination under oath, which was requested by the insurer shortly after the fire pursuant to the “Your Duties After Loss” provision contained in the policy.³⁹ The charges were subsequently dismissed at a preliminary hearing for lack of probable cause.⁴⁰ Shortly after disposal of the criminal matter, the policyholder informed the insurer that he was now prepared to submit to an examination under oath.⁴¹ The insurer declined to depose the policyholder based on its contention that the coverage was void because the policyholder had previously breached a condition in the policy requiring the insured to submit to an examination under oath at the insurer’s request.⁴²

Arguing that the insurer had unreasonably suggested that he intentionally caused the loss, the policyholder sought both compensatory and punitive damages.⁴³ In adopting the negligence standard in this first-party case, the California Supreme Court reasoned that the third-party context cannot be distinguished from the first-party context.⁴⁴ In third-party claims, the insurer has a “duty to accept reasonable settlements,” whereas in a first-party claim, the insurer has a “duty not to withhold unreasonably payments due under a policy.”⁴⁵ The court observed that “these are merely two different aspects of the same duty.”⁴⁶ When an insurer “[refuses], without proper cause, to compensate its insured for a loss covered by the policy, such conduct may give rise to a cause of action in tort for breach of an implied covenant of good faith and fair dealing.”⁴⁷

³⁷ *Id.*

³⁸ *Id.*

³⁹ *Id.* at 1035.

⁴⁰ See *Gruenberg*, 510 P.2d at 1035.

⁴¹ *Id.*

⁴² *Id.*

⁴³ *Id.*

⁴⁴ *Id.* at 1036.

⁴⁵ *Id.* at 1037.

⁴⁶ *Gruenberg*, 510 P.2d at 1037.

⁴⁷ *Id.* The California court did address the issue pertaining to recovery for mental distress. Given that the policyholder alleged substantial economic losses (e.g., loss of earnings, loss associated with bankruptcy) apart from damages for mental distress, the policyholder was entitled to make a claim for mental distress. Generally, when a policyholder substantially prevails in a first-party claim against

2. Intentional Tort Standard

An “intentional tort” standard was first adopted in *Anderson v. Continental Insurance Co.*, a first-party homeowner’s insurance case in Wisconsin in 1978.⁴⁸ Like the California Supreme Court in *Gruenberg*, the Supreme Court of Wisconsin ruled that the theoretical underpinnings of the bad faith tort in the third-party claim context apply equally in the first-party claim context.⁴⁹ Importantly, however, the Wisconsin Supreme Court departed from the California legal precedent, ruling that “the tort of bad faith is not a tortious breach of contract. It is a separate intentional wrong, which results from a breach of duty imposed as a consequence of the relationship established by contract.”⁵⁰

This dichotomy is the foundation of the intentional tort standard; the denial of a claim may constitute a breach of contract, but it does not constitute bad faith. In other words, an insurer is entitled to contest a claim so long as it has a reasonable basis grounded in law or fact. Whether the insurer ultimately is correct in its position is of no consequence in resolving the bad faith issue. Denying a claim whose validity is “fairly debatable” does not necessarily constitute bad faith, even if the insurer ultimately is incorrect in its position.⁵¹ Rather, the issue is first, whether the insurer undertook a proper investigation, and second, whether the results of the investigation were subjected to a reasonable evaluation and review.⁵² If these conditions are met, the insurer will have established that its denial of the claim was reasonably grounded in law or fact.

Because the intentional tort standard is more stringent than the negligence standard, insurers are more likely to be successful in pretrial pleadings. Judges are more likely to dismiss as a matter of law an allegation of bad faith that involves nothing more than an insurance coverage dispute. Specifically, the Wisconsin court ruled that “there must be a showing of an evil intent deserving of punishment or of something in the nature of special ill-will or wanton disregard of duty or gross or

an insurer, the policyholder is entitled to damages for aggravation and inconvenience. *See, e.g., Hayseeds, Inc. v. State Farm Fire & Cas. Ins. Co.*, 352 S.E.2d 73 (W. Va. 1986).

⁴⁸ *See Anderson v. Cont’l Ins. Co.*, 271 N.W.2d 368 (Wis. 1978).

⁴⁹ *Id.* at 374.

⁵⁰ *Id.*

⁵¹ *Id.* at 375-76.

⁵² *Id.* at 377.

outrageous conduct” in order to recover punitive damages.⁵³ The court added, “[An insurer] must not only intentionally have breached [its] duty of good faith, but in addition must have been guilty of oppression, fraud, or malice”⁵⁴ This heightened standard means that only a small subset of bad faith claims will warrant punitive damages. Direct proof must be presented establishing either that the misconduct was extreme, or that the misconduct was the result of a deliberate, company-wide practice of underpaying claims.

3. Quasi-Criminal Standard

In 1984 the Arkansas Supreme Court overturned a jury verdict that, in the view of the Court, did not require sufficiently stringent standards of conduct and proof to support an award of punitive damages in a first-party bad faith case.⁵⁵ In *Aetna Casualty and Surety v. Broadway Arms*, the policyholder alleged bad faith in the handling of a fire insurance claim.⁵⁶ The evidence presented to the jury to support a finding of intentional oppressive conduct was the claim representative’s statement to the policyholder that he might be asked by the Internal Revenue Service (IRS) to explain why the insurance carrier would pay \$75,000 for loss of inventory when the policyholder’s financial statement showed an inventory valued at only \$23,000.⁵⁷ Apparently convinced that the claim representative had made a thinly veiled threat to report the policyholder to the IRS if the policyholder refused a reduced settlement offer, the jury awarded the policyholder \$5 million in punitive damages.⁵⁸ The judgment was reversed on appeal, and the case was remanded for a new trial based on a “quasi-criminal” standard of conduct.⁵⁹

⁵³ *Id.* at 379.

⁵⁴ *Anderson*, 271 N.W.2d at 379.

⁵⁵ *See Aetna Cas. & Sur. v. Broadway Arms Corp.*, 664 S.W.2d 463 (Ark. 1984). Arkansas is the only state that has embraced the quasi-criminal standard. This standard was upheld in *Columbia Nat’l Ins. Co. v. Freeman*, 64 S.W.3d 720 (Ark. 2002). In this particular case, which involved a property insurer, the court held that the conduct of the insurer must be carried out with a state of mind characterized by hatred, ill will, or a spirit of revenge.

⁵⁶ *Broadway Arms*, 664 S.W.2d at 464.

⁵⁷ *Id.* at 469.

⁵⁸ *Id.* at 465-66.

⁵⁹ *Id.* at 470.

In adopting this standard, the court declared that “evidence of bad faith must be sufficient to show affirmative misconduct of a nature which is malicious, dishonest, or oppressive.”⁶⁰ As articulated by the court, the quasi-criminal standard has three elements. First, the court noted that a single violation of the Arkansas Trade Practices Act does not necessarily constitute bad faith.⁶¹ At minimum, there must be multiple violations in the handling of the claim. Alternatively, a pattern of institutional misconduct (e.g., a company-wide practice of deliberately underpaying claims) would constitute bad faith.⁶² Assuming multiple violations in the handling of the claim, or institutional misconduct, an inference can be made that the evidence is “sufficient to show affirmative misconduct of a nature which is malicious, dishonest, or oppressive.”⁶³ Second, the court ruled that the purpose of the tort of bad faith is not to address the situation where the insurance carrier simply refuses or fails, through nonfeasance, to pay an insurance claim.⁶⁴ In cases of this sort, an adequate remedy already exists under Arkansas law (See Arkansas Stat. Ann. Section 66-3001 et seq. (Repl. 1980)).⁶⁵ Third, the court reasoned that the public interest demands that the tort of bad faith, which includes a substantial punitive damages exposure, be carefully confined to extreme cases of misconduct.⁶⁶ Otherwise, insurers will be inappropriately discouraged from questioning false, suspicious, or inflated claims – a result that will increase insurers’ claim costs and raise policyholders’ premiums. The court suggested that alternative remedies should be used to assure that policyholders are appropriately compensated in those cases where an insurer simply refuses or fails to pay a valid insurance claim.⁶⁷

⁶⁰ *Id.* at 467.

⁶¹ *Id.* at 466.

⁶² *Broadway Arms*, 664 S.W.2d at 466 (citing ARK. STAT. ANN. § 66-3005(9) (current version at ARK. CODE ANN. § 23-66-206(13) (2009))).

⁶³ *Id.* at 467.

⁶⁴ *Id.* at 468.

⁶⁵ ARK. STAT. ANN. §§ 66-3001–3014 (current version at ARK. CODE ANN. §§ 23-66-201–215 (2009)).

⁶⁶ *Broadway Arms*, 664 S.W.2d at 469.

⁶⁷ *Id.* at 468.

B. LEGAL DEVELOPMENT OF CONTRACT ACTION WITH A BROAD DEFINITION OF DAMAGES

The contract standard was embraced in the landmark *Beck v. Farmers Insurance Exchange* decision,⁶⁸ a case in Utah in which the insurer rejected the insured's claim for uninsured motorist benefits without explanation and without conducting an investigation to determine the validity of the insured's claim.⁶⁹ In this particular case, the court reasoned that "[a]lthough the policy limits define the amount for which the insurer may be held responsible in performing the contract, they do not define the amount for which it may be liable upon a breach."⁷⁰ Confinement of the good faith/bad faith inquiry to the realm of contract assures compensation in the situation where the insurer fails to (1) diligently investigate the facts to enable it to determine whether a claim is valid, (2) fairly evaluate the claim, or (3) act promptly and reasonably in rejecting or settling the claim.⁷¹ However, the contract standard forecloses the possibility of a punitive damages award in the absence of proof that an independent tort such as fraud or intentional infliction of emotional distress occurred.⁷² Indeed, in *Beck v. Farmers Insurance Exchange*, the court reasoned that "the practical end of providing a strong incentive for insurers to fulfill their contractual obligations can be accomplished ... through a contract cause of action, without the analytical straining necessitated by the tort approach and with far less potential for unforeseen consequences to the law of contracts."⁷³

⁶⁸ *Beck v. Farmers Ins. Exch.*, 701 P.2d 795 (Utah 1985). At least six states have embraced the contract standard for first-party bad faith claims. These states include Maine (*Marquis v. Farm Family Mut. Ins. Co.*, 628 A.2d 644 (Me. 1993)); Maryland (*Johnson v. Fed. Kemper Ins. Co.*, 536 A.2d 1211 (Md. Ct. Spec. App. 1987)); New York (*Bi-Econ. Mkt., Inc. v. Harleysville Ins. Co.*, 886 N.E.2d 127 (N.Y. 2008)); Oregon (*Nw. Pump & Equip. Co. v. Am. States Ins. Co.*, 925 P.2d 1241 (Or. 1996)) and Virginia (*A & E Supply Co. v. Nationwide Mut. Fire Ins. Co.*, 798 F.2d 669 (4th Cir. 1986) (interpreting Virginia law), *cert. denied*, 479 U.S. 1091 (1987)).

⁶⁹ *Beck*, 701 P.2d at 796-97.

⁷⁰ *Id.* at 801.

⁷¹ *Id.*

⁷² *Id.* at 800-801.

⁷³ *Id.* at 799.

C. LEGAL DEVELOPMENT OF PRIVATE CAUSE OF ACTION BASED ON STATUTE

In a majority of states, a private cause of action is not statutorily or judicially permitted under the state's Unfair Trade Practices Act. In a small number of states, however, either the state legislature has written the law to permit a private cause of action or a court has recognized an implied private cause of action under the law. For example, the Connecticut statute identifies specific types of conduct that constitute bad faith, sets forth the burden of proof, and specifies the damages that can be recovered.⁷⁴ Furthermore, in many states where the courts have failed to recognize a common law cause of action for first-party bad faith, the state legislatures have responded by enacting a statute that permits a private cause of action for the first-party bad faith. Typically, these statutes identify the standard of conduct, the burden of proof, and the damages that can be recovered in a first-party bad faith action.⁷⁵

There is considerable variation among state bad-faith statutes with respect to the standard of conduct, burden of proof, and damages that can be recovered. Some statutes, for example, only allow for limited recovery of damages (e.g., prejudgment interest and attorney fees).⁷⁶ Other statutes contain language that has been broadly construed by at least one court to permit unlimited punitive damages in those cases where the insurer has engaged in more than one listed prohibited practice with respect to the processing of a single claim.⁷⁷

The statutory basis for first party insurance bad faith is still evolving. In recent years, a number of states have enacted new legislation creating or modifying the first-party bad faith liability exposure for insurers. For example, Minnesota passed legislation in 2008⁷⁸ that creates a new private cause of action for first-party insurance bad faith where one previously did not exist. The statute codifies the intentional tort standard, providing for damages if the insured can show (1) the absence of a reasonable basis for denying the benefits of the insurance policy, and (2) that the insurer knew or acted in reckless disregard of the lack of a

⁷⁴ See CONN. GEN. STAT. §§ 38a-816, 42-110q (2009).

⁷⁵ See, e.g., FLA. STAT. ANN. § 624.155(1)(b)(1) (West 2004).

⁷⁶ See, e.g., MICH. COMP. LAWS SERV. § 500.3148(1) (LexisNexis 2001).

⁷⁷ See, e.g., *Maier v. Cont'l Cas. Co.*, 76 F.3d 535 (4th Cir. 1999) (applying West Virginia law).

⁷⁸ MINN. STAT. ANN. § 604.18 (West 2008).

reasonable basis for denying the benefits of the insurance policy.⁷⁹ The law allows policyholders to be awarded up to \$250,000 in “taxable costs” if an insurer is found to be acting in bad faith and up to \$100,000 in attorney’s fees, but specifically precludes punitive damages in the absence of an independent tort such as fraud or intentional infliction of emotional distress.⁸⁰

Recent Colorado legislation lowered the legal standard for asserting a first-party bad faith claim and increased the penalties levied against an insurer, relative to existing common law.⁸¹ The new legislation adopts the negligence standard, whereas the intentional tort standard applies under common law.⁸² In addition, under common law, consequential, or incidental, damages for economic loss and mental distress can be recovered, but the cost of litigation cannot be recovered.⁸³ The new legislation allows for the recovery of the cost of litigation and caps the damages award at two times the policy benefit that was unreasonably denied.⁸⁴

Recent first-party insurance bad-faith legislation in Maryland⁸⁵ applies exclusively to property/casualty insurance policies and allows policyholders to initiate bad-faith claims through the Maryland Insurance Administration (MIA), the state agency responsible for enforcing Maryland’s insurance laws.⁸⁶ The new law adopts the negligence standard and caps damages the insured can recover at the policy limit.⁸⁷ In addition, it provides for recovery of pre-judgment interest and allows recovery of attorney’s fees, but limits the recoverable amount to one third of the actual

⁷⁹ § 604.18(2)-(3).

⁸⁰ § 604.18(3)

⁸¹ COLO. REV. STAT. § 10-3-1113 (2008).

⁸² *Id.*

⁸³ *See* Am. Family. Mut. Ins. Co. v. Allen, 102 P.3d 333, 342 (Colo. 2004) (applying Colorado common law).

⁸⁴ COLO. REV. STAT. §§ 10-3-1116 & 10-16-106.5. Moreover, the new legislation imposes a special penalty on health insurers that unreasonably delay the payment of the policy benefit (i.e., the penalty is 20 percent of the policy benefit, the payment of which was delayed 90 days or longer past the submission of the claim).

⁸⁵ MD. CODE ANN., [INS.] § 27-1001(e)(2)(i) – (ii) (West 2006).

⁸⁶ The law does not apply to claims that fall under the small-claims jurisdiction of district courts or to commercial insurance policies with policy limits exceeding \$1 million.

⁸⁷ §§ 27-303(9) & 27-305(c)(2).

damages.⁸⁸ Previously, an insured could recover only the amount of actual damages, subject to the policy's limits. Such actions could be pursued either through the MIA or as a civil action for breach of contract, but neither avenue allowed recovery of litigation expenses.⁸⁹

Legislation adopted in Washington⁹⁰ expands the definition of first-party insurance bad faith and increases the damages awards available to policyholders in cases alleging insurer bad faith. The remedies specified in the act are separate and distinct from the remedies provided under common law as well as those prescribed in the state's Consumer Protection Act.⁹¹ The new legislation provides for a private cause of action in the event an insurer "unreasonably" denies or delays payment of a policy benefit or commits a specified unfair claims settlement practice, recovery of "actual damages sustained," recovery of the cost of reasonable attorney's fees, and treble actual damages sustained, at the discretion of the trial judge.⁹² This legislation represents a significant departure from most other states' statutory approaches to first-party insurance bad faith, because it permits both unlimited punitive damages and does not contain a stringent standard of conduct for the awarding of such damages.

III. ECONOMIC PERSPECTIVES

A. THE ECONOMIC RATIONALE FOR TORT LIABILITY

Applying economic analysis to first-party bad faith insurance law leads to important insights regarding the purpose of first-party bad faith actions. A key concept is that allowing the courts to impose extra contractual liability on insurers in cases of intentional or unintentional bad faith denial of claims serves the obvious purpose of compensating

⁸⁸ § 27-305(4).

⁸⁹ § 27-305(3)(1)-(2).

⁹⁰ WASH. REV. CODE ANN. § 48.30.010.

⁹¹ § 19.86.090 Washington common law provides for the tort of bad faith with a negligence standard, and the Consumer Protection Act provides for recovery of actual damages sustained, the cost of litigation, and treble damages, subject to a cap of \$10,000 in the event the insurer violates a claims handling regulation.

⁹² *Id.* The specific unfair claims settlement practices covered by the legislation include misrepresentation of policy provisions, failure to acknowledge pertinent communications, failure to meet standards for prompt investigation of claims, and failure to meet standards for prompt, fair, and equitable settlements applicable to all insurers.

policyholders for their unwarranted losses; but it may also serve the broader economic purpose of enhancing the efficiency of insurance contracting. In competitive insurance markets, systematic bad faith in claim settlement practices will reduce demand for insurance from a company that engages in such practices. However, reputation penalties may not be sufficient to guarantee that an insurer will never have an incentive to engage in intentional bad faith claim settlement practices. Nor can reputation penalties guarantee that an insurer will never engage in behaviors that lead to unintentional bad faith denial of a claim settlement.⁹³

An isolated example of intentional bad faith would include, for example, if an insurer strategically denies or delays the settlement of a particularly large insurance claim for the purpose of coercing the policyholder to accept a reduced claim settlement.⁹⁴ Market sanctions alone may not deter this kind of behavior, because the potential cost savings on the claim could outweigh the cost of reputation penalties meted out in the market in the form of reduced demand for insurance. In such cases, the potential for tort litigation creates an incentive for an insurer to avoid unwarranted strategic denial or delay of a claim settlement, by imposing a potentially large financial penalty for such conduct.⁹⁵ Indeed, the mere threat of substantial extra contractual liability will reduce the incentive for an insurer to strategically deny or delay a claim settlement. Most importantly, because a practical mechanism does not exist for an

⁹³ See Joseph M. Belth, *Two Recent Court Decisions Critical of UNUM's Disability Insurance Claims Practices*, THE INS. FORUM, Mar. 2009, at 161. Perhaps the classic case that demonstrates this point concerns disability insurance claim settlement practices implemented by several operating companies of the UNUM Group. These companies implemented a "claims management philosophy" in 1993; previously these companies had a "claims payment philosophy." Various claim settlement practices used to deny, terminate, and settle disability insurance claims resulted in numerous lawsuits alleging bad faith. The volume of lawsuits was such that they resulted in widespread media attention, two far-reaching regulatory probes, and eventual settlements with regulatory authorities under which UNUM agreed to reassess some denied claims and improve claim procedures. While the settlements did not include an admission to the effect that a statute or regulation had been violated, UNUM did agree to pay a \$15 million fine divided among 49 jurisdictions that had signed on to the settlements. *Id.*

⁹⁴ Alan O. Sykes, *"Bad Faith" Breach of Contract by First-Party Insurers*, 25 J. OF LEGAL STUD. 405, 412-13 (1996).

⁹⁵ *Id.* See also KENNETH S. ABRAHAM, DISTRIBUTING RISK: INSURANCE, LEGAL THEORY, AND PUBLIC POLICY 183-85 (1986).

insurer to provide a credible and binding contractual commitment not to deny or delay payment of a legitimate claim, the threat of substantial extra contractual liability for intentional bad faith may reassure policyholders that a valid claim will be paid in a timely manner. This assurance may improve the insurance contracting environment, thereby benefiting both policyholders and insurers.

Because insurance contracts do not contain a binding contractual commitment not to deny payment of a valid insurance claim, and because insurance claims may be complex and policy language cannot fully anticipate all of the details or nuances of a loss, coverage disputes are inevitable.⁹⁶ Some coverage disputes involve differing interpretations of the policy language by the insurer and the policyholder, leading to different conclusions about whether a loss is covered. For example, under a builder's risk policy, the reporting form may instruct the policyholder to deduct the value of land when reporting the estimated completed value of structures. The insurer may interpret "land" to include just the value of the land itself; the policyholder may interpret "land" to include not only the value of the land but also the value of land improvements including paving, gutters, and curbs, for example. Assuming a loss caused by the collapse of paving, gutters or curbs,⁹⁷ the insurer may deny the claim because the policyholder did not report the completed value of these items and thus did not pay a premium for coverage on them. The policyholder may insist that coverage exists for these items based on the instruction contained in the reporting form to the effect that land should be deducted from total estimated completed value of structures, which presumably would include land improvements as well as the value of land itself.

A particularly interesting special case is the potentially fraudulent claim, in which the insurer believes that the policyholder may have "manufactured" false information about the loss event or the amount of loss.⁹⁸ One such case is when an insurer may incorrectly but reasonably believe that the policyholder intentionally caused a loss.⁹⁹ For example, if a summary judgment concerning a mortgage default was issued by a court the day preceding a fire, an insurer may have reasonably believed that the

⁹⁶ Sykes, *supra* note 94, at 429.

⁹⁷ For example, the soil was improperly compacted when the sewer lines were laid.

⁹⁸ Sykes, *supra* note 94, at 425-26.

⁹⁹ *Id.* at 426.

policyholder conspired with a friend to set the fire, notwithstanding the fact that a jury ultimately ruled that the friend acted alone in setting the fire.

With respect to some coverage disputes, further claim investigation can be helpful in enabling the insurer to reach the correct conclusion concerning the existence or non-existence of coverage. With respect to other coverage disputes, however, further claim investigation is not viable because of cost considerations or factual issues that simply cannot be resolved.¹⁰⁰ In these cases, the discovery process attendant to litigation may be the best mechanism for bringing a claim to resolution.¹⁰¹ For example, if an insurer reaches the conclusion that further claim investigation is not viable, the insurer may offer a nuisance settlement even though it believes that coverage does not exist.

In the case of suspected fraud, if we assume that it is more costly for a policyholder to pursue a fraudulent claim than a valid claim, a policyholder with a fraudulent claim will be less likely to persist in the face of a claim denial than a policyholder with a valid claim.¹⁰² This fact implies that costly litigation may be used as a screening device by insurers for the purpose of sorting valid claims from fraudulent ones. For this reason, even though the denial of a claim may lead to litigation, an insurer may find that the denial of some fraction of suspicious claims is efficient.¹⁰³ The benefit of this approach for the insurer is that a policyholder pursuing a fraudulent claim may drop it rather than engage in costly litigation; policyholders with legitimate claims will be more likely to pursue the litigation.¹⁰⁴ While some policyholders needlessly suffer because they are forced to litigate legitimate claims, policyholders as a group benefit from this approach because it minimizes unwarranted claim costs and results in reduced insurance premiums.¹⁰⁵ In specifying a claims denial fraction, the insurer will balance the expected reduction in claims fraud with the expected costs of litigating denied claims.

A similar claims settlement dynamic may occur if we consider the insurer's decision regarding the amount of payment, rather than the decision regarding claim denial. Again, in cases where establishing the complete truth through investigation is not practical, and assuming that the

¹⁰⁰ *Id.*

¹⁰¹ *Id.* at 425-29.

¹⁰² *Id.* at 428.

¹⁰³ *Id.*

¹⁰⁴ Sykes, *supra* note 94, at 428.

¹⁰⁵ *Id.* at 426-27.

“manufacture” of documentation to support an overstated claim is costly for a policyholder, underpayment of suspicious claims may be an optimal fraud-deterrent strategy for an insurer.¹⁰⁶ The benefit of this approach for the insurer is that underpaying suspicious claims reduces policyholders’ incentives to exaggerate the claimed amount.¹⁰⁷ While some policyholders are forced to litigate in order to collect the full amount of a valid claim, the benefit of this approach for policyholders as a group is reduced fraud and reduced insurance premiums. The cost to the insurer is the potential for litigation due to underpayment. Thus, the insurer must balance the amount of claim underpayment against the expected costs of dispute resolution and litigation when specifying a claims payment strategy for suspicious claims.

These perspectives on insurers’ claims payment strategies for responding to suspicious claims provide an economic rationale for permitting first-party private actions for insurer bad faith failure to settle, and for allowing extra contractual damages in these cases. First, the threat of bad faith litigation serves to mitigate insurers’ incentives to strategically deny or delay the payment of valid claims.¹⁰⁸ Second, the threat of bad faith litigation places an appropriate limit on insurers’ use of claim denial or underpayment as a fraud-deterrent strategy.¹⁰⁹ In the absence of potential extra contractual liability, an insurer will consider only the benefits of this strategy to its claims operation, ignoring the costs imposed on policyholders who have legitimate claims that were denied or underpaid. If an insurer faces the possibility of a damage award in excess of the benefit specified in the policy, the insurer is given an incentive to take into account the costs imposed on a policyholder when a legitimate claim is denied, delayed or underpaid. The result is an efficient balance (of costs and benefits) – i.e., between the costs associated with unwarranted claims denial, delay or underpayment,¹¹⁰ and the benefits of reduced insurance claim fraud, and thus the avoidance of excessive costs in the insurance system.¹¹¹

¹⁰⁶ Keith J. Crocker & Sharon Tennyson, *Insurance Fraud and Optimal Claims Settlement Strategies*, 45 J.L. & ECON. 469, 470 (2002).

¹⁰⁷ *Id.* at 504.

¹⁰⁸ *Id.* at 472-73.

¹⁰⁹ *Id.* at 475.

¹¹⁰ Similar arguments can be made regarding insurers’ claim settlement strategies for claims disputes centering on contractual language or factual disputes not involving suspected fraud.

¹¹¹ Crocker & Tennyson, *supra* note 106, at 504-505.

B. POTENTIAL UNINTENDED EFFECTS OF TORT LIABILITY

Although allowing tort actions for the purpose of addressing insurer bad faith in claims settlement may be efficient in theory, practical considerations concerning implementation of tort law have important implications for whether the tort system is in fact efficiency-enhancing. If the standard applied by the court for a finding of insurer bad faith is too lax, and/or if the damages award is too high relative to the actual damages sustained by a policyholder whose claim was denied or underpaid, substantial incentive distortions may arise. If first-party bad faith laws create substantial incentive distortions, the benefits of these laws will be lessened because of increases in insurance costs due to fraud. Because it is the possibility (more specifically, the expected value) of damages from bad faith actions that affects insurer and policyholder decisions, interactions between insurers and policyholders may be distorted irrespective of whether an injured policyholder actually files a suit.¹¹²

A major concern is the increased pressure on insurers to pay reasonably disputable claims.¹¹³ Insurers balance the benefits of reduced fraud costs with the expected costs of litigation.¹¹⁴ If the expected costs of litigation to insurers are sufficiently high that they exceed the expected cost-savings from reduced fraud costs, insurers will have less incentive to employ fraud reduction strategies.¹¹⁵ Specifically, claim investigations may lead to insurer actions that bring accusations of bad faith, and thus an excessive threat of bad faith liability may reduce the number and scope of claim investigations below what they should be.¹¹⁶

This effect on insurer incentives will raise the costs of fraud in both the immediate term because fewer fraudulent claims will be detected, and over the longer term because of reduced fraud deterrence. This latter point is particularly important. A key insight gleaned from economic theory is that the largest savings to an insurer from investigating claims fraud may stem from the deterrence of fraudulent claiming, rather than from cost

¹¹² See Sykes, *supra* note 94.

¹¹³ ABRAHAM, *supra* note 95, at 184.

¹¹⁴ *Id.*

¹¹⁵ *Id.*

¹¹⁶ *Id.*

savings associated with the detection of fraudulent claims.¹¹⁷ By reducing insurer resistance to potentially fraudulent claims and thus increasing the expected payoff to policyholders from filing fraudulent claims, excessive liability for insurer bad faith may increase policyholders' incentives to engage in claims fraud or exaggeration.¹¹⁸

IV. EMPIRICAL EVIDENCE

A. RELATIONSHIP TO PREVIOUS STUDIES

While the impact of first-party bad faith law on insurer claims settlement behavior has not been extensively studied, anecdotal evidence exists from case law to the effect that in some cases tort liability standards have been too lax and/or damages awards have been too high. Based on an examination of a variety of cases involving application of first-party bad faith law, Sykes reached the conclusion that the courts have made substantial errors in applying the law.¹¹⁹ Specifically, in some of these cases, the court found insurer bad faith even though the dispute arose as a result of the insurer's reasonable suspicion of claims fraud;¹²⁰ in other cases, the intentional tort standard was misapplied by the court, with the result being a finding of insurer bad faith even though the claim was reasonably debatable;¹²¹ in still other cases, the size of a punitive damages award appeared to be disproportionately high in comparison to the offense of the insurer.¹²² Of all the cases examined by Sykes, perhaps the most perplexing were those cases in which the court found bad faith based on an insurers' strict reading of the policy language.¹²³ Sykes concluded that "the remedy may be worse than the problem, as the courts seem to find bad faith on the part of insurers who have genuine and reasonable disputes with their

¹¹⁷ Picard provides an excellent discussion of the theoretical literature on insurance fraud. See Pierre Picard, *Economic Analysis of Insurance Fraud*, in HANDBOOK OF INSURANCE 337, 339 (Georges Dionne ed., 2000).

¹¹⁸ See *id.* at 337.

¹¹⁹ Sykes, *supra* note 94, at 407-408.

¹²⁰ See *T.D.S. Inc. v. Shelby Mut. Ins. Co.*, 760 F.2d 1520, 1527-29 (11th Cir. 1985); *Capstick v. Allstate Ins. Co.*, 998 F.2d 810, 815-16 (10th Cir. 1993).

¹²¹ See *Aetna Life Ins. Co. v. Lavioe*, 475 U.S. 813, 823 (1986).

¹²² See *id.* and *Nationwide Mut. Ins. Co. v. Clay*, 525 So. 2d 1339, 1342, 1344 (Ala. 1989).

¹²³ See *Silberg v. Cal. Life Ins. Co.*, 521 P.2d 1103, 1112 (Cal. 1974); *Sparks v. Republic Nat'l Life Ins. Co.*, 647 P.2d 1127, 1137 (Ariz. 1982).

policyholders,”¹²⁴ and that “the ability of the courts to identify opportunistic behavior...is very much in doubt.”¹²⁵

In addition, empirical analysis of insurance claims data reveals that tort liability for first-party insurer bad faith is associated with higher claim payments. Browne, Pryor and Puelz analyze a large dataset of first-party automobile insurance claims settled in 38 different states in 1992.¹²⁶ They find that, even after controlling for a wide array of claim characteristics and for other features of states’ legal and claims environments, claim payments are significantly higher in states that allow tort actions for insurer bad faith compared to states that do not.¹²⁷

A Rand Institute study of the impact of the Royal Globe case (the case that allowed third-party bad faith tort liability claims in California) on insurance claim payments found similar effects.¹²⁸ This study found that when third-party bad faith tort liability claims were allowed, claim payments for automobile bodily injury liability (BIL) claims in California were 25 percent higher than similar claims in other states and that this trend was reversed after the Royal Globe ruling was overturned.¹²⁹ The Rand study also found that the number of BIL claims was higher in California when third-party bad faith tort liability claims were allowed, and this frequency declined when the Royal Globe ruling was overturned.¹³⁰

Of course, higher claim payments or claim filing rates should not be construed negatively if, in the absence of bad faith liability, a tendency

¹²⁴ Sykes, *supra* note 94, at 405.

¹²⁵ *Id.* at 443.

¹²⁶ Mark J. Browne et al., *The Effect of Bad-Faith Laws on First-Party Insurance Claims Decisions*, 33 J. LEGAL STUD. 355, 357 (2004). These authors study uninsured and underinsured motorist claims using data compiled by the Insurance Research Council from a survey of closed claims obtained from insurance companies.

¹²⁷ *Id.* at 386.

¹²⁸ ANGELA HAWKEN ET AL., RAND INST. FOR CIVIL JUSTICE, PUB. NO. MR-1199-ICJ, THE EFFECTS OF THIRD-PARTY, BAD FAITH DOCTRINE ON AUTOMOBILE INSURANCE COSTS AND COMPENSATION 49-53 (2001) (analyzing the impact of *Royal Globe Ins. Co. v. Super. Ct.*, 592 P.2d 329 (Cal. 1979)). In this landmark case, the California Supreme Court held that third-party claimants may sue insurers for violating the Unfair Claim Settlement Practices Act. *Id.* Ultimately, this decision was overruled in *Moradi-Shalal v. Fireman’s Fund Ins. Co.*, 758 P.2d 58 (Cal. 1988).

¹²⁹ HAWKEN, *supra* note 128, at 49-50.

¹³⁰ *Id.* at 49.

exists for insurers to underpay or to wrongfully deny claims. Indeed, in this circumstance, bad faith liability would be performing its intended function. However, if higher payments or claim filings are occurring because insurers are less inclined to investigate potential fraud (that is, unwarranted amounts are paid in order to avoid potential bad faith liability), this phenomenon should be a source of concern for policymakers.

In order to explore the relevance of this concern, we undertake an analysis of the relationship between a state's first-party bad faith regime and the settlement of insurance claims. We focus on the legal regime because this affects the expected value of potential damages faced by an insurer, by affecting both the likelihood of a finding of bad faith and the damages awarded in the event of such a finding. Tort-based bad faith – especially using the California rule – generally leads to higher expected damages for an insurer than contract-based or statutory-based laws, due to the expanded possibility for a punitive damages award. States in which bad faith actions are not permitted, or states in which the law is silent on such actions, will impose lower expected penalties on insurers than other states.

We examine two aspects of insurance claims that may be affected by bad faith liability: the characteristics of claims (specifically the accident, injury and medical treatments), and the claim settlement behavior of insurers (specifically the claim investigations). Consistent with the economic analysis above, we hypothesize that claims will be more likely to exhibit characteristics consistent with (possible) fraud in states where insurers face greater potential liability for bad faith in claims settlement. We further hypothesize that insurers will be less likely to engage in vigorous investigation of claims in those states.

We utilize a large database of paid automobile insurance claims to test these hypotheses. Because the database that we use includes only claims that are closed with some payment by the insurer, a higher prevalence of fraud suspicion indicators and a lower prevalence of insurer investigations among the claims may indicate that insurer claim settlement behavior is different in states that impose liability for bad faith as compared to other states.

B. THE DATA

1. Insurance Claims Data

Individual claim data pertaining to uninsured motorist (UM) claims collected by the Insurance Research Council (IRC) are analyzed.¹³¹ UM coverage is an element of the private passenger automobile insurance policy, and it provides coverage for bodily injury to the policyholder with respect to an accident in which the other driver was at fault, but the other driver was the owner or operator of an “uninsured motor vehicle” (as defined in the policy). Under this scenario, the injured policyholder files a UM claim with his own insurer and receives compensation for both economic and non-economic damages. UM insurance is considered a first-party insurance contract and, consequently, courts in a number of states have specifically upheld the applicability of first-party bad faith remedies in the UM context.¹³²

The data are obtained from a national sampling of claims from a large number of insurance companies. The original dataset includes nearly 6,000 UM claims closed in 1997 (the latest year data are available to us), from accidents occurring throughout the entire United States. Most claims arise from accidents occurring in 1996 or 1997, but accident dates extend back to 1986.¹³³ The closed claim survey provides a wealth of information for each claim, including claim characteristics, insurer investigations of the claim, the claimed amount and the paid amount.

¹³¹ See Insurance Research Council Home Page, <http://www.ircweb.org>. The IRC is an independent, not-for-profit organization founded in 1977, and it is supported by leading property-casualty insurance organizations, including property-casualty insurance carriers and trade associations that represent property-casualty insurance carriers. Its purpose is to provide timely and reliable information that examines public policy issues affecting property-casualty insurers, their customers and the general public. The IRC is devoted solely to research and the communication of its research findings to interested parties; it does not advocate for property-casualty insurers on public policy issues as such.

¹³² Browne, *supra* note 126, at 360-61.

¹³³ Overall, 76.3 percent of accidents in the dataset occur in 1996-1997, 21.9 percent of accidents occur in 1993-1995 and 1.8 percent of accidents in the dataset occur prior to 1993.

2. Claim Characteristics

Studies of automobile insurance fraud have developed a catalog of fraud suspicion indicators or “red flags” that most claim professionals find to indicate potential fraud.¹³⁴ The claim characteristics identified as suspicion indicators encompass a variety of characteristics of the insured, the accident, the injury and the injury treatment. One fraud suspicion indicator is the *lack of a police report* for the accident that produced the claim.¹³⁵ The rationale is that in the normal course of an accident, the police will be called and a report will be filed. If there is no police report, it is more likely that the accident (and hence the injury) is fictitious. Another fraud suspicion indicator is the *lack of a visible injury* at the scene of the accident.¹³⁶ While it is possible that a policyholder could realize his or her injuries only with some delay, if no injury was apparent at the scene of the accident, the likelihood that the injury is fictitious or exaggerated is enhanced.

Soft tissue injuries such as sprains and strains are difficult to medically verify and, therefore, fall into the category of claims that may not lend themselves to discovery through investigation.¹³⁷ As a result, this sort of injury is notorious for being prone to falsification and exaggeration, and a claim involving *only or primarily a sprain injury* is a fraud suspicion indicator for insurers.¹³⁸ Appropriate treatment of sprain injuries is also difficult to determine, providing an additional avenue for a policyholder to falsify the treatment or to exaggerate the amount of treatment. Thus, a *large number of visits to a chiropractor* for treatment of injuries allegedly sustained in an accident is another fraud suspicion indicator or “red flag.”¹³⁹

¹³⁴ See Herbert I. Weisberg & Richard A. Derrig, *Fraud and Automobile Insurance: A Report on Bodily Injury Claims in Massachusetts*, 9 J. INS. REG. 523, 534 (1991); Herbert I. Weisberg & Richard A. Derrig, *Detection de la Fraude: Methodes Quantitatives*, 35 RISQUES 75-99 (1998) (in English translation).

¹³⁵ See, e.g., Weisberg & Derrig, *supra* note 134, at 523.

¹³⁶ *Id.* at 534.

¹³⁷ See Georges Dionne & Pierre St-Michel, *Workers' Compensation and Moral Hazard*, 73 REV. ECON. & STAT. 236, 238-39 (1991).

¹³⁸ See, e.g., Weisberg & Derrig, *supra* note 134, at 534, 537.

¹³⁹ *Id.* at 534, 536.

3. Claim Investigations

Insurers have several methods at their disposal to investigate the validity of medical claims. One method is a medical audit, which entails having a medical professional (usually a nurse) review the medical treatment and bills submitted.¹⁴⁰ This review will provide information from a medical perspective on whether the treatment and billed amounts are appropriate. Another, more costly and detailed, investigative method is an independent medical exam (IME).¹⁴¹ An IME is an examination of the injured policyholder by a medical professional (usually a doctor) chosen by the insurance company. An IME provides a second medical opinion concerning the nature and severity of the injuries to the policyholder. An IME is more expensive than a medical audit and necessitates the cooperation and involvement of the policyholder.

1. *State Bad Faith Liability Regimes*

We combine the data on UM claims with data on each state's legal regime for first-party insurance bad faith to facilitate a comparison of outcomes across states with different bad faith regimes. For each claim, we identify the bad faith regime in effect in the state and year that the accident occurred. After omitting claims for which the bad faith legal regime cannot be determined, our sample for analysis contains 5,338 claims from 48 states and the District of Columbia.¹⁴²

State bad faith laws are compiled from Jeffrey W. Stempel, *Stempel on Ins. Contracts 10-87* (Aspen Publ's 3d ed. 2009) (1994).¹⁴³ The Appendix table displays the bad faith regimes in effect in each state during the sample period for the study. Twenty-four states permitted tort-based bad faith actions during the entire sample period,¹⁴⁴ and 4 states permitted

¹⁴⁰ Weisberg & Derrig, *supra* note 134, at III.1.

¹⁴¹ *Id.*

¹⁴² We omit claims for which the accident state or the accident date is missing. We also omit claims that arise in Pennsylvania, Montana, Puerto Rico and the Virgin Islands because of ambiguities surrounding the treatment of first party bad faith claims in these states and territories.

¹⁴³ STEMPEL, *supra* note 26, at 10-3 to 10-20; BARRY R. OSTRAGER & THOMAS R. NEWMAN, *HANDBOOK ON INSURANCE COVERAGE DISPUTES* 990-1012 (14th ed. 2008) (1988); GENRE & EDWARDS ANGELL PALMER & DODGE LLP, *BAD FAITH LAWS FOR PROPERTY/CASUALTY CLAIMS* (2008).

¹⁴⁴ *Id.*

tort-based bad faith actions for a portion of the sample period.¹⁴⁵ Seven states permitted contract-based bad faith actions during the entire sample period,¹⁴⁶ and 2 states permitted contract-based bad faith actions for a portion of the sample period.¹⁴⁷ First-party bad faith actions were permitted by statute in 5 states during the entire sample period;¹⁴⁸ first-party bad faith actions were not permitted in 4 states during the entire sample period;¹⁴⁹ and first-party bad faith actions were not authorized by either statute or legal precedent in 4 states during the entire sample period.¹⁵⁰

Table 1 displayed below indicates the number of claims in the dataset that were filed under each of the bad faith regimes. The majority of claims in the dataset stem from accidents in states that permit tort actions for first-party bad faith.

Table 1: Distribution of Claims by Bad Faith Regime

Law Regime	Number of Claims	Percent of Claims
No Private Actions Allowed	182	3.4%
No Private Actions Defined	164	3.1%
Contract Law Actions	592	11.1%
Statutory Actions	930	17.4%
Tort Actions, Intentional Standard	832	15.6%
Tort Actions, Negligence Standard	2,638	49.4%
Total Claims	5,338	100.0%
Source: Authors' calculations from IRC survey data.		
Note: Arkansas is included in the intentional tort category.		

¹⁴⁵ *Id.*

¹⁴⁶ *Id.*

¹⁴⁷ *Id.*

¹⁴⁸ *Id.*

¹⁴⁹ STEMPER, *supra* note 26, at 10-3 to 10-20; OSTRAGER & NEWMAN, *supra* note 143, at 990-1012; GENRE & EDWARDS ANGELL PALMER & DODGE LLP, *supra* note 143.

¹⁵⁰ *Id.*

Specifically, 3,470 claims (65.0% of the sample) arise in states that permit tort-based bad faith actions, under either a negligence or an intentional tort standard. A large minority of claims arise in states that permit contract-based bad faith actions (592 claims or 11.1% of the sample) or statute-based bad faith actions (930 claims or 17.4% of the sample). A small minority of claims arise in states that either specifically do not permit bad faith actions (182 claims or 3.4% of the sample) or are silent with respect to whether a bad faith action is permitted (164 claims or 3.1% of the sample).

C. ANALYSIS OF CLAIM CHARACTERISTICS AND CLAIM INVESTIGATIONS

We investigate the effects of first-party bad faith liability on claim characteristics and claim investigations by conducting t-tests of differences in mean values of relevant variables across states with different bad faith regimes. We first compare the states that recognize bad faith actions (through tort, contract or statute) to the states, which do not recognize (either do not permit or have not specifically authorized) bad faith actions. The results of the comparisons are reported in Table 2.

Table 2: Comparison of Bad Faith versus No Bad Faith

	Means		T-Test Statistics
	A	B	A vs B
	States that Allow Bad Faith Actions	States that do not Allow Bad Faith Actions	Bad Faith vs No Bad Faith
Police at the scene	0.833	0.895	-2.98***
Any police report	0.879	0.958	-4.32***
On scene police report	0.809	0.855	-2.14**
No visible injury at scene	0.681	0.556	4.79***
Any sprain injury	0.833	0.769	3.07***
Worst injury is sprain	0.666	0.586	3.04***
Any chiropractor visits	0.355	0.364	-0.34
Number chiropractor visits	25.440	26.300	-0.42
Chiropractor cost/total cost	0.235	0.160	3.22***
Any medical audit	0.367	0.261	3.95***
External medical audit	0.064	0.086	-1.61
Independent medical exam	0.066	0.263	-13.01***

Source: Authors' calculations from IRC survey data.

Note: *** indicates statistically different from zero at the 1 percent confidence level; ** indicates statistically different from zero at the 5 percent confidence level; * indicates statistically different from zero at the 10 percent confidence level. All are two-sided tests. The number of observations with non-missing data differs by variable.

The first two columns of the table report mean values of variables in states that permit bad-faith actions and states that do not, respectively. The third column of the table reports t-statistics and significance levels for comparisons of the means across the two sets of states. The t-statistics provide a test of whether the differences in means across these groups of states are significantly different from zero at the one percent (***) , five percent (**) or ten percent (*) confidence level.

The comparisons reveal significant differences in claim characteristics in states that permit bad faith actions. We observe that police verification of accidents is less prevalent for claims in states that permit bad faith actions; police are less likely to be at the scene of the accident, and accidents are less likely to be the subject of a police report, as compared to other states. Specifically, 80.9 percent of claims paid have a police report from the scene of the accident in states that permit bad faith while this is true for 85.5 percent of claims paid in states that do not permit

bad faith actions. The right hand columns of the table confirm that the differences are statistically significant.

Claims that involve no visible injury at the scene of the accident are more prevalent in bad-faith states (68.1 percent) than in states with no bad faith (55.6 percent). Claims in bad-faith states are more likely to involve a sprain injury (by a margin of 83.3 percent to 76.9 percent). Similarly, sprain injuries are more likely to be the most severe injury experienced in bad-faith states (66.6 percent in bad-faith states compared to 58.6 percent in states with no bad faith). All of these differences are statistically significant, as indicated in the right hand column of the table.

The use of chiropractors by injured policyholders is about the same across the two sets of states, as is the number of chiropractor visits, and the differences are not statistically significant. However, the proportion of the total claimed amount that arises from chiropractor care is larger in bad-faith states (23.5 percent) compared to states with no bad faith (16.0 percent), and this difference is statistically significant.

The data also suggest differences in insurer investments in claim investigation in states that permit bad-faith actions relative to investigations in other states. Insurers faced with potential bad-faith actions are more likely to conduct a medical audit (36.7 percent of claims versus 26.1 percent of claims in states with no bad faith actions), and this difference is statistically significant. However, this result is entirely due to a greater propensity to conduct in-house medical audits. Insurers in bad-faith states are slightly less likely to invest in external medical audits than in other states (although the difference is not statistically significant). This greater use of in-house medical audits may indicate greater investments in claim processing bureaucracy so that a defense can be mounted in the event of a bad-faith lawsuit.

The above interpretation is reinforced by the fact that the potential for a bad faith claim has the opposite effect on insurers' IME use; the proportion of claims for which insurers request an IME is only 6.6 percent in states that permit bad faith actions, while it is 26.3 percent in states with no bad faith. This difference is both large and statistically significant. Because an IME requires the notification and cooperation of the policyholder, insurers may be particularly reluctant to undertake this type of investigation when faced with the potential for bad-faith lawsuits. Such is the case because an IME may lead to an allegation that the insurer unnecessarily engaged in delay tactics with respect to resolution of the claim so as to coerce a reduced claim settlement.

Because there are important differences in legal standards and potential damage awards across different legal regimes for insurance bad

faith, we also separately compare claims filed under tort-based bad faith regimes to those filed under non-tort-based regimes (contract-based or statute-based), and to those filed in states that do not recognize first-party bad faith. Table 3 presents the results of these comparisons.

Table 3: Comparison of Bad Faith Regimes

	Means			T-Test Statistics		
	A	B	C	A vs B	A vs C	B vs C
	States with Tort-based Bad Faith	States with Contract or Statute	States with No Private Actions	Tort vs Contract or Statute	Tort vs No Actions	Contract or Statute vs No Actions
Police at the scene	0.814	0.875	0.895	-5.31***	-3.69***	-1.01
Any police report	0.865	0.911	0.958	-4.53***	-4.84***	-2.82***
On scene police report	0.794	0.843	0.855	-4.12***	-2.73***	-0.56
No visible injury at scene	0.699	0.641	0.556	4.08***	5.48***	2.94***
Any sprain injury	0.846	0.804	0.769	3.61***	3.71***	1.49
Worst injury is sprain	0.690	0.612	0.586	5.43***	3.96***	0.89
Any chiropractor visits	0.359	0.346	0.364	0.87	-0.19	-0.63
Number chiropractor visits	23.390	30.230	26.300	-6.59***	-1.59	-1.56
Chiropractor cost/total cost	0.246	0.211	0.160	2.66***	3.87***	1.89*
Any medical audit	0.395	0.305	0.261	6.06***	4.85***	1.59
External medical audit	0.064	0.063	0.086	0.16	-1.55	-1.54
Independent medical exam	0.043	0.120	0.263	-10.04***	-16.33***	-6.72***

Source: Authors' calculations from IRC survey data.

Note: *** indicates statistically different from zero at the 1 percent confidence level; ** indicates statistically different from zero at the 5 percent confidence level; * indicates statistically different from zero at the 10 percent confidence level. All are two-sided tests. The number of observations with non-missing data differs by variable.

The three left-hand columns of the table report mean values of variables in states that permit tort-based bad-faith actions, states that permit contract-based or statute-based actions, and states that do not recognize bad faith actions, respectively. The three right-hand side columns of the table report t-statistics and significance levels for comparisons of mean values across these sets of states. As in the previous table, the t-statistics provide a test of whether the differences in means across these groups of states are significantly different from zero at the one percent (***), five percent (**) or ten percent (*) confidence level.

The comparisons are striking, and are consistent with the hypothesis that the impact of permitting tort-based bad faith actions is greatest. Claims in tort-based states are notably different from claims in the other two sets of states. These differences are in the hypothesized directions and are statistically significant.

Relative to states that permit contract or statute-based actions, insurance claims in states that permit tort-based bad-faith actions are significantly more likely to include fraud suspicion indicators (i.e. “red flags”), including, less prevalence of police verification of accidents, greater prevalence of claims that involve no visible injury at the scene of the accident, greater prevalence of claims that involve only sprain injuries, and greater proportion of treatment costs from chiropractors. The second of the right hand columns of the table confirms that these differences between tort-based states and contract or statute-based states are statistically significant at better than the 1 percent confidence level. Insurer investigation patterns in states that permit tort-based faith actions also differ from those in states with contract or statute-based actions. Insurers faced with potential tort-based bad faith are more likely to conduct medical audits but are less likely to conduct IME’s than insurers in states with contract or statutory bad faith regimes, and these differences are statistically significant. Insurers in tort-based bad-faith states are no more likely to undertake external medical audits than insurers in states with contract or statute-based bad faith.

The differences between states that allow tort-based bad faith actions and states that allow contract or statute-based actions are generally smaller than those between tort-based states and states that do not recognize bad faith actions. For example, claims that involve no visible injury at the scene of the accident are more prevalent in tort-based states (69.9 percent) than in contract-based or statute-based states (64.1 percent) or states with no bad faith (55.6 percent). Similarly, sprain injuries are more likely to be the most severe injury experienced in tort-based states (69.0 percent in tort-based states, 61.2 percent in contract-based or statute-based states and 58.6 percent in states with no bad faith). The proportion of the total claimed amount that arises from chiropractor care is significantly larger in tort-based states (24.6 percent) compared to contract-based or statute-based states (21.1 percent) or states with no bad faith (16.0 percent). There is also a large difference in the propensity of insurers to use IME’s across the different bad faith regimes. The proportion of claims for which insurers request an IME is only 4.3 percent in states that permit tort-based bad faith actions, while it is 12.0 percent in states that permit contract or statute-based claims and 26.3 percent in states with no bad faith.

Furthermore, there are fewer significant differences between states that permit contract-based or statute-based bad faith and states that do not recognize bad faith actions. These patterns are consistent with the hypothesis that the expected value of potential damages faced by an insurer is an important influence in the claim settlement process, and that the

prospect of punitive damages is of particular importance. In states where insurers face a greater potential for punitive damages in bad faith actions, “red flags” are more prevalent in paid claims and insurers are less likely to employ certain claim investigation techniques.

To further investigate the role of potential punitive damages, we compare UM claims in states that permit tort-based bad faith actions under a “negligence” standard to states that permit bad faith actions but require more stringent standards of proof for punitive damages (i.e., intentional tort, contract law or statutory law). The results of this comparison, displayed in Table 4, yield conclusions that are identical to those obtained from the previous comparisons. This suggests that differences in the standards required for a finding of bad faith, and the attendant penalties arising from such a finding, play a significant role in the claim settlement process.

1. Robustness Checks

One caveat to the previous analysis is that characteristics of automobile accidents may differ across the states. If accidents tend to be less severe in states that permit tort-based bad faith actions, this factor could partially explain the differences in claim characteristics and insurers’ use of investigative techniques. For example, less severe accidents may be more likely to result in sprain injuries (or only sprain injuries). Or, with respect to a small claim, an investigation may be cost-prohibitive even if the claim has “red flags.” Having acknowledged this, observing a larger proportion of small claims (as opposed to large claims) filed and paid in states that permit bad faith actions may itself constitute evidence of incentive distortions induced by bad faith laws. That is, policyholders may file small, illegitimate claims knowing that the insurer will not fight the claim because of the exposure to a bad faith lawsuit. On the other hand, in states that permit contract-based bad faith actions, or no bad faith actions, insurers may “nickel and dime” small claims because they know that an insufficient incentive exists for a policyholder to sue for bad faith. However, these hypotheses cannot be tested because our data do not report accident or claim frequency by state.

Table 4: Comparison of Negligence Standard and Other Standards

	Means		T-Test Statistics
	A	B	A vs B
	States with Negligence- based Tort Actions	States with Other Basis for Bad Faith Actions	Negligence Tort vs Other Bad Faith
Police at the scene	0.779	0.894	-10.93***
Any police report	0.836	0.927	-9.82***
On scene police report	0.757	0.867	-9.93***
No visible injury at scene	0.720	0.638	6.22***
Any sprain injury	0.866	0.796	6.62***
Worst injury is sprain	0.712	0.614	7.36***
Any chiropractor visits	0.397	0.309	6.49***
Number chiropractor visits	23.553	28.183	-4.70***
Chiropractor cost/total cost	0.274	0.191	6.92***
Any medical audit	0.416	0.313	7.55***
External medical audit	0.068	0.058	1.42
Independent medical exam	0.045	0.091	-6.33***

Source: Authors' calculations from IRC survey data.

Note: *** indicates statistically different from zero at the 1 percent confidence level; ** indicates statistically different from zero at the 5 percent confidence level; * indicates statistically different from zero at the 10 percent confidence level. All are two-sided tests. The number of observations with non-missing data differs across variables.

We are nonetheless able to examine whether the differences in claim characteristics and insurer investigations remain statistically significant for claims that are of equivalent size in the different states. We compare claim characteristics and insurer investigations for claims in the third quarter (the quarter in which claim amounts fall above the median claim amount) and the second quarter (the quarter in which claim amounts fall below the median claim amount) of the distribution of claimed amounts in our database, respectively. The benefit of this approach is that it controls for claimed loss amount and focuses attention on the “typical” claim by removing both very small claims and very large claims from the comparisons. One drawback of this approach is that it reduces the sample sizes for our statistical tests.

Table 5 reports the results of the comparison of claims of similar size in states with different bad faith regimes. Because all of our previous

results are consistent across the various alternative methods of categorizing states by bad faith regime, we report only the comparison of states with tort-based bad faith regimes to all other states. The left-hand columns of the table report mean values and t-test statistics for characteristics of claims in the second quarter of the claim size distribution, and the right-hand columns report means and t-test statistics for claims in the third quarter.

Table 5: Comparison of Claims by Quarter

	Claims in 2nd Quarter			Claims in 3rd Quarter		
	Means		T-tests	Means		T-tests
	States with		Tort vs All Other	States with		Tort vs All Other
	Tort-based Bad Faith	All Other States		Tort-based Bad Faith	All Other States	
Police at the scene	0.805	0.880	-3.50***	0.745	0.874	-4.82***
Any police report	0.850	0.932	-3.82***	0.808	0.912	-4.35***
On scene police report	0.783	0.868	-3.36***	0.727	0.84	-4.12***
No visible injury at scene	0.723	0.679	1.50	0.739	0.659	2.74***
Any sprain injury	0.908	0.885	1.20	0.949	0.915	2.14**
Worst injury is sprain	0.758	0.764	0.21	0.797	0.706	3.38***
Any chiropractor visits	0.456	0.399	1.73*	0.616	0.519	3.06***
Number chiropractor visits	15.657	15.539	0.13	27.468	27.349	0.09
Chiropractor cost/total cost	0.356	0.306	1.75*	0.470	0.336	4.80***
Any medical audit	0.445	0.317	4.04***	0.575	0.361	6.67***
External medical audit	0.069	0.067	0.11	0.104	0.084	1.03
Independent medical exam	0.019	0.073	-4.60***	0.043	0.117	-4.57***

Source: Authors' calculations from IRC survey data.
 Note: *** indicates statistically different from zero at the 1 percent confidence level; ** indicates statistically different from zero at the 5 percent confidence level; * indicates statistically different from zero at the 10 percent confidence level. All are two-sided tests.
 The number of observations with non-missing data differs by variable.

The table shows the same general results as the comparisons based on all claims in the database. Indeed, the comparisons of claims in the third quarter follow the exact patterns of signs and statistical significance as those for the full sample of claims. For the smaller claims (those in the second quarter) there are some differences from the full sample. Most notable is that there are no significant differences in injury characteristics (visible injuries, prevalence of sprain) across states with tort-based bad faith and other states. This suggests that the relevant difference in claim characteristics is the greater prevalence of *large-valued* claims that involve sprain injuries in states with tort-based bad faith, consistent with the hypothesis that claim exaggeration may be more prevalent in those states. Overall, it appears from the table that differences in claim size across states with different bad faith regimes is not the determining factor in explaining

differences in claim characteristics and claim investigations, although results for claim characteristics may be somewhat less robust.

A second caveat to the preceding analysis is that results are based on the pooling of claims from different states into categories based solely on first-party insurance bad faith regime.¹⁵¹ One concern is that there may be other important sources of heterogeneity across states; a second concern is the potential for a large state to dominate the comparisons. A specific issue that has been considered in other studies is that the compensation system for automobile accident claims may affect the claim settlement process.¹⁵² To explore whether claim and investigation characteristics differ across bad faith legal regimes when the accident compensation regime is held constant, we modify our sample to exclude claims that were settled under an automobile no-fault compensation system. Additionally, we investigate the influence of a large state on our previous results by excluding claims from California (a state with tort-based bad faith) from the sample. This is done because claims from this state make up a nearly one-quarter of the sample and thus may be influential on the results.

The results of comparisons of claim characteristics and claim investigations in states with tort-based bad faith to all other states, for both of these alternative samples of claims, are displayed in Table 6. The left-hand columns display means and t-test statistics for the sample that omits claims settled under no-fault insurance; the right-hand columns display means and t-test statistics for the sample that omits California.

¹⁵¹ To more fully account for this issue, additional investigations were undertaken using a logistic regression approach that can account for the clustering of claims by state. Like the t-tests, this approach analyzes whether different bad faith regimes are associated with significantly different mean values of each relevant claim characteristic, but takes into account that claims occur within a state and the bad faith regime varies only by state and not for each claim. Thus, clustering by state reduces the likelihood that differences across bad faith regimes are statistically significant, by allowing for the possibility that variable values are correlated across claims within each state. Nonetheless, the results of this analysis remain generally consistent with the t-test results, and are most similar to the t-test results for the sample omitting the state of California. Claim injury characteristics remain statistically significant and higher in tort-based states, and insurer investigations also remain statistically significant and of the same signs as in the t-test analysis. Police presence and police reports and use of chiropractic treatments become statistically insignificant in the logistic regression analysis.

¹⁵² Browne, *supra* note 126, at 360-61; HAWKEN, *supra* note 128.

The results for the sample which omits claims settled under a no-fault regime are extremely similar to results obtained from the full sample of states: the signs and statistical significance of all variables remain the same. This confirms that the automobile accident compensation regime is not the determining factor in explaining differences in claim characteristics and claim investigations.

Table 6: Comparison of Alternative Samples

	Sample omitting Nofault States			Sample omitting California		
	Means		T-tests	Means		T-tests
	States with Tort-based Bad Faith	All Other States	Tort vs All Other	States with Tort-based Bad Faith	All Other States	Tort vs All Other
Police at the scene	0.813	0.868	-4.55***	0.901	0.879	2.01**
Any police report	0.864	0.908	-4.122***	0.942	0.919	2.72***
On scene police report	0.793	0.833	-3.20***	0.879	0.846	3.04***
No visible injury at scene	0.701	0.633	4.54***	0.664	0.625	2.53**
Any sprain injury	0.847	0.799	3.96***	0.802	0.798	0.32
Worst injury is sprain	0.693	0.645	3.15***	0.641	0.607	2.15**
Any chiropractor visits	0.360	0.300	3.94***	0.275	0.35	-5.05***
Number chiropractor visits	23.412	23.272	0.13	22.318	29.636	-5.58***
Chiropractor cost/total cost	0.247	0.192	3.96***	0.174	0.201	-2.08**
Any medical audit	0.396	0.280	7.46***	0.343	0.27	3.02***
External medical audit	0.063	0.054	1.20	0.059	0.067	1.08
Independent medical exam	0.040	0.069	-4.10***	0.048	0.146	-10.31***

Source: Authors' calculations from IRC survey data.

Note: *** indicates statistically different from zero at the 1 percent confidence level; ** indicates statistically different from zero at the 5 percent confidence level; * indicates statistically different from zero at the 10 percent confidence level. All are two-sided tests.

The number of observations with non-missing data differs by variable.

In contrast, results for the sample from which California claims are omitted differ from previous comparisons in several ways. Most notable are the changes in sign for the police report and chiropractor use variables. In the sample without California, police are more likely to be at the accident scene and to submit a report of the accident in states with tort-based bad faith than in other states. The percent of claim costs stemming from chiropractor use is also smaller in tort-based bad faith states than in other states, once California is omitted from the sample. Thus it appears that these claim characteristics may be prevalent in California rather than in states with tort-based bad faith more generally. On the other hand, several of the comparisons without California remain statistically significant and consistent with the findings for the full sample of states. Claims with no visible injury at the accident scene and claims in which the worst injury is a sprain remain significantly more prevalent in states with tort-based bad

faith than in other states. Moreover, the pattern of insurer claim investigation activities remains the same in this sample as in the full sample; insurers in states with tort-based bad faith are more likely to invest in internal routine medical audits and less likely to invest in independent medical examinations. Thus, there remain important differences in claim “red flags” and insurer investigations between states with tort-based bad faith and other states, even after the influence of California is removed from the sample.

V. CONCLUSION

This article has examined first-party insurance bad faith remedies under common law and the recent legislative expansion of such remedies. Theory predicts that allowing policyholders to recover damages over and above the value of the insurance benefit owed will provide insurers with added incentives to engage in fair claim settlement practices, with the result being an enhancement in the efficiency of insurance market contracting. However, theory also predicts that uncertain bad faith standards for insurers and excessive punitive damages awards for policyholders will undermine the benefits of the bad faith remedy by distorting insurers’ claim settlement practices and policyholders’ claim filing incentives, in ways that will lead to more borderline (or even fraudulent and exaggerated) claims and unwarranted increases in insurance premiums.

Previous empirical studies have found that tort-based standards for insurer bad faith are associated with higher insurance claim payments. This article notes that higher claim payments may be evidence of beneficial effects of bad faith liability, if in its absence insurers would underpay claims. A more pertinent concern is whether tort liability for insurer bad faith deters insurers from appropriately challenging potentially fraudulent or otherwise invalid claims, leading to greater amounts of fraud and to unwarranted costs and higher insurance premiums. The article provides new evidence that tort liability for first-party bad faith may reduce insurers’ incentives to monitor for claim fraud, leading to less intensive use of investigative techniques and to more paid claims that contain characteristics often associated with fraud. Although constructing a baseline comparison for determining the appropriateness of claim investigations is difficult, these findings are consistent with the predictions of theory when bad faith liability is uncertain and/or excessive. This raises questions about whether tort liability facilitates efficient claim settlement practices in insurance markets. Additional empirical study of the

relationship between bad faith liability standards and the insurance claim settlement process would be useful.

Appendix Table
State First-Party Bad Faith Regimes for Sample Period of Claims Data

State	First Party Bad Faith Law	State	First Party Bad Faith Law
Alaska	Tort Actions	Mississippi	Tort Actions
Alabama	Tort Actions	Nebraska	Tort Actions
Arkansas	Tort Actions	New Hampshire	Contract Law Actions
Arizona	Tort Actions	New Jersey	Tort Actions (until 1993)
California	Tort Actions		Contract Law Actions (since 1993)
Colorado	Tort Actions	New Mexico	Tort Actions
Connecticut	Tort Actions	Nevada	Tort Actions
District of Columbia	No Private Actions Defined	New York	No Private Actions Allowed
Delaware	Tort Actions (until 1995); Contract Law Actions (since 1995)	North Carolina	Tort Actions
Florida	Statutory Actions	North Dakota	Tort Actions
Georgia	Statutory Actions	Ohio	Tort Actions
Hawaii	No Private Actions Defined	Oklahoma	Tort Actions
Iowa	Tort Actions	Oregon	Contract Law Actions
Idaho	Tort Actions	Rhode Island	Tort Actions
Illinois	Tort Actions	South Carolina	Tort Actions
Indiana	Tort Actions (since 1993)	South Dakota	No Private Actions Defined
Kansas	No Private Actions Allowed	Tennessee	Statutory Actions
Kentucky	Tort Actions	Texas	Tort Actions
Louisiana	Statutory Actions	Utah	Contract Law Actions
Massachusetts	Statutory Actions	Virginia	Contract Law Actions
Maryland	Contract Law Actions	Vermont	Tort Actions
Maine	Contract Law Actions	Washington	Tort Actions (since 1992)
Michigan	No Private Actions Allowed	Wisconsin	Tort Actions
Minnesota	No Private Actions Allowed	West Virginia	Contract Law Actions
Missouri	No Private Actions Defined	Wyoming	Tort Actions (since 1990)

Source: Authors' calculations from GenRe (2008), Stempel (2006) and Ostrager and Newman (2008).

Note: The sample period includes years 1986 through 1997.

REGULATION OF LARGE FINANCIAL INSTITUTIONS: LESSONS FROM CORPORATE FINANCE THEORY

John P. Harding*

Stephen L. Ross**

This article applies a model of firm capital structure to the current financial crisis and summarizes the insights the model provides regarding the regulation of large financial institutions in a post-crisis world. Firm capital structure is evaluated by studying how firms finance their activities using debt and equity, which in turn captures an important element of firm risk taking. First, the simple model is briefly summarized. Second, the model's results are used in order to interpret the evolution of the current financial crisis and to put it into perspective. Finally, the article presents forward-looking observations and suggestions for future regulation. The article concludes that an essential element of a new regulatory framework must include an effective method for dealing with the extension of the "Too Big to Fail" umbrella, which has extended the risk of moral hazard beyond depository institutions. It asserts that a successful new framework must include stringent capital standards for financial institutions combined with regulators that have the authority and commitment to enforce those standards putting owners and managers at risk when capital standards are violated, even during financial crises when there are strong incentives for regulator accommodation to preserve asset value. The new framework must be flexible in order to adapt to changing financial conditions especially developments that affect franchise value, and contain provisions that expose uninsured debtors to risk when capital standards are violated so that debt holders have an incentive to monitor the activities of very large financial firms.

* John P. Harding, Professor, Finance Department, University of Connecticut

** Stephen L. Ross, Professor, Economics Department, University of Connecticut. The authors benefited from conversations with Dwight Jaffe, Ed Kane, Pat McCoy and seminar participants at the Federal Reserve Bank of Boston.

I. INTRODUCTION

Equity capital is the shock absorber for our financial system and the current financial crisis, like a bumpy road for an auto designer, provides a unique opportunity for financial regulators to evaluate the predictions of theory and improve the design of the regulatory system. The purpose of this paper is to apply a simple model of firm capital structure to the current situation and summarize the insights it provides regarding the regulation of large financial institutions in a post-crisis world. The paper begins with a brief summary of the model and uses the results of that model to place the evolution of the current crisis into perspective. The paper concludes with forward-looking observations and suggestions for future regulation.

II. A SIMPLE MODEL OF FIRM CAPITAL STRUCTURE

The study of firm capital structure is basically the analysis of how firms finance their economic activities – in particular, the allocation of funding between equity and debt or firm leverage (the ratio of debt to equity). In much of the literature, the fundamental risk of the firm's assets is taken as given by economic forces outside the control of the firm and the choice of leverage can be viewed as allocating that risk between different providers of capital. In general, the use of debt financing results in creating a lower risk/lower return investment opportunity for debt providers and a higher risk/higher return investment for equity providers.

Leland develops a simple model of the choice of firm capital structure that has proven to provide valuable insights into important questions.¹ In Leland's model, a private firm in a given industry must fund its initial investment in productive assets with a combination of debt and equity. The return on the firm's assets evolves stochastically, but the firm has limited ability to adjust its capital structure in response to the actual asset outcomes.² As a result, the firm's managers must choose their initial leverage to maximize the expected present value of the firm's future operations in the presence of taxes, bankruptcy costs and tax-deductible

¹ Hayne E. Leland, *Corporate Debt Value, Bond Covenants, and Optimal Capital Structure*, 49 J. OF FIN. 1213 (1994).

² *Id.* at 1248. While such a restriction may seem unrealistic, the recent crisis has demonstrated that firms experiencing financial difficulties often have no access to additional equity capital at the time of distress. This fact places more importance on the forward-looking choice of initial leverage. *Id.*

interest payments on the debt. In this model, firms rationally choose to fund themselves with some debt and some equity as they trade-off expected bankruptcy costs that could be avoided if the firm chose all equity and the tax advantage of debt, which is maximized by choosing all debt.

A key lesson of this paper arises from the author's consideration of the firm's choice of a bankruptcy threshold – the level of asset value at which the firm's owners voluntarily turn over the assets to the debt holders for liquidation. In financial terms, the firm's owners have a call option on the assets with a strike price equal to the debt payoff. Because an option can never have a negative value, ignoring the periodic debt service payment, firm equity could never be negative and the owners would never choose bankruptcy.³ However, to keep their option alive, the owners must make the periodic debt service payments and thus choose bankruptcy when the cost of the required debt service payment exceeds the current call option value. Two important insights from Leland's analysis are that for reasonable parameter assumptions owners will choose to continue to operate the firm (maintain their option value) until the market value of assets falls well below the face value of debt and the key factor that limits this behavior of owners is the required debt service to bond holders.

Because equity holders can choose the best time (from their perspective) to turn over the firm to debt holders, the bond holders are exposed to greater risk by the continued operation of an insolvent firm. Once declines in asset values have made a firm insolvent, the equity holder's decision whether to continue operating the firm only depends upon the subset of outcomes where asset values recover sufficiently for the firm value to exceed debt obligations. The return to equity holders is zero in all other scenarios, no matter how bad the outcome. However, debt holders care about the expected future value of assets over all states of the world, but are especially concerned about situations where asset values continue to erode because they will be the owner of the assets in those states of the world. Furthermore, insolvency is likely to affect the equity holders' choice of investment risk because, from their perspective, they still benefit from exceptionally good outcomes and do not bear increased losses if the

³ A call option provides the holder the right, but not the obligation, to purchase a security or asset at a fixed price. Such an option always has a positive value regardless of the current value of the asset as long as there is some chance that the value of the asset may rise above the strike price. For a complete introduction to options and their valuation see JOHN C. HULL, *OPTIONS, FUTURES, AND OTHER DERIVATIVES* (Prentice-Hall) (2008).

bad outcomes become worse. Such increased investment risk exposes debt holders to substantially more risk than the likely investment decisions of a solvent firm.⁴

The providers of debt capital can anticipate this behavior on the part of owners and demand fair compensation for bearing the risk of significant losses from asset declines by increasing the required coupon rate on the debt. However, this creates “risky” debt and limits the ability to allocate risk that was one of the original motivations for leverage.⁵ A common solution to this problem is for equity and debt providers to negotiate bond covenants whereby owners contract to transfer control of the firm to debt holders at a higher threshold value of assets. For example, a positive net worth covenant would transfer control when the market value of assets first falls below the face value of debt.⁶

A final important result from Leland’s model is that whether or not the debt is protected by bond covenants, the optimal choice of initial leverage provides a substantial equity “cushion” above the bankruptcy threshold. Leverage levels near the bankruptcy threshold raise the risk of incurring the transactions costs associated with bankruptcy. Since debt holders will bear those costs in bankruptcy, firm owners must pay for those costs with higher interest rates and so have an incentive to hold additional equity in order to lower the interest rate on debt.

⁴ This behavior arises from the limited liability exposure of equity holders, so that owners of an insolvent firm face no downside risk and benefit from large risks that have a low probability of success, but, if successful, create the potential for future positive net worth of the firm. FRANK H. EASTERBROOK & DANIEL R. FISCHEL, *THE ECONOMIC STRUCTURE OF CORPORATE LAW* 49-50 (1996).

⁵ To see the intuition underlying this claim, envision a continuum of risk measured along a straight line where the firm’s assets are located in the middle of the line. Adding leverage to the firm creates one security with less risk (debt) and another with more risk (leveraged equity) – splitting the single center point into two, one to the safer side of center and one to the riskier side. To the extent that more of the risk of future outcomes falls on the debt holders, the two new points must remain closer to the original center point.

⁶ There are many possible variations of this basic idea. For example, some covenants might limit the managers’ ability to finance new growth or sell assets in lieu of transferring total control to debt holders. Another variation might provide debt holders representation on the board when certain financial thresholds are crossed. Clifford W. Smith & Jerold B. Warner, *On Financial Contracting: An Analysis of Bond Covenants*, 7 *J. OF FIN. ECON.* 117 (1979).

Regulated depository institutions have two significant differences from the firms Leland studies: they are able to issue debt at the riskless rate without regard to their individual financial condition or the selected leverage, and they are subject to binding capital regulation. These capital regulations can be viewed in Leland's terms as an exogenously determined bankruptcy threshold at which level the debt holders are paid in full and the assets turned over to the regulator for liquidation. As a result of the reduced market discipline via debt cost, many scholars have discussed the concern that deposit insurance creates a moral hazard that encourages owners of insured depositories to select very high levels of leverage, i.e., hold the minimum capital allowed by the capital regulations.⁷ On the other hand, Buser, Chen and Kane (1981) point out that banks incur significant costs that are not explicitly priced attributable to regulations, investment restrictions and monitoring.⁸ Merton (1978) and Markus (1984) argue that these costs increase when banks have weak capital positions creating an equity favoring factor that offsets the incentives created by deposit insurance and could lead to banks choosing to hold high levels of equity and so provide very limited deposit services.⁹

Ideally, regulation should preserve incentives for owners to operate as deposit issuing banks, while selecting a prudent level of leverage tailored to the riskiness of their asset strategy – voluntarily holding capital in excess of the specified minimum. It is unreasonable and impractical to expect regulators to be able to establish capital standards tailored to an individual bank's operating strategy and asset risk. Similarly, regulators cannot monitor the portfolios of all financial institutions closely enough to accurately classify banks by the riskiness of their assets. Consequently, it is important that the capital regulation framework provides owners and managers sufficient incentives to provide a capital cushion that reflects the

⁷ See Jean-Pierre Gueyie & Van Son Lai, *Bank Moral Hazard and the Introduction of Official Deposit Insurance in Canada*, 12 INT'L REV. OF ECON. & FIN. 247 (2003); Michael C. Keeley, *Deposit Insurance, Risk, and Market Power in Banking*, 80 THE AM. ECON. REV. 1183 (1990); and David A. Marshall & Edward S. Prescott, *Bank Capital Regulation With and Without State-Contingent Penalties* (Federal Reserve Bank of Chicago, Working Paper No. 2000-10, 2000).

⁸ Stephen A. Buser, Andrew H. Chen & Edward J. Kane, *Federal Deposit Insurance, Regulatory Policy and Optimal Bank Capital*, 36 J. OF FIN. 51, 56 (1981).

⁹ Alan J. Marcus, *Deregulation and Bank Financial Policy*, 8 J. OF BANKING AND FIN. 51, 59 (1984); Robert C. Merton, *On the Cost of Deposit Insurance When There are Surveillance Costs*, 51 J. OF BUS. 439, 448 (1984).

operating strategy of the bank and appropriately reduces the risk of violating the standards in adverse times.

Harding, Liang, and Ross (2008) (HLR) extend Leland's model to incorporate the additional factors influencing bank capital choices.¹⁰ They find that capital regulation alone is not effective for establishing an incentive for banks to hold capital reserves in excess of the minimum capital standards. Bankruptcy costs and insurance benefits both vary with capital structure in such a way that banks would rationally choose to operate with the minimum capital allowed or forego being a bank entirely (i.e., choose all equity financing), depending upon the strictness of capital regulation. A bank will only choose to hold capital reserves in excess of minimum capital requirements if there is some additional firm franchise value that is placed at risk by the capital regulation.

HLR find that tax-advantaged debt as parameterized by Leland (1994), or any franchise value that depends upon the total amount of an institution's deposits, creates such a franchise value and, when this franchise value is considered, banks will hold excess capital without shifting to the extreme of foregoing issuing deposits. Counter-intuitively, the tax benefits of debt or the franchise benefits of the deposit base lead to lower leverage which is quite striking given that tax-advantaged debt traditionally is viewed as increasing the incentive for firms to take on more debt. The result is explained by the fact that a firm's franchise value is placed at risk by capital standards causing firms to choose lower levels of leverage in order to protect this value for equity holders. Overall, the combination of capital standards with the power of the regulator to wipe out equity when the standard is violated acts like the bond covenants discussed in Leland encouraging managers to operate with a capital cushion above the threshold that would trigger takeover.

HLR demonstrate that it is the regulator's ability to move swiftly to place a firm that violates capital standards into receivership that leads firms to hold excess capital. Alternative structures such as warning thresholds that trigger additional regulatory burdens have little effect on leverage. It is the ability to liquidate the firm, placing the franchise value at risk, that leads to more conservative choices of leverage. Consistent with these predictions, the authors note that commercial banks, that are subject to dissolution if capital standards are violated, hold a substantial capital

¹⁰ John P. Harding, Xiaozhong Liang & Stephen L. Ross, *Bank Capital Requirements and Capital Structure*, 4-6 (Univ. of Conn., Dept. of Econ, Working Paper No. 2009-09, 2008).

cushion above the minimum capital requirement while the mortgage GSEs, Freddie Mac and Fannie Mae, who prior to the summer of 2008 were free from the risk of receivership, held little or no capital cushion through the same period.¹¹

III. UNDERSTANDING THE CURRENT CRISIS

The origins of the current financial crisis are complex and no single factor can be singled out as the primary cause. However, most observers believe that increasing use of leverage, broadly defined, was a contributing factor. The U.S. Government Accountability Office (GAO) report to Congress on the origins of the crisis shows that total debt in the U.S. economy rose significantly in the years preceding the crisis. Measured as a ratio to nominal GDP, total debt rose from roughly two times GDP in 2002 to 2.25 times GDP in 2007.¹²

¹¹ Fannie Mae and Freddie Mac first began operating under the statutory minimum capital requirement in 1993, and in that year they held excess capital of roughly \$1 billion and \$0.7 billion, respectively. These amounts expressed as a percentage of assets plus Mortgage Backed Securities (MBS) outstanding were .14% and .13%, respectively. In most of the eight subsequent years, Fannie Mae held excess capital well under \$1 billion and in 1998 and 1999 its excess capital was 1/100th of a percent of the assets and MBS. Freddie Mac's excess capital, while slightly higher when measured as a percentage of assets and MBS, was smaller when measured in dollars and is also consistent with the claim that the firm intended to meet, but not exceed its capital standard. On average over the period from 1993 through 2001, Fannie Mae and Freddie Mac held less than 1/10th of a percent of excess capital. The period from 2002 through 2007 is distorted by the effects of financial restatements arising from accounting problems experienced in the period from 2003 through 2005. However, the numbers for 2007 most likely reflect the firms' contemporaneous intentions, and they still suggest that the firms were not holding precautionary excess capital. See FHFA AND OFHEO ANNUAL REPORTS TO CONGRESS 1994 - 2008 available at <http://www.fhfa.gov/Default.aspx?Page=240>. From 2001 through 2008, commercial banks held total capital (Tier I plus Tier II) of 12.3% of risk-based assets. During this period, a bank was deemed to be well-capitalized with a total capital ratio of 10% of risk-based assets and adequately capitalized with a ratio of 8%. Data on commercial bank capital ratios is available at <http://www2.fdic.gov/SDI>.

¹² See, e.g., U.S. GOV'T. ACCOUNTABILITY OFFICE, FINANCIAL MARKETS REGULATION: FINANCIAL CRISIS HIGHLIGHTS NEED TO IMPROVE OVERSIGHT OF LEVERAGE AT FINANCIAL INSTITUTIONS AND ACROSS SYSTEM, GAO-09-739, p.13,

By their nature, financial institutions have always been some of the most highly leveraged firms. The five largest U.S. investment banks together had an average leverage ratio of about 30 to 1 during the Asian Financial Crisis in 1997. While this ratio declined to 22 to 1 in the period following 1997, it had risen back to 30 to 1 by 2007. The five largest bank holding companies had an average leverage ratio of about 13 to 1 throughout the same period.¹³

However, the effective growth of leverage at financial institutions in recent years is difficult to measure precisely because recent developments in financial assets and derivatives allow institutions numerous opportunities to effectively leverage their risk “off balance sheet” while still maintaining excess capital as measured by traditional measures of on-balance sheet risks. For example, before the development of credit default swaps, an institution willing to bear the credit risk of an industrial firm would make a loan or purchase the debt of that company and report the loan as an asset, thereby increasing its required capital. In recent years, institutions could take on that same credit risk without making a loan by writing an insurance policy through a credit default swap.

Leverage in the economy also increased with greater use of hedge funds, Special Purpose Entities (SPE) as part of holding company structures and the ability to issue structured debt securities. One sector of the economy that made prominent use of structured debt to expand credit was the housing sector where subprime mortgages were used to expand the population of mortgagors. The resulting pools of subprime mortgages were structured in ways to attract investment from a wide array of non-traditional investors. While most large investment banks and bank holding company subsidiaries that originated subprime mortgages operated with the intent to pool and sell mortgage-backed securities as soon as a sufficient number of loans had been originated, at any given time, they nevertheless had significant exposure to subprime loans because they were holding mortgages as inventory awaiting future sales or holding securities as part of their underwriting and trading operations. In some cases these institutions

Fig.3 (2009) (hereinafter GAO REPORT) *available at* www.gao.gov/new.items/d09739.pdf.

¹³ *Id.* at 19. At the time, the five largest investment banks or broker-dealer holding companies were Bear Stearns, Goldman Sachs, Lehman Brothers, Merrill Lynch, and Morgan Stanley. The five largest bank holding companies were Bank of America, Citigroup, JPMorgan Chase, Wachovia, and Wells Fargo. *Id.* at 18-19 figs.4 & 5.

also bought mortgage related securities as investments through special purpose entities or subsidiaries.¹⁴

The role of SPE's is especially instructive. If a regulated financial institution committed to provide contingent funding for a SPE, such as liquidity facilities or credit enhancements, the institution would have been required to hold capital against that commitment. In many cases, however, the institutions establishing the SPE's provided no such guarantees and so the business activities of the SPE were not considered in calculating capital requirements. Nevertheless, these institutions still faced both reputation and legal risks associated with the possibility of an SPE failure. Many of these SPE's increased returns by investing in long-term assets like mortgage-backed securities and financed these investments with short term commercial paper. As the financial crisis began in 2007, many of these SPE's could not renew their debt financing and the regulated financial institutions either extended financing themselves or directly brought the SPE's onto their books in order to avoid the reputation damage associated with SPE failure.¹⁵

The first signs of the current crisis arose in the subprime mortgage market where delinquencies and defaults began to increase in the first half of 2007. In June, 2007, Standard and Poor's and Moody's began to downgrade structured debt backed by subprime mortgage securities. This, in turn, led Bear Stearns to suspend redemptions in certain subprime investment funds it was managing, and later in July to liquidate two hedge funds that invested in subprime mortgage-backed securities. By March 2008, Bear Stearns was taken over by JPMorgan Chase in a deal brokered by the U.S. government.¹⁶ Throughout this period, subprime mortgage securities market values fell sharply as uncertainty increased about the ability of the structures to withstand higher levels of delinquency and default.¹⁷ These concerns were exacerbated when home prices began to fall in several major markets.¹⁸

¹⁴ *Id.* at 19-20.

¹⁵ *Id.* at 56-58.

¹⁶ For a detailed timeline of events, see The Federal Reserve Bank of St. Louis, *The Financial Crisis: A Timeline of Events and Policy Actions*, <http://timeline.stlouisfed.org/index.cfm?p=timeline>.

¹⁷ See, e.g., Roddy Boyd, *How the Bear Stearns Deal Got Done*, CNN MONEY.COM, Mar. 17, 2008, http://money.cnn.com/2008/03/17/news/companies/boyd_bear.fortune/index.htm?postversion=2008031717. See also, *Economic Outlook Before the S. Comm. On Banking, Housing, and Urban Affairs*, 110th Cong. (2008) (statement of Federal Reserve Chairman Ben S. Bernanke)

While much of the public attention was focused on “toxic assets”, the combination of declining asset values with high leverage and increasing uncertainty and risk aversion resulted in a rapid spread of the crisis throughout the global financial system. Asset value declines forced institutions to deleverage initially by raising capital,¹⁹ but as raising capital became more difficult institutions were forced to deleverage by selling assets into markets dominated by sellers.²⁰ Brunnermeier and Kashyap, Rajan, and Stein²¹ emphasize the role of deleveraging in asset price declines suggesting that deleveraging could trigger downward spirals as asset sales depress asset prices requiring further deleveraging. Further, Mark-to-Market²² rules could exacerbate a deleveraging based price spiral as financial institutions are forced to re-value assets in the face of rapid declines in asset prices and then due to Mark-to-Market rules must respond to those lower values with further deleveraging. However, it is important to recognize that the falling prices of assets are not simply the result of

available at http://banking.senate.gov/public/index.cfm?FuseAction=Files.View&FileStore_id=5a34e5df-2de7-4c9b-a5a7-4f0893acd79e.

¹⁸ See, e.g., Lee Christie, *Home Prices See Another Record Plunge*, CNN MONEY.COM, Oct. 28, 2008, http://money.cnn.com/2008/10/28/real_estate/August_Case_Shiller/index.htm. See also Lee Christie, *Home Prices in Record Decline*, CNN MONEY.COM, Nov. 25, 2008, http://money.cnn.com/2008/11/25/real_estate/third_quarter_case_shiller/index.htm.

¹⁹ GAO REPORT, *supra* note 13, at 20.

²⁰ If a firm is levered 30:1 and experiences a 10% decline in value on 10% of its assets, its equity base declines by 30% and its leverage increases to more than 40:1. In order to return leverage to the original ratio of 30:1, it must sell close to one third of its assets.

²¹ Markus K. Brunnermeier, *Deciphering the 2007-08 Liquidity and Credit Crunch*, 1 J. OF ECON. PERSPECTIVES 23, 77-100 (2009); Anil K., Kashyap, Raghuram G. Rajan, & Jeremy C. Stein, *Rethinking Capital Regulation*, Jackson Hole Symposium: “Maintaining Stability in a Changing Financial System” (2008), available at <http://kansascityfed.org/publicat/sympos/2008/KashyapRajanStein.03.12.09.pdf>.

²² Mark-to-Market is a way to measure assets and liabilities that appear on a company’s balance sheet and income statement that involves an attempt to measure companies’ assets and liabilities at fair or market value. For more detailed information, see SEC. AND EXCH. COMM’N, OFFICE OF CHIEF ACCOUNTANT & DIV. OF CORP. FIN., REPORT AND RECOMMENDATIONS PURSUANT TO SECTION 133 OF THE EMERGENCY ECONOMIC STABILIZATION ACT OF 2008: STUDY ON MARK-TO-MARKET ACCOUNTING (2008) available at www.sec.gov/news/studies/2008/marktomarket123008.pdf.

psychology and forced sales, but also reflect declines in value due to greater economic uncertainty in general and/or a higher likelihood of extreme or tail events in the fundamental markets on which those assets draw their value. Further, even without a deleveraging price spiral, real declines in complex asset values such as subprime mortgage-backed securities could be exacerbated by a lemons problem where firms have private information on the quality of securities and sell the lowest quality assets as they deleverage.²³

The initial regulatory reaction focused on the immediate symptom – the freezing up of markets – by providing liquidity in the hope of stabilizing the markets by stimulating buyers of assets. In August 2007, the Federal Reserve publicly emphasized its intention to provide reserves as necessary to meet the needs of depository institutions,²⁴ and then lowered the federal funds rate throughout the fall of 2007. In December 2007, the Federal Reserve Board announced the creation of a Term Auction Facility (TAF) that would auction fixed amounts of term funds to depository institutions allowing those institutions to use a wide variety of assets including mortgage-backed securities as collateral. This action was followed by related efforts throughout the spring and summer of 2008 including legislative authority to extend credit to the Government Sponsored Enterprises (GSEs) Fannie Mae and Freddie Mac.²⁵ These efforts proved to be ineffective in containing the crisis and repeated waves

²³ See GAO REPORT, *supra* note 13, at 15-23 for a general discussion of the deleveraging of financial institutions during this period.

²⁴ The Federal Reserve is often viewed as limiting the use of its discount window to meet reserve requirements by requiring detailed explanation for substantial requests. ROBERT E. HALL & JOHN B. TAYLOR, *MACROECONOMICS: THEORY, PERFORMANCE, AND POLICY*, 327 (2nd ed., 1988). Banks usually meet this requirement via inter-bank lending. See OECD Glossary of Statistical Terms, <http://stats.oecd.org/glossary/detail.asp?ID=1385>. Accordingly, this statement addressed the increasing illiquidity in inter-bank lending markets as indicated by an increasing London Inter-Bank Offer Rate (LIBOR). Asani Sarkar, *Liquidity Risk, Credit Risk, and the Federal Reserve's Response to the Crisis*. Federal Reserve Bank of New York Staff Report, available at http://www.newyorkfed.org/research/staff_reports/sr389.pdf.

²⁵ Credit was not granted prior to the GSEs being placed in receivership in September. However, the recovery act passed in the summer of 2008 did authorize the U.S. Treasury Department to extend credit to the GSEs if necessary and was intended to improve the GSEs borrowing ability by increasing investor confidence. See Timeline, *supra* note 17.

of deleveraging asset sales simply led to further asset value declines and the need for additional deleveraging. By the end of the summer of 2008, U.S. regulators shifted focus from providing liquidity and shoring up specific markets to working to preserve the solvency of financial institutions through the infusion of capital. Most notably, this shift was signaled by the U.S. Treasury Department proposed legislation to purchase “troubled assets” of financial institutions that was eventually passed by congress in October as the Troubled Asset Relief Program (TARP). Soon after passage, the plan to purchase “troubled assets” was abandoned and replaced by efforts to directly infuse capital into financial institutions through the purchase of preferred shares of stock.²⁶

Clearly leverage was a key element of the financial crisis and understanding the drivers that led to the sharp increase in leverage is critical for planning the future regulatory regime.²⁷ HLR’s model of bank capital regulation provides useful insights. First, consider entities such as investment banks, the two large mortgage GSEs and other non-depository financial institutions. Most of the larger firms in this category had significant benefits from relationships with the federal government and arguably were protected by the “Too Big to Fail” principle. Such firms differ significantly from the purely private firms studied by Leland in that they benefit from an implicit government guarantee of liabilities due to the risk their failure would pose to the national economy. This implicit guarantee meant that these institutions were similar in some respects to depository institutions in that they could issue debt at lower cost and with less market scrutiny than private firms. Significantly, however, there was

²⁶ *See id.*

²⁷ As noted earlier, leverage cannot be singled out as the sole or even primary contributing factor. Like a “perfect storm” several factors came together to create the financial crisis. Low interest rates encouraged institutions and investors to seek out higher risk/higher return investment opportunities. Technology enabled investment banks to quickly create and sell structured securities. The global trade imbalances led to rapid growth of dollar balances in the portfolios of developing nations and as those investors sought out higher return investments they financed increased leverage by consumers and institutions in developed countries. Ethan Cohen-Cole et al., *Looking Behind the Aggregates: A Reply to “Facts and Myths about the Financial Crisis of 2008”* (Fed. Reserve Bank of Boston, Working Paper No. QAU08-5, 2008) available at <http://www.bos.frb.org/bankinfo/qau/wp/2008/qau0805.pdf>; V.V. Chari et al., *Facts and Myths about the Financial Crisis of 2008* (Fed. Reserve Bank of Minneapolis, Working Paper No. 666, 2008) available at <http://www.minneapolisfed.org/research/WP/WP666.pdf>.

no direct mechanism for regulatory authorities to take control of these firms and wipe-out the equity holders' claims if the firms' operations became too risky, and therefore these firms did not face the direct threat to franchise value that HLR emphasize is essential to motivate maintaining a capital cushion.

As discussed earlier, the two mortgage GSEs operated throughout the 2002-2007 period with essentially the minimum capital required by their regulator and returned excess capital to equity holders through dividends and share repurchases. The largest commercial banks (those with assets in excess of \$10 billion) maintained an average total capital ratio of 11.8% from 2001 through 2008 compared to the 12.3% average for all commercial banks, which is a substantial difference when compared to the 10% standard for well-capitalized banks. Many of the banks in this category might have viewed themselves as being too big to fail and were therefore willing to carry a smaller capital cushion.²⁸ HLR points out that this behavior is entirely rational and predictable for owners of firms in their position. The debt-favoring factors (tax-benefits and implicit insurance benefits) significantly outweigh the expected cost of lost franchise value when early intervention by a regulator is unlikely. Thus, the optimal leverage for these firms lies above that implied by the capital standard and they rationally choose to hold no more capital than required. In addition, these institutions were able to further circumvent regulation by using derivatives and special purpose subsidiaries of holding companies to undertake additional financial risk in ways that entailed less risk of early regulatory takeover and loss of franchise value. Further, while subject to capital regulations, neither bank holding companies nor investment banks are subject to statutory threat of receivership for violating capital standards under Prompt Corrective Action (PCA).²⁹

During a crisis that involves substantial deleveraging that is destabilizing markets, regulators may have an incentive not to enforce capital standards. Enforcing capital standards during a period of rapid asset price decline and thin securities markets will lead to additional asset sales exacerbating asset price declines. As noted by HLR, it is the threat of receivership that causes financial institutions to hold a capital buffer, and those institutions will not hold a sufficient buffer if they know that

²⁸ For data on commercial bank capital ratios, see Federal Deposit Insurance Corporation, *Statistics on Depository Institutions*, <http://www2.fdic.gov/sdi/index.asp> (last visited Nov. 3, 2009).

²⁹ GAO REPORT, *supra* note 13, at 28-42.

regulators are unlikely to enforce capital standards during a financial crisis.³⁰ PCA might be viewed as a reasonable policy to address this problem in that the government has committed to act when banks violate capital standards by placing a legislative mandate on bank regulators.³¹

Finally, the HLR model suggests that an additional factor contributing to the trend toward higher leverage was diminished franchise value. Franchise value has many different sources, but a common source is the restriction of competition. In the last two decades, there has been a strong trend toward deregulation and the elimination of barriers to competition. Before 1999, in the U.S. the Glass-Steagall Act limited competition between commercial banks and investment banks for the provision of certain financial services. In 1999, the Gramm-Leach-Bliley Act reduced those barriers. While increased competition may well lead to the elimination of excess profits and more competitive pricing of services, it also eliminates a factor that contributes to franchise value.³² In saying this, we are not attributing blame to the Gramm-Leach-Bliley Act, but rather simply pointing out that regulators need to monitor all sources of franchise value and react appropriately. The loss of franchise value due to increased competition domestically and globally might help explain the increasing levels of leverage and other risk taking behavior of well established financial institutions, such as the major investment banks and the mortgage GSEs, over the past decade as global competition, in the case

³⁰ In game theory, this situation is known as a non-credible threat where an individual cannot commit to an action in the future (a priori) because all players know that the action will be irrational in the future (ex post). A standard solution in such situations is a commitment device that the player imposes on themselves requiring the action in the future even though irrational ex post. PRAJIT K. DUTTA, *STRATEGIES AND GAMES: THEORY AND PRACTICE* (M.I.T. Press 1999).

³¹ While a minority view, some have argued that the federal government is not enforcing PCA in violation of federal law during this crisis. See Bill Moyer's Journal, *William K. Black on the Prompt Corrective Action Law*, http://www.pbs.org/moyers/journal/blog/2009/04/william_k_black_on_the_prompt.html (April 6, 2009 8:28 EDT).

³² For a discussion of how increased competition can reduce franchise value and increase risk taking by financial institutions, see Thomas F. Hellmann et al., *Liberalization, Moral Hazard in Banking, and Prudential Regulation: Are Capital Requirements Enough?*, 90 AM. ECON. REV. 147 (2000).

of investment banks, and an expanding subprime market, in the case of the GSEs, eroded the franchise values of those firms.³³

IV. IMPLICATIONS FOR REGULATORY REFORM

Recent actions to prevent the failure of many large financial institutions have significantly extended the “Too Big to Fail” umbrella and suggest an extension of the moral hazard problem well beyond the realm of depository institutions. We believe an essential element of the new regulatory framework must include an effective method for dealing with this extension. Based on our work, we believe a critical element of the new regulatory framework must be the implementation of substantive capital regulation for all large highly-levered financial institutions (i.e., all institutions “Too Big to Fail”). The U.S. Treasury Department’s recent legislative proposal to create a single regulator responsible for the capital regulation of systematically important firms provides a fairly simple and straightforward mechanism for accomplishing this goal.³⁴

Moreover, to achieve the appropriate incentive structure, regulators must have the clear authority to move quickly to takeover firms that violate those standards and wipe-out the shareholder equity claims. Only by putting significant franchise value at risk will the capital standards provide the incentive for owners and managers to maintain excess capital appropriate to the unique risks of their firms. As discussed earlier, we believe that it is not feasible to expect regulators to have as much detailed knowledge of a financial institution's investment strategies and risk exposure as the institution’s employees, and so it is important to have a regulatory system that causes the institutions themselves to maintain a capital cushion out of self-interest based their own mix of financial assets and risk exposure.

A credible threat of receivership requires that regulators be committed to enforcing capital standards for large financial institutions in

³³ See Bharat N. Anand et al., *Does Competition Kill Relationships? Inside Investment Banking*, December 2001, http://www.webmanager.cl/prontus_cea/cea_2002/site/asocfile/ASOCFILE120030327160830.pdf.

³⁴ For a general outline of this proposal see Press Release, U.S. Dep’t of the Treasury, Treasury Outlines Framework for Regulatory Reform (Mar. 26, 2009) available at <http://www.treasury.gov/press/releases/tg72.htm>. Also, for a discussion of non-legislative actions to tighten regulation of financial institutions, see Binyamin Appelbaum, *Obama Administration Pushing for Regulatory Reform on Many Fronts*, WASH. POST, Sept. 23, 2009, at A25.

the midst of serious financial turmoil. If large institutions believe that regulators will be unwilling to enforce capital regulations during major economic downturns they will not view the threat of receivership as credible. Legislative requirements of Prompt Corrective Action or PCA is one possible commitment tool that might be imposed on “Too Big to Fail” institutions in order to create the proper incentives for these large financial institutions.

At the same time, regulators must consider the implications of enforcing capital standards during economic crises on deleveraging and asset prices when designing policies for capital regulation. During the crisis, financial institutions may not be able to raise capital and might be forced to sell assets further depressing asset values if capital regulation is strictly enforced. One solution to deleveraging is to actually take over the failed institution allowing regulators to hold onto assets until the market has stabilized much like the Resolution Trust Corporation did during the savings and loan crisis.³⁵ Further, it may be very difficult to value assets in order to “Mark-to-Market” during these periods as a result of the lemons problem and/or very few trades in the asset markets. There are no clear solutions to the problem of marking assets to market during periods of financial turmoil, but at a minimum we believe that “Mark-to-Market” should be based on Generally Accepted Accounting Principles (GAAP).³⁶

Further, the new regulatory framework must be comprehensive and flexible enough to adapt to changing financial instruments. For example, we must avoid the problem of basing capital standards on balance sheet assets and liabilities while institutions use off balance sheet derivatives to take on equivalent risks. Compartmentalizing regulatory authority based on type of instrument will provide institutions tempting opportunities for circumventing the intent of the regulation. However, recognizing that no regulatory scheme can keep ahead of the market in terms of security design just re-emphasizes the importance of a regulatory capital scheme that

³⁵ The Resolution Trust Corporation was created by the federal government during the savings and loan crisis in order to hold and dispose of assets of insolvent thrifts. Lee Davidson, *Politics and Policy: The Creation of the Resolution Trust Corporation*, 17(2) FDIC BANKING REV. 17 (2005).

³⁶ During the savings and loan crisis, S&L's were subject to Regulatory Accounting Principles rather than GAAP. In May 1987, the Federal Home Loan Bank Board began phasing in the use of GAAP in response to concerns that regulatory accounting principles had contributed to the crisis. See Federal Department Insurance Corporation, *The S&L Crisis: A Chrono-Bibliography*, <http://www.fdic.gov/bank/Historical/s&l/> (last visited Nov. 3, 2009).

provides incentives for firm managers themselves to assess the risk of new securities and provide a sufficient capital cushion against future shocks.

Finally, the new framework should provide that when capital standards are violated, unsecured debt holders are not necessarily fully protected as they have been in the current crisis. Properly motivated, unsecured, debt holders can provide significant market discipline to an institution. Throughout the public debate over “bailing out” various firms, most attention has been given to the fact that in some cases the owners have retained a valuable equity stake (e.g., Bear Stearns, Merrill Lynch, Countrywide and AIG) and managers have received significant bonuses (famously, AIG).³⁷ However, little attention has been placed on the fact that in most cases, debt holders were the major beneficiaries of the bailouts. The government intervention and apparent willingness to fund future operating losses via further infusions of capital provides assurance of payment in full to debt holders.³⁸ Regulators should have powers similar to those of a bankruptcy court to impose modifications to unsecured debt contracts to avoid such windfalls and in fact to impose a share of the loss on unsecured debt holders so that debt holders will have an incentive to monitor and assess the investment and financial decisions of large, highly levered financial institutions.

Finally, because franchise value plays such an important role in encouraging prudent managerial choices, the new regulatory framework must be prepared to react to market developments that affect franchise value. Global competition, domestic and international tax policies, technological changes and regulatory changes all have the potential to erode franchise value. Clearly, concerns about financial institution franchise value cannot and should not drive policy in these areas. However, regulators need to be cognizant that financial institution managers will rationally react to declines in franchise value by taking on additional risk. Capital standards and regulatory monitoring activities need to reflect this reality.

³⁷ David Goldman, *AIG Bonuses: \$235 Million To Go*, CNN MONEY.COM, July 10, 2009, http://money.cnn.com/2009/07/10/news/companies/aig_bonuses/index.htm.

³⁸ For example, the first business day after the mortgage GSEs were placed in receivership, the market value of GSE debt and MBS securities jumped significantly providing a windfall profit to recent purchasers.

PREDATORY LENDING AND ITS INSURANCE CONSEQUENCES

Erin O'Leary*

This note distinguishes predatory from subprime lending, while focusing on the insurance consequences of predatory lending. It considers how single premium credit insurance (SPCI) and private mortgage insurance (PMI), two mortgage-related insurance products, have affected the current predatory lending crisis. This note argues for reform that eliminates SPCI and makes PMI a more feasible option for insureds. Such reform would allow subprime lenders to offer mortgages to qualified borrowers, while reducing the amount of predatory lending and foreclosures. The introduction of this note presents some background information regarding subprime lending and predatory lending. The second part examines several issues concerning the role of insurance in the subprime mortgage market. Third, reform measures necessary to alleviate the issues with mortgage insurance are discussed. Finally, the fourth section studies recent actions by the Federal Reserve Board and analyzes whether they can be expected to bring meaningful change. It concludes that, although the Fed's new regulations are a step in the right direction, there needs to be an outright ban of SPCI and predatory must be stopped completely.

I. INTRODUCTION

There is no question that a crisis is gripping the subprime market. As of January 2009, 1.5 million homes had been lost to subprime foreclosure.¹ Another two million subprime mortgage holders are currently

* University of Connecticut School of Law, Juris Doctor Candidate for the Class of 2010. The author wishes to thank John McGrath for his assistance and topical expertise.

¹ SONIA GARRISON ET AL., CTR. FOR RESPONSIBLE LENDING, CONTINUED DECAY AND SHAKY REPAIRS: THE STATE OF SUBPRIME LOANS TODAY 2-3 (2009), available at http://www.responsiblelending.org/mortgage-lending/research-analysis/continued_decay_and_shaky_repairs.pdf.

delinquent.² The fallout from the crumbling subprime market affects not only subprime borrowers but also their communities.³ The crisis has turned “subprime” into a dirty word. Many Americans have heard enough on the news to know that this type of lending is largely to blame for the credit crisis. However, the subprime market has valid and socially valuable applications; thus, before one can identify the most troubled areas of the mortgage industry in order to aim reform measures at these problem spots, the critical distinction one must draw is between *subprime* and *predatory* lending, which represents only a narrow but dangerous sliver of the subprime market.

There is a legitimate need for subprime lending. Not all borrowers can qualify for prime loans, typically because of a negative credit event in their history or because they lack the cash for a down payment.⁴ Loan qualification is typically based on credit scores, as calculated by the Fair Isaac Credit Organization (FICO). Some examples of negative credit events that would disqualify a borrower from prime-rate loans include a history of default, bankruptcy, or low or no credit. Such borrowers are considered to present too great a risk of delinquency, default, or foreclosure to justify the prime rate. Nevertheless, there is a strong policy argument in the United States for enabling as many people as possible to become homeowners. Therefore, in order for these somewhat risky borrowers to become homeowners, they may need to resort to other types of loans. These loans are usually considered subprime. There is no one accepted

² *Id.*

³ The Center for Responsible Lending (CRL) estimates that because of what is known as the “spillover effect,” subprime foreclosures will drain properties surrounding foreclosed homes of nearly \$352 billion in value from 40 million nearby families, averaging out to almost \$9,000 per family. CTR. FOR RESPONSIBLE LENDING, UPDATED PROJECTIONS OF SUBPRIME FORECLOSURES IN THE UNITED STATES AND THEIR IMPACT ON HOME VALUES AND COMMUNITIES 1 (2008), available at <http://www.responsiblelending.org/mortgage-lending/research-analysis/updated-foreclosure-and-spillover-brief-8-18.pdf>. These estimates were made in August 2008 and update an earlier study made in January 2008 that was considered “wildly pessimistic” in its estimate of a \$202 billion loss in neighboring home values and an average loss of \$5,000 for 40 million nearby families. See CTR. FOR RESPONSIBLE LENDING, SUBPRIME SPILLOVER: FORECLOSURES COST NEIGHBORS \$202 BILLION; 40.6 MILLION HOMES LOSE \$5,000 ON AVERAGE 1 (2008) available at <http://www.responsiblelending.org/mortgage-lending/research-analysis/subprime-spillover.pdf>

⁴ David Anderson, *The Subprime Lending Crisis*, 71 TEX. B.J. 20 (2008).

definition of subprime, but it is generally thought to encompass this category of loans made to borrowers whose backgrounds make them a riskier investment for lenders.

Because these borrowers present an increased risk to the lender, the terms of the loans they receive are generally less favorable than those offered to prime borrowers.⁵ For example, 70% of subprime loans come with a prepayment penalty, while only 2% of prime loans do.⁶ Subprime loans also tend to have higher fees than their prime counterparts. Most subprime loans have interest rates that change, usually with little predictability as to the direction or magnitude of the change. Additionally, prime loans tend to be similar from lender to lender, facilitating comparison. Subprime loans are considered by Fannie Mae and Freddie Mac to be non-conforming loans and tend to vary widely among originators. A higher annual percentage rate (APR) is the price that subprime borrowers must pay in order for the lender to take on the risk of the borrower defaulting, going into delinquency, or being forced to surrender his home to foreclosure.

Subprime loans have been increasing in popularity at a rapid pace. In 2003, just 8% of all mortgage originations were subprime loans. By 2006, that figure had jumped to 28%.⁷ Several factors contributed to this rise. One important example is the recent housing bubble. Soaring home

⁵ See Kathleen C. Engel & Patricia A. McCoy, *A Tale of Three Markets: The Law and Economics of Predatory Lending*, 80 TEX. L. REV. 1255, 1265–66 (2002). (explaining the costs that justify higher rates and fees for subprime loans). It is important to note that many of the costs that justify higher costs and fees, such as a more thorough review of a customer's income and credit, lower loan principal amounts that lead to higher origination costs as a percentage of the total loan, and higher incidence of prepayment, are much lower in predatory loans, if they exist at all. For example, predatory lenders often fail to verify a customer's income, debt obligations, and assets before making a loan, thereby erasing most of the cost of income verification. *See id.*

⁶ CRL, A SNAPSHOT OF THE SUBPRIME MARKET 1–3 (2007), available at <http://www.responsiblelending.org/mortgage-lending/tools-resources/snapshot-of-the-subprime-market.pdf>. That number is probably low; Professors Engel and McCoy estimate that at least 98% of subprime home loans contain “substantial prepayment penalties.” Engel & McCoy, *supra* note 5, at 1285.

⁷ CTR. FOR RESPONSIBLE LENDING, *supra* note 6, at 2.

values reassured lenders that if the properties for which they were issuing mortgages would at the very least retain their value.⁸

Predatory lending, on the other hand, has no legitimate basis. The vast majority of predatory loans occur in mortgages that would otherwise be considered subprime. While subprime lenders have a legitimate argument that they must increase the fees and rates associated with their products in order to hedge against the increased risk of default presented by borrowers with less than ideal credit, predatory lending practices have no place in the world of legitimate loans. They are characterized by practices that strip homeowners of the equity they have in their properties extract as much in fees, rates, and other charges as possible from the homeowner.

The subprime lending crisis has had obvious and devastating ramifications for millions of subprime borrowers. There have also been less apparent consequences for the insurance industry. This industry has become closely intertwined with the mortgage industry. As the problems in the subprime market have become exacerbated, so too have the problems in mortgage-related insurance products. This paper considers how two mortgage-related insurance products, namely, single premium credit insurance (SPCI) and private mortgage insurance (PMI), have affected the predatory lending crisis. I contend that through reform measures that eliminate SPCI and make PMI a more feasible option, the subprime lending market can continue to offer mortgages to qualified subprime borrowers while simultaneously reducing the tide of predatory lending and foreclosures.⁹

In Part II of this Note, I examine two issues concerning the role of insurance in the subprime mortgage market. The first is the declining centrality of the role played by providers of private mortgage insurance (PMI) to insure many subprime mortgages. I will explain the effect that predatory lending has had on legitimate PMI policy providers. The second issue is the abuse of credit insurance in subprime mortgage loans. In particular, I will focus on how single premium credit insurance functions as one of the more abusive terms that predatory lenders employ. I will also explore the related nature of these two problems. Moreover, I will provide a brief synopsis of some of the most common abusive tactics used by

⁸ See Christopher Barlow, *Developments in Banking & Financial Law, 2006-2007: The Subprime Mortgage Crisis: Macroeconomic Implications of the Subprime Lending Crisis*, 27 REV. BANKING & FIN. L. 57, 57-58 (2008).

⁹ This paper considers primarily legislation-based reform rather than litigation.

predatory lenders that have led to the complications that the mortgage industry is facing today.

In Part III, I discuss the reform measures that will be necessary to alleviate the problems related to mortgage insurance. I will also focus on two reform measures: first, the Federal Reserve Board's newly finalized amendments to Regulation Z, which implements the Truth in Lending Act (TILA) and the Home Ownership and Equity Protection Act (HOEPA); and second, H.R. 3915, or the Mortgage Reform and Anti-Predatory Lending Act of 2007, which was approved by the House but was never voted on by the Senate. Additionally, I suggest that PMI is an important product that can be used to stem the tide of foreclosures because of the mutual interests shared by borrowers and insurers.

Part IV examines whether the Fed's new amendments to Regulation Z can be expected to bring meaningful change and explores what other measures must be taken in order to stop predatory lending for good, including the effect that H.R. 3915 would have if re-introduced in the next session of Congress.

II. MORTGAGE-RELATED INSURANCE PROBLEMS IN THE SUBPRIME AND PREDATORY LENDING CONTEXT

A. PRIVATE MORTGAGE INSURANCE¹⁰

1. Overview

A private mortgage insurance (PMI) policy is one that protects the lender by paying the costs of foreclosing and guaranteeing a certain portion of the debt; in most cases, that portion is 20 per cent¹¹. PMI is the only

¹⁰ There are multiple types of PMI, but the type that is the focus of this article is primary PMI, which protects against mortgage default. PMI in this article is used to refer to primary PMI. Other kinds of PMI not discussed in this article include pool insurance, which insures groups of individual mortgages and provides 100 per cent coverage for any default losses on mortgages in the pool, subject to a total loss limit on all mortgages in the pool, and PMI reinsurance. Quinton Johnstone, *Private Mortgage Insurance*, 39 WAKE FOREST L. REV. 783, 783 (2004).

¹¹ OFFICE OF THE COMPTROLLER OF THE CURRENCY, Answers About Credit Insurance, http://www.helpwithmybank.gov/faqs/insurance_credit_life.html. Policies that pay 20-30% of the amount of the outstanding debt typically allow lenders to retain title to the property after foreclosure. Mortgage INSURANCE

type of mortgage insurance that lenders can require before making a loan.¹² Often, a potential borrower cannot afford a 20 per cent down payment on a home mortgage.¹³ On the other hand, most lenders are hesitant to make such a loan because these borrowers tend to have a higher risk of defaulting. PMI has served a conciliatory function and permitted these individuals to become homeowners, while alleviating the fears of lenders.¹⁴

Not only is the PMI industry regulated at the state level by laws that vary from state to state¹⁵, but also in part by the federal Homeowner Protection Act of 1998, which took effect on July 29, 1999.¹⁶ If PMI is required as a condition for a mortgage, this statute protects homeowners by requiring both automatic cancellation of the PMI and notice of cancellation rights with respect to PMI.¹⁷ For borrowers with good payment history who are current on their mortgage payments and who can show that the property has not declined below its original value, the PMI can be cancelled on a predetermined cancellation date.¹⁸ This date typically occurs when the balance of the mortgage reaches 80 per cent, which occurs by the borrower making sufficient payments of principal on the mortgage, through appreciation in the value of the property value, or a combination of both.¹⁹ High risk loans are exceptions, however: lenders must give

COMPANIES OF AMERICA, FACT BOOK (2008-09), *available at* <http://www.privatemi.com/news/factsheets/2008-2009.pdf>.

¹² *Id.*

¹³ SAN FRANCISCO FEDERAL RESERVE BANK, PRIVATE MORTGAGE INSURANCE (2008).

¹⁴ *Id.*

¹⁵ Johnstone, *supra* note 11, at 802-03. Although all states impose their own financial requirements on PMI companies doing business within their borders, there are several similar requirements: all states have substantial reserve requirements, including contingency reserves, loss reserves, and unearned premium reserves. *Id.* at 813-14. Many have paid-in capital and paid-in surplus requirements. *Id.* at 814. Some require that the risk to capital ratio does not exceed a certain threshold; for example, it must not be greater than 25:1. *Id.* at 815. Most also restrict the kinds of investments that PMI companies can make and require that master policies of PMI companies be filed for approval with the state's regulatory authority. *Id.* at 815-16.

¹⁶ 12 U.S.C.A. §§ 4902-4903.

¹⁷ 47 C.J.S. *Interest & Usury* § 515 Homeowners Protection Act (2008).

¹⁸ *Id.*

¹⁹ WILLIAM F. GALVIN, SECRETARY OF THE COMMONWEALTH OF MASSACHUSETTS, QUESTIONS AND ANSWERS ON PRIVATE MORTGAGE INSURANCE

borrowers notice of the automatic cancellation provisions, which state that while PMI *may* be cancelled when the borrower has reached 20 per cent equity, it *must* be cancelled by the lender when the loan reaches 22 per cent equity or 78 per cent of the loan outstanding.²⁰ The statute also requires that, at the time the home is purchased, the lender must give written notice of when the borrower may cancel PMI.²¹

The impact of the subprime crisis is becoming more apparent, as well as the effect on PMI providers. Most prime lenders require PMI when the loan-to-value ratio is 80 per cent.²² This helps to ensure that if the borrower is forced into foreclosure, the lender will not lose the difference between the selling price and the balance on the loan as well as the foreclosure fees.²³ If the borrower does default, the insurance policy pays out 20 per cent of the loan amount and the bank can recoup the rest through the foreclosure sale. PMI insurers are left holding the bag if borrowers default early into their mortgages.

The subprime crisis has caused two central problems for PMI providers: loss of market share and significant financial losses. First, at the height of subprime lending in 2005 and 2006, PMI providers lost significant market share. With the relatively new risk-spreading maneuver known as the 80-10-10 or "piggyback loan", a buyer could finance enough of the purchase price such that the first mortgage holder would not require PMI, discussed *infra*.²⁴ Borrowers sought to avoid PMI for a variety of reasons. Until recently, PMI could not be used as a tax writeoff, but the

AND THE FEDERAL HOMEOWNER PROTECTION ACT OF 1998, *available at* <http://www.sec.state.ma.us/cis/cispmi/pmiidx.htm>.

²⁰ *Id.*

²¹ 12 U.S.C.A. § 4902-4903.

²² SAN FRANCISCO FEDERAL RESERVE BOARD, *supra* note 14.

²³ NATIONAL CONSUMER LAW CENTER, COMMENTS TO THE FEDERAL RESERVE BOARD'S PROPOSED REVISIONS TO REGULATION Z TRUTH IN LENDING REGARDING PROPOSALS TO ADDRESS PREDATORY MORTGAGE LENDING (Mar. 9, 2001).

²⁴ Patrick Rucker, *Subprime crisis will boost mortgage insurers*, REUTERS, Aug. 23, 2007, *available at* <http://www.reuters.com/article/gc06/idUSN2263259520070823> (acknowledging the loss in market share experienced by PMI companies, but predicting that those mortgage insurers who can "weather" the current financial crisis will be in a good position to regain market share).

piggyback loan did.²⁵ Also, PMI is paid for and maintained by the borrower, thus adding an obvious cost to the loan.²⁶ For many borrowers, this cost can be especially difficult to swallow because PMI is not first party insurance; rather, it is in place to protect the *lender*. By contrast, the costs of piggyback lending are more subtle, such as higher interest rates. Additionally, underwriting standards employed by PMI companies tend to be much stricter than those used by piggyback lenders.²⁷ As I will explain, this is due in part because PMI companies and borrowers share a unity of interest that simply does not exist between brokers of piggyback loans and their borrowers.

An 80-10-10 loan is actually two loans: one is a first-lien mortgage on the property that covers 80 per cent of the purchase price. The second loan is, in most cases, a ten per cent loan that covers part of the remaining 20 per cent so that the borrower does not have to purchase PMI.²⁸ Thus, with the increase in predatory loans in the past five years, the mortgage industry has seen a concurrent increase in 80-10-10 loans. Because the 80-10-10 loan is a tool invented in part to *avoid* PMI, it is easy to see how a sharp increase in its usage would translate into decreased profits for PMI providers.²⁹ As an illustration of this point, piggyback loans were virtually nonexistent in 2000, but by 2006 about 22% of owner-occupied houses had piggyback subordinate lien mortgages, and the number and dollar volume

²⁵ See Robert B. Avery, Kenneth P. Brevoort & Glenn B. Canner, The 2006 HMDA Data, 93 Fed. Reserve Bulletin A73, A84-A85 (2007), available at <http://www.federalreserve.gov/pubs/bulletin/2007/pdf/hmda06final.pdf>.

²⁶ Todd J. Zywicki & Joseph D. Adamson, *The Law & Economics of Subprime Lending*, 80 U. COLO. L. REV. 1, 42 (2009).

²⁷ *Id.*

²⁸ While most prime loans require PMI if the borrower puts less than 20 per cent down, many subprime lenders, although they are willing to lend to customers who cannot put very much down, will demand either PMI or some other type of credit insurance if the borrower is not able to put down 25 or even up to 30 per cent. NATIONAL CONSUMER LAW CENTER, INC., *supra* note 24.

²⁹ *Id.* Private mortgage insurers have recognized the threat to their product that is posed by this 80-10-10 or “piggyback” loan. On the website of the Mortgage Insurance Companies of America (MICA), an industry organization, there is an entire section dedicated to warning consumers of the dangers of this type of loan. See MICA, QUICK FACTS ABOUT 80-10-10 LOANS (2009), <http://www.privatemi.com/news/factsheets/quickfacts.cfm>; see also MICA, MYTHS AND FACTS ABOUT PIGGYBACK LOANS (2009), <http://www.privatemi.com/news/factsheets/myths.cfm>.

of piggyback loans had risen dramatically during that period.³⁰ By contrast, the number of homes purchased backed by PMI declined about 6% from 2005 to 2006 alone.³¹

The second major problem that predatory lending has created for PMI companies is an unprecedented financial loss.³² Traditionally, PMI has been relatively inexpensive,³³ for the simple reason that people tend to have tremendous incentives and motivation to pay off their home mortgage loans.³⁴ However, as predatory lending creates an increasing number of dangerous loans whose terms quickly become unaffordable, mortgages are entering foreclosure at unprecedented rates.

It became clear that many subprime loans were dangerous investments and carried with them a high risk of borrower default, and providers shied away from them. Insurers who had issued policies for any of these predatory loans suddenly found themselves paying claims on the policies at rapidly increasing rate.

Most PMI companies did not want to be associated with predatory lending, particularly as it became an increasingly well-known phenomenon. Moreover, because predatory lending often leaves a borrower with a loan that he or she cannot afford, from a business standpoint, these loans were much more likely to result in default and to trigger a payout. Now, several insurers have publicly refused to insure loans with certain characteristics.³⁵ Many of these are characteristics that often lead commentators to label loans as “predatory.”³⁶

An examination into the evolution of underwriting standards of PMI companies reveals their scramble to limit their losses and regain financial stability. For example, Genworth Mortgage Insurance has issued statements explaining that, as a response to predatory lending, it will not insure loans that have excessive fees and costs, prepayment penalties without a correlating borrower benefit, a history of repetitive financing

³⁰ Zywicki & Adamson, *supra* note 27.

³¹ *Id.*

³² Vikas Bajaj, *More Lenders Feeling Pain from Defaults*, N.Y. TIMES, July 31, 2007.

³³ Bankrate.com, *The Basics of Private Mortgage Insurance* (June 1, 2001), <http://www.bankrate.com/brm/news/mtg/20010601b.asp> (last visited July 19, 2009).

³⁴ Brian M. Heat, *Note, Hoosier Inhospitability: Examining Excessive Foreclosure Rates in Indiana*, 39 IND. L. REV. 87, 100-01 (2005).

³⁵ *Id.*

³⁶ *See* § 2.a.ii, *infra*; *see also* Engel & McCoy, *supra* note 6, at 1260.

(“flipping”), or SPCI attached to the loan.³⁷ The policy in place at the Mortgage Guaranty Insurance Corporation (“MGIC”) is very similar to Genworth’s.³⁸ A series of updates and changes to the underwriting policies of the Radian Guaranty, Inc. reflect the company’s progression toward increasingly strict standards. The first major change to Radian’s guidelines became effective on September 17, 2007 and eliminated some of the more obviously dangerous practices, such as the Radian 103% LTV program, which presumably insured loans that were made for up to 103% of the value of the property.³⁹ Effective February 1, 2008, the maximum LTV ratio for subprime loans became 95%.⁴⁰ Effective March 31, 2008, the minimum credit score for a subprime loan was raised to 660, and interest only loans as well as cash-out refinances became ineligible for insurance.⁴¹ Radian’s increasingly strict guidelines provides an effective illustration of how the fallout from the subprime crisis became slowly apparent to PMI companies over a period of time.

Mortgage-related insurance products and predatory lending are inextricably intertwined in that as long as predatory lenders continue to make extremely high-cost loans that are likely to cause borrower default, PMI providers will continue to be reluctant to insure these policies because of the likelihood that they will have to pay out the policy. PMI providers have identified several practices that are associated with predatory lenders and, for reasons both practical and ethical, they often refuse to insure any loans with some of the following properties.

³⁷ See GENWORTH MORTGAGE INSURANCE, POLICY STATEMENT ON “PREDATORY” LENDING (2008), <http://www.mortgageinsurance.genworth.com/Legal/PredatoryLendingGuidelines.aspx> (discussing practices that will cause the company to refuse to insure the loan in which the practices are found).

³⁸ MORTGAGE GUARANTY INSURANCE CORPORATION, ANTI-PREDATORY LENDING POLICY (2008), <http://www.mgic.com/emergingmkt/antipredatorylending.html>.

³⁹ eBulletin from Radian Guaranty, Inc. (Aug. 28, 2007), <http://www.radian.biz/pdf/EBULLETIN%202007-02%20GUIDELINE%20CHANGES%20EFFECTIVE%2009-17-07%20Final.pdf>. Because the program is no longer available, Radian does not thoroughly explain it anywhere.

⁴⁰ eBulletin from Radian Guaranty, Inc. 2 (Dec. 17, 2007), <http://www.radian.biz/pdf/Radian%20eBulletin%202007-03.pdf>.

⁴¹ eBulletin from Radian Guaranty, Inc. 2 (Mar. 7, 2008), <http://www.radian.biz/pdf/Radian%20eBulletin%202008-02.pdf>.

2. Characteristics of Predatory Loans

a. Lack of Verification

No-income, no-asset (NINA) loans have gained notoriety as a significant cause of the subprime crisis, and rightfully so. Brokers were so eager to generate loans that they offered mortgages while requiring little to no documentation of a borrower's income, job, assets, or other obligations,⁴² and they did so at a startling rate: it is estimated that as of 2007, up to half of all subprime mortgages made between 2004 and 2006 had been made without fully documented income.⁴³ Thus, even if they wanted to, brokers could not ensure that these borrowers would be able to meet their monthly payments, particularly when the initial low "teaser" rate reset to a much higher rate.⁴⁴

These practices are all harmful to borrowers but often have almost no cost to the brokers and, to a lesser degree, to the lenders. Once the broker sells the mortgage on behalf of the lender, often earning a sizeable kickback for doing so, he is no longer responsible for servicing the loan, or collecting on the payments each month. Therefore, a broker has no interest in ensuring that a borrower can actually make the loan payments. Rather, the broker's main interest is in maximizing his own profit, which is best accomplished by selling loans at higher rates than the borrower qualifies for and by including prepayment penalties. These are both actions that loan originators reward with a kickback.⁴⁵ Sometimes, brokers are even rewarded with a commission if the borrower refinances and incurs the prepayment penalty.

The loan originators themselves often have no more incentive to write affordable loans than do the brokers. Most lenders bundle these loans into mortgage-backed securities and sell them on the secondary market.⁴⁶ Some retain servicing rights, but others sell these rights, leaving them with virtually no exposure if the borrower defaults. This resale of the mortgage

⁴² AMERICAN ASSOCIATION OF RETIRED PERSONS, AARP COMMENTS TO FEDERAL RESERVE BOARD ON THE HOME OWNERSHIP AND EQUITY PROTECTION ACT (Aug. 15, 2007), available at http://www.federalreserve.gov/SECRS/2007/August/20070816/OP-1288/OP-1288_51_1.pdf.

⁴³ Center for Responsible Lending, *supra* note 7.

⁴⁴ See *infra* § II.a.ii.2.

⁴⁵ See *infra* § II.a.ii.7

⁴⁶ See Engel & McCoy, *supra* note 6, at 1273-74.

bundles serves two crucial functions: it gets the mortgages off of the originator's books and also provides an inflow of capital so the originator can continue to make loans.⁴⁷ Secondary market purchasers of these securities were unable to inspect the individual loans underlying the securities they were purchasing. The securities, however, were rated investment-grade for the most part by the ratings agencies and were thought to be safe investments because they were diverse and because it was thought that with their homes at stake, borrowers would generally not default. Even if one or two failed, investors reasoned, the rest of the portfolio would be sound. When they failed in large numbers, the so-called "subprime crisis" ensued.

However, unlike most brokers or lenders, PMI companies have a strong interest in their insured's ability to repay his or her mortgage. The company will have to pay a claim on the policy if the borrower does not repay; thus, it is in the company's best interest to fully investigate a borrower's ability to repay the loan. The requirement that a customer document his or her ability to repay is now well-reflected in the underwriting standards of PMI companies.⁴⁸

b. "Exploding" Adjustable Rate Mortgages

Many predatory loans are advertised at a very low introductory or "teaser" rate.⁴⁹ Frequently, advertisements claim that the loan is at a low fixed rate. What these advertisements do *not* always disclose, however, is that once the introductory period, which typically lasts no more than two to three years, is over, an estimated 90% of these mortgages suddenly switch to some type of adjustable rate mortgage.⁵⁰ On a 2/28, which is a 30-year mortgage whose rate is fixed only for the first two years and then adjustable for the next 28, a borrower's monthly payments will increase by, on average, 30 to 50 percent in the first month of the third year, hence the

⁴⁷ *See id.*

⁴⁸ *See, e.g.,* RADIAN GUARANTY, INC., UNDERWRITING GUIDELINES, 6-7, 10-11 (2009), *available at* http://www.radian.biz/pdf/Radian_Standard_Guidelines_08_17_09.pdf.

⁴⁹ CENTER FOR RESPONSIBLE LENDING, *supra* note 7, at 1-2, 4.

⁵⁰ *Id.* at 1. These adjustable rate mortgages typically float in relation to an index like the London Interbank Offered Rate (LIBOR), but subprime borrowers often pay a premium over the rate that prime borrowers would pay for a similar product. Anderson, *supra* note 5, at 20.

term “exploding.”⁵¹ Sometimes, a loan is structured as an optional adjustable rate mortgage, which gives a borrower flexibility in the amount he wants to pay on his mortgage each month. These can end up being devastating for the borrower, however: at times, a borrower may be permitted to pay *less* than the monthly interest rate each month, thereby actually adding to his principal. This is a process known as negative amortization.⁵²

c. Excessive Fees

Predatory loans virtually always come with numerous fees, deftly packaged into the loan such that a typical borrower would have a difficult time spotting them.⁵³ Up until now HOEPA has structured its definition of “high cost mortgages” such that lenders could avoid falling within HOEPA’s regulations simply by hiding large fees in their mortgages rather than simply increasing the interest rate.⁵⁴ One of the most onerous and frequently used of these fees is the prepayment penalty.

d. Prepayment Penalties

A prepayment penalty is a provision of the mortgage contract that states that if the borrower pays off the loan entirely or in part, either through payments against the principal or, as is more common, through refinancing, the borrower will incur a penalty. Penalties are usually expressed as a percentage of the mortgage balance at the time of prepayment or alternatively as a set number of months’ worth of interest,

⁵¹ CENTER FOR RESPONSIBLE LENDING, *supra* note 7, at 2. While the average increase in rate is only 30 to 50 per cent, a loan resetting from 7 to 12 per cent would cause the borrower’s payments to increase by 70 per cent. *See Id.*

⁵² Engel & McCoy, *supra* note 6, at 1263. *See also* Mincey v. World Sav. Bank, FSB, 614 F. Supp. 2d 610, 635-638 (D. S.C. 2008) (holding that where a lender violated TILA where he failed to disclose that if a borrower chose to pay off his mortgage as an option-ARM, it would cause negative amortization).

⁵³ These fees greatly exceed the amounts justified by the costs of the services provided and the credit and interest rate risks involved. David Reiss, *Subprime Standardization: How Rating Agencies Allow Predatory Lending to Flourish in the Secondary Mortgage Market*, 33 FLA. ST. U. L. REV. 985, 999 (2006).

⁵⁴ *See* CENTER FOR RESPONSIBLE LENDING, *supra* note 7, at 1-3.

and can easily total five percent of the principal balance of the loan.⁵⁵ The penalties usually can only be invoked in the first part of a loan, i.e. within the first four or five years. Brokers defend these particular penalties with the argument that by applying a prepayment penalties, lenders are willing to lower the interest rate because they are more confident that the borrower will not prepay, such as through a cash-out refinance. Thus, contend brokers, because of the lower interest rate, there is a net benefit to borrowers. However, research indicates that even where lenders do lower the interest rate on mortgage with prepayment penalties, the difference is made up in added fees, evidencing a lack of correlation between the penalty and the purported savings to the consumer. Instead, the cost of the penalty to the average borrower is three to four times the average savings in interest payments.⁵⁶

Prepayment penalties are particularly troublesome for borrowers in two situations. First, if a borrower has successfully improved his credit score such that he now qualifies for a better loan at a prime rate, he invokes a prepayment penalty. Those borrowers who cannot afford to pay the penalty- and there are many- feel they have no choice but to stay with their current mortgages.⁵⁷ Many subprime lenders market their product as a loan that a borrower can use while he or she tries to improve his credit rating such that he can qualify for a prime loan. Given the fact that prepayment penalties lock a consumer into the loan, this claim seems ironic at best and downright fraudulent at worst.⁵⁸ If the borrower chooses to refinance, he incurs a large penalty for doing so. In the words of the Center for Responsible Lending, the borrower incurs “punishment for obtaining a better loan.”⁵⁹

The second situation in which prepayment penalties become onerous for borrowers occurs when the mortgage “explodes” at the end of the introductory period. The borrower abruptly finds himself owing on

⁵⁵ A typical prepayment penalty would be six month’s interest, or about 4.5 per cent of the initial loan balance. *Supra* note 45, at 5. This translates into thousands of additional dollars required for refinancing. *Id.*

⁵⁶ *Id.*

⁵⁷ DEBBIE GOLDSTEIN & STACY STROHAUER SON, CRL, WHY PREPAYMENT PENALTIES ARE ABUSIVE IN SUBPRIME HOME LOANS 1, 4 (2003), available at http://www.responsiblelending.org/mortgage-lending/research-analysis/PPP_Policy_Paper2.pdf.

⁵⁸ *See id.*

⁵⁹ *Id.*

average 30 to 50 percent more each month than he did previously.⁶⁰ Faced with this financial burden, many borrowers attempt to refinance their loan. However, in doing so, they incur the prepayment penalty. Many can't afford to pay the penalty and so are forced to remain in the mortgage. For others, the original lender offers to reduce the prepayment penalty in exchange for an agreement to refinance with the same lender. The prepayment penalty is then added to the principal, along with other fees charged by the lender for refinancing, and the borrower's principal increases.

This tactic is used almost exclusively in subprime loans: 70% of subprime loans have prepayment penalties, while less than two per cent of prime loans have them.⁶¹ This is partially fueled by lenders who pay a premium to brokers who can pad a loan with a high prepayment penalty.⁶²

e. Equity Stripping & Loan Flipping

One of the most prevalent practices among predatory lenders is equity stripping. Equity stripping can take many forms, but in the context of mortgage lending the method of choice for accomplishing this is known as loan flipping.⁶³ When one of these mortgages becomes onerous or unaffordable for the borrower, the lender is quick to offer the opportunity to refinance repeatedly and at short intervals.⁶⁴ The borrower incurs a large prepayment penalty, described below, as well as other fees in addition to the penalty. These costs usually get added to the principal and the borrower begins paying interest on the fees themselves.⁶⁵ Some lenders increase the fees every time they refinance. With each refinancing, the lenders pocket more in fees and the borrower's equity dwindles.

⁶⁰ See Ctr. for Responsible Lending, *supra* note 6, at 2.

⁶¹ *Id.* Prepayment penalties have been limited in the prime market because of better competition among lenders. *Id.* But see Engel & McCoy, *supra* note 6, at 1285 (estimating that at least 98% of subprime loans contain substantial prepayment penalties).

⁶² See *infra* § II.a.6; William F. Bennett, *Mortgage Brokers Get Fatter Payoffs for Selling Riskier Loans*, NORTH COUNTY TIMES, May 5, 2007 (considering the case of California mortgage brokers who are legally obligated to act in the best interest of the borrowers, but who often do not because there is no enforcement mechanism).

⁶³ See Engel & McCoy, *supra* note 5, at 1263.

⁶⁴ *Id.*

⁶⁵ *Id.*

Sometimes the prepayment penalty added to principal borrower pays for every refinancing is less than the cash received during that transaction. Ultimately, having leveraged it all, homeowners are left with little to no equity and owe enormous sums because fees have been folded into the principal amount of their loan.⁶⁶ Thus, the equity has been stripped.

Most of the wealth of the average American is found in the equity of his home.⁶⁷ To unscrupulous lenders, this represents a veritable jackpot, and they attack it as such with aggressive advertising practices, as described below.

f. Aggressive Advertising

Some scholars characterize the tactics employed by predatory brokers and lenders as one calculated to overcome free will.⁶⁸ Whereas borrowers of prime loans tend to seek out a lender, the opposite is true with predatory loans. Predatory lenders tend to use mass mailings, cold calls, and other aggressive advertising techniques to seek out people who often are not even in the market for a loan to begin with.⁶⁹ When seeking out these borrowers, brokers and lenders focus on the most vulnerable populations.⁷⁰ They use public records to identify which individuals or households may owe back taxes. Many prey on elderly homeowners who have built up significant amounts of equity in their homes.⁷¹ The CRL estimates that borrowers age 65 and older have five times the odds of receiving a subprime loan than borrowers younger than 35.⁷² Others use census data to identify low to moderate-income neighborhoods, typically targeting minority groups within those neighborhoods. In 2006, over 50% of loans made to African-Americans were subprime, 40% of loans made to people of Hispanic descent were subprime, but only 22% of loans made to Caucasians were subprime.⁷³ These populations often feel that the prime

⁶⁶ See Goldstein & Son, *supra* note 57, at 3.

⁶⁷ *Id.* at 12..

⁶⁸ See Engel & McCoy, *supra* note 6, at 1346

⁶⁹ See *id.* at 1296.

⁷⁰ See *id.*

⁷¹ AARP, *supra* note 42.

⁷² CTR. FOR RESPONSIBLE LENDING, FACT SHEET: PREDATORY MORTGAGE LENDING ROBS HOMEOWNERS & DEVASTATES COMMUNITIES (2006), available at <http://www.responsiblelending.org/mortgage-lending/research-analysis/2b003-mortgage2005.pdf>.

⁷³ Ctr. for Responsible Lending, *supra* note 6, at 2

market is not available to them because of poor credit history or other major credit event. Of course, as discussed above, up to 50% of these borrowers may in fact qualify for a prime loan.

Brokers and lenders rely on high-pressure marketing tactics, and may go door to door or cold-call.⁷⁴ Their target borrowers are often not even seeking a loan, but marketers tend to be charming and friendly, convincing the target that they are going to help them. Sometimes, would-be borrowers, especially elderly ones, have little to no debt on their home. For these groups, brokers focus on convincing them that they can “help” them by giving them some extra cash. Of course, the customer is then subject to prepayment penalties, exorbitant fees that may get folded into the principal itself, and other abusive practices. For these groups, brokers focus on convincing them that they can “help” them by giving them some extra cash. Of course, the customer is then subject to prepayment penalties, exorbitant fees that may get folded into the principal itself, and other abusive practices. In the worst scenarios, people who once owned their homes free and clear (or close to it) end up the victims of foreclosure.⁷⁵

Predatory lenders act quickly once they have identified their targets. They restrict the amount of information they give out and what information borrowers do receive is confusing and hard to read. Closings are often rushed and accompanied by stacks of paperwork that are insurmountable to the average borrower. Sometimes, brokers misrepresent what the borrower must do in order to close. For example, the broker may tell the borrower that he has to purchase life insurance in conjunction with his loan when that is not the case at all.⁷⁶

g. Yield-Spread Premiums

Yield-spread premiums are kickbacks paid by lenders or originators to brokers if the broker “upsells” a loan. By steering a customer who qualifies for a lower cost subprime or even a prime rate into a much costlier loan, originators can increase the rates and fees on their loans. However, this means that brokers are trying to sell borrowers loans at rates much higher than the customer would otherwise qualify for since these loans are more suitable for customers with poorer credit.⁷⁷ Usually, Fannie

⁷⁴ Engel & McCoy, *supra* note 5, at 1295.

⁷⁵ See AARP, *supra* note 42.

⁷⁶ See Engel & McCoy, *supra* note 5, at 1267.

⁷⁷ See *id.* at 1264-66.

Mae and Freddie Mac studies have found that up to about 50% of all borrowers steered into subprime loans could have qualified for a prime rate.⁷⁸ This figure is partially a result of brokers working in concert with lenders to drive up mortgage rates. In addition, many lenders offer these kickbacks if a brokers can pad the loans with large prepayment penalties.⁷⁹ Lenders are willing to pay a premium for a prepayment penalty that functions as a guarantee that the borrower will either stay in the overpriced predatory loan or that the lender will make money when the borrower is forced to pay the penalty in order to get out of his expensive loan.⁸⁰

These kickbacks are dangerous for another reason. Many unscrupulous brokers employ tactics designed to convince the potential borrower that they, the brokers, are working on behalf of the borrower to obtain the best rate possible. Borrowers are lulled into a false sense of comfort and a belief that the broker is laboring under some sort of fiduciary duty. In reality, exactly the opposite is true: brokers are motivated to sell the borrower the most expensive mortgage possible.⁸¹

h. Over-appraisal and Overselling

Lenders and brokers often work hand in hand with appraisers: when appraisers return favorable, overstated estimates of home values, the lenders and brokers continue to use their services.⁸² Because the cost of their home has been overstated, borrowers end up with loans whose principal is greater than the true value of their home.⁸³ This benefits the lenders, who make more money from interest, fees, and charges if the underlying principal is higher. Many of these loans were made during the housing bubble, when people expected home prices to continue to rise.

⁷⁸ See Goldstein and Son, *supra* note 57, at 1 n.3.

⁷⁹ See *id.* at 1.

⁸⁰ For example, a ContiMortgage Corporation rate sheet shows that for loans with prepayment penalties, the maximum yield-spread premium is 2.5 per cent. For mortgages with prepayment penalties, the maximum premium jumps to 4.25 per cent. AARP, *supra* note 42, at 5-6.

⁸¹ Bennett, *supra* note 62.

⁸² See 73 Fed. Reg. 44522, 44566 (proposed Jul. 30, 2008) (to be codified at 12 C.F.R. pt. 226).

⁸³ See Christopher L. Peterson, *Predatory Structured Finance*, 28 CARDOZO L. REV. 2185, 2223 (2007) (describing a predatory lender who used overstated appraisals to justify loans packed with excessive fees and charges).

Instead, in 2007, the bubble burst and home prices fell.⁸⁴ At the same time, many rates on subprime loans reset from the introductory “low fixed rates” to a higher rate. Borrowers who relied on increasing home values were simultaneously faced with increased mortgage payments, leading to increased foreclosures.

B. THE ABUSE OF CREDIT INSURANCE

Among the many tools in a predatory lender’s arsenal is the use of insurance to extract even more money out of subprime borrowers.⁸⁵ Credit insurance is linked to a specific debt or loan and will pay off that debt if the borrower becomes unable to do so.⁸⁶ Although such a policy protects the lender, it is paid for by the borrower. There are various types of insurance against the borrower’s failure to pay. Among others, some examples include credit health insurance, which protects the lender in case the borrower becomes ill and can no longer make payments on the loan; credit life insurance to protect against the risk of the borrower’s death prior to repaying the loan; and credit unemployment insurance protects against the borrower’s inability to make payments due to job loss.⁸⁷ Whereas it is not uncommon for prime lenders to require private mortgage insurance for loans in which the borrower pays less than twenty per cent of the home’s value as a down payment, these types of credit insurance are rarely found in prime loans, but have been aggressively marketed in the subprime market.

Single premium credit insurance (SPCI) is particularly abusive. With this type of insurance, borrowers are forced to pay a one-time premium that is very high, tending to cost four to five times as much as credit insurance whose premiums are paid on a monthly basis.⁸⁸ This premium is nearly always financed with the mortgage leaving the borrower paying a hefty interest rate in addition to an already enlarged insurance premium.⁸⁹ For these reasons, the Consumer Federation of America has

⁸⁴ Anderson, *supra* note 4, at 20.

⁸⁵ For a description of other abusive practices found in predatory loans, see *supra* Part II.a.

⁸⁶ ASSOCIATION OF COMMUNITY ORGANIZERS FOR REFORM NOW, PREDATORY LENDING PRACTICES, ¶ 15 (2004), <http://www.acorn.org/index.php?id=754>.

⁸⁷ *Id.*

⁸⁸ *Id.* at ¶ 16.

⁸⁹ *Id.* at ¶ 17.

called SPCI on mortgages “the worst insurance rip-off in the nation,” and Fannie Mae and Freddie Mac refuse to purchase loans that included financed credit insurance.⁹⁰

While onerous for the borrowers who must pay the premiums, mortgage credit insurance is lucrative for the lenders. Credit insurance companies typically sell group insurance products to home mortgage and other lenders.⁹¹ One author suggests that some lenders make up to 50% of their pre-tax income from the sale of credit insurance.⁹² The Consumer Federation of America also found in a 1999 report that the credit insurance industry had unusually low loss ratios, with the ratio of claims to premiums often no more than 40% for credit life and disability insurance policies.⁹³ This exceeds the National Association of Insurance Commissioners’ (NAIC) recommended 60% claims to premiums ratio.⁹⁴ The implication from these numbers is that lenders selling credit insurance are profiting at the disproportionate expense of the borrowers who are often deceived into purchasing the policies.⁹⁵

In theory, these products are voluntary. The only form of insurance that a lender may legally require in order to obtain a loan is private mortgage insurance, but once a borrower does purchase credit insurance, it becomes part of his contract.⁹⁶ In many cases, the relatively meager consumer benefits provided by credit insurance last only a few years, while the borrower continues to pay a high interest fee on a large premium for years after the benefits have stopped.⁹⁷ In previous

⁹⁰ *Id.* See also Broderick Perkins, *California Enforcing New Predatory Loan Rules*, REALTY TIMES, July 10, 2002, available at http://realtytimes.com/rtpages/20020710_predatoryloans.htm

⁹¹ Ronald H. Silverman, *Toward Curing Predatory Lending*, 122 BANKING L.J. 483, 504 (2002).

⁹² *Id.* at 504-505

⁹³ *Id.* at 505.

⁹⁴ *Id.*

⁹⁵ *See id.*

⁹⁶ NATIONAL ASSOCIATION OF INSURANCE COMMISSIONERS, CREDIT INSURANCE: SAFETY NET OR NO NET GAIN? (Nov. 2006), http://www.naic.org/documents/consumer_alert_credit_insurance.htm; see also Office of the Comptroller of the Currency, *supra* note 11.

⁹⁷ Edward M. Gramlich, Governor, Federal Reserve Board, Remarks at the Housing Bureau for Seniors Conference (Jan. 18, 2002).

amendments made to HOEPA, the Federal Reserve Board has attempted to regulate SPCI without prohibiting it altogether, but it persists.⁹⁸

At least one appellate-level court has agreed with the proposition that SPCI is inherently unfair. In *Richardson v. Bank of America*, the North Carolina Court of Appeals held that the lender's illegal sale of SPCI was an unlawful and deceptive trade practice under North Carolina law, that it was not made in good faith and did not represent fair dealing, and that it constituted willful and wanton tortuous activity sufficient to support the imposition of punitive damages.⁹⁹

III. REGULATIONS AND REFORM MEASURES

A. OBJECTIVES OF REFORM

1. Competing Interests

Any reform measures aimed at the insurance industry in the context of predatory lending, whether directly or indirectly, must seek to further some important goals. Many of these goals in fact consist of striking a balance between two equally compelling but competing interests. The first and perhaps the most frequently cited set of competing interests involves the conflict between providing consumers with adequate protection against predatory loans and promoting the availability of credit. In December, 2007, Ben Bernanke, Chairman of the Federal Reserve Board, referenced the need to balance these interests when introducing a proposal to amend existing regulations, stating that the proposed rules were crafted with an eye toward deterring improper lending and advertising practices without unduly restricting mortgage credit availability.¹⁰⁰

A broad scope of regulations, whether at the federal or state level, is appealing to consumer advocacy groups but carries a greater risk of encompassing loans and lenders that the drafters of the regulations never contemplated. Stricter regulations tend to cause legitimate subprime lenders to restrict credit for fear of falling into regulated categories. At a time when the country desperately needs credit, any reform measures must

⁹⁸ *See id.*

⁹⁹ 643 S.E.2d 410, 424-425 (N.C. Ct. App. 2007).

¹⁰⁰ Ben S. Bernanke, Chairman, Federal Reserve, Statement of Jul. 14, 2008, available at <http://www.federalreserve.gov/newsevents/press/bcreg/bernankereg20080714.htm>

be just as careful to preserve lenders' willingness to lend as to protect consumers. Of course, on the other hand, most regulatory reform measures are motivated by a desire to stop the abuses inflicted upon subprime borrowers by predatory lenders. To that end, any measures that did not increase protections for consumers would simply be ineffective.

There is also a conflict between forcing predatory brokers and lenders to internalize the harm they cause and requiring borrowers to accept greater personal financial responsibility. On the one hand, when lenders make irresponsible loans with abusive terms that cause a consumer to default and frequently to lose his home to foreclosure, society often must bear the costs.¹⁰¹ On the other hand, the hallmark of responsible borrowing is knowing how much one can afford. According to one author, "the social and moral question centers around who should determine that a particular borrower cannot afford to pay the proposed mortgage."¹⁰²

Especially in today's real estate situation, where it is not uncommon for a borrower to owe more on his home than it is worth, when a home is lost to foreclosure the borrower may be left with nothing and suddenly become dependent on social programs such as welfare to provide for him and his family. Strict regulations, such as stringent requirements regarding verification of a borrower's ability to pay or an escrow requirement¹⁰³ would be costly to implement but ultimately would be likely to reduce foreclosures. Since the lender is in the best position to prevent the harm caused by predatory mortgage loans and their abusive terms, one

¹⁰¹ See, e.g. Manny Fernandez, *Helping to Keep Homelessness at Bay as Foreclosures Hit More Families*, N.Y. TIMES, Feb. 4, 2008, at B6 (discussing a program adopted by New York charitable organizations to help homeowners who are facing foreclosure avoid eviction by giving them up to \$10,000 to pay for the costs of moving, the first month's rent at their new apartments, and other related costs).

¹⁰² David Schmudde, *Responding to the Subprime Mess: The New Regulatory Landscape*, 14 FORDHAM J. CORP. & FIN. L. 709, 730 (2009).

¹⁰³ An escrow requirement like the one proposed in the Fed's finalized amendments to Regulation Z requires lenders to set up an escrow to cover the costs of property taxes and insurance. See CENTER FOR RESPONSIBLE LENDING, PROPOSED RULES REGARDING UNFAIR, DECEPTIVE, ABUSIVE LENDING AND SERVICING PRACTICES PURSUANT TO HOME OWNERSHIP AND EQUITY PROTECTION ACT at 52 (Apr. 8, 2008), available at <http://www.responsiblelending.org/mortgage-lending/policy-legislation/regulators/fed-udap-comments-final-040808.pdf>.

side of the policy argument says that any regulation should favor the consumer in order to force lenders to internalize the harm they cause.¹⁰⁴

The other side of that same policy argument is the side favored by Senator Phil Gramm, among others. Gramm attributed the increase in foreclosures, particularly in the subprime market, to “predatory borrowers,” not predatory lenders.¹⁰⁵ With the spectrum of available subprime products, many borrowers were able to obtain mortgages that they ultimately could not afford.¹⁰⁶ People such as Senator Gramm favor regulations that require the *borrower* to be responsible for ensuring that he only borrows what he can afford. The problem with this, however, is that predatory brokers and lenders often structure their marketing strategies around pressure and removing free will.

2. Policy Goals

The mortgage insurance industry can be reformed by measures aimed specifically at SPCI and PMI reform. However, it is equally if not more important that broader reforms encompassing predatory lending generally be passed.

The preferred method for avoiding PMI is the use of the 80-10-10 loan, which is often a favorite target of predatory lenders.¹⁰⁷ The type of borrower who seeks out these loans is often one who has very little cash and may feel excluded or disenfranchised from the prime market. There may be an element of desperation: often, borrowers want to avoid incurring a PMI premium so badly that they accept credit from any lender willing to offer it. Some may feel that because the loan is quite small in comparison to the larger 80% loan, it is less risky to accept more onerous terms. The 80% loans, particularly when made to subprime borrowers, may also be predatory. Whatever the reason, legislative efforts toward eliminating

¹⁰⁴ See generally Lauren E. Willis, *Against Financial-Literacy Education*, 94 IOWA L. REV. 197 (2008) at A1 (discussing the interaction of foreclosures and societal reaction to the same and positing that regulations are the best way to attack foreclosures, which are costly to the borrower and to society).

¹⁰⁵ Eric Lipton & Stephen Labaton, *Deregulator Looks Back, Unswayed*, N.Y. TIMES, Nov. 17, 2008 (emphasis added)

¹⁰⁶ See Fernandez, *supra* note 102, at B6 (discussing the example of a woman who was able to obtain a \$486,000 mortgage with \$4,000 monthly payments despite the fact that her monthly income was only \$2,800).

¹⁰⁷ Johnstone, *supra* note 10, at 787.

predatory lending are a critical first step toward improving the PMI outlook and making mortgages safer and more affordable for subprime borrowers.

PMI providers avoid predatory loans for reasons that are both ethical and practical. From an ethical standpoint, it is likely that these providers do not want to be associated with a product that has gained so much notoriety as a major cause of rising foreclosure costs and the seizing up of the capital markets. And from a practical point of view, insurance companies are loath to insure loans that carry with them a high likelihood of a borrower default that would trigger payout on the policy. As discussed above, predatory loans carry with them a higher rate of default. Therefore, curbing predatory lending would make subprime loans safer and less expensive for PMI providers to insure. Although there are a relatively small number of PMI providers,¹⁰⁸ there is intense competition among this limited group of companies. Therefore, if the costs of PMI can be reduced across the board, then costs of PMI policies are likely to fall, making it a much more feasible option for borrowers. This will help drop demand for the 80-10-10 loans. If faced with decreased demand and more of the right regulations, along with the increased public awareness, predatory lenders would find it much more difficult to find willing victims.

Perhaps the most appealing aspect of PMI is that it aligns the borrower's interest with the insurer's interest. Restrictive or requirement-based regulations, such as prohibitions on prepayment penalties or mandatory disclosures, are often difficult because the brokers and lenders targeted by these regulations realize that it is frequently not in their best interests. Brokers in particular have little incentive to ensure that the loans they write on behalf of their lenders are sustainable and affordable.¹⁰⁹ If the borrower defaults, the broker is not implicated: he or she has likely already collected his or her yield-spread premium and is unaffected by the foreclosure.

One author has identified the moral hazard faced by brokers and explains that at the same time that the law permits and even encourages a mortgage broker to bargain for the most advantageous outcome, it also asks that broker to serve as the conduit for information that is intended to reduce

¹⁰⁸ Members of MICA, the industry association for mortgage insurers includes AIG United Guaranty, Genworth Mortgage Insurance Corporation, Mortgage Insurance Guaranty Corporation, PMI Mortgage Insurance Co., Radian Guaranty, and Republic Mortgage Insurance Company. Private MI, ABOUT MICA (2009), available at <http://www.privatemi.com/about.cfm>.

¹⁰⁹ See Bennett, *supra* note 62.

his advantage. Setting up such a conflict between the duties we expect a mortgage broker to discharge and that broker's self-interest is ill-conceived.¹¹⁰

Furthermore, in the age of increasing securitization, fewer lenders have kept the mortgage on their books, although some have kept the servicing rights.

PMI is exactly the opposite: here, there is a *unity* of interest between the borrower and the seller of the product. If the policy is still in effect, a foreclosure triggers a payout. Indeed, some PMI providers have even paid off an insured's monthly mortgage debt if the insured loses his job or is temporarily unable to pay, or have otherwise negotiated or mediated between the insured and the lender.¹¹¹ This is just one example of a practice that is mutually beneficial to the insurer and the policyholder: the insured avoids a devastating foreclosure while the insurer does not have to pay out the policy's limits.

For this reason, PMI does not need as much heavy federal regulation that is required by loan originators and brokers; regulatory efforts should thus remain focused on predatory lending tactics. PMI companies are already extensively regulated as insurance at the state levels.¹¹² Furthermore, while the government must implement restrictions and requirements to force brokers and lenders to conduct business in a manner that is fair for consumers, the common desire to avoid foreclosure shared by PMI providers and policyholders ensures that insurers are already motivated to act in the interests of the consumer. In short, if reforms to the predatory lending practices are effectuated such that loans become more affordable and sustainable, private mortgage insurers can once again insure those risks without the fears of foreclosure that have been brought on by the subprime crisis.

¹¹⁰ Lloyd T. Wilson, Jr., *Effecting Responsibility in the Mortgage Broker-Borrower Relationship: A Role for Agency Principles in Predatory Lending Regulation*, 73 U. CIN. L. REV. 1471, 1500 (2005).

¹¹¹ Genworth Financial has stepped in to facilitate mediations and workouts that will keep their insureds in their homes and avoid foreclosure and claims. See Shannon Behnken, *Insurance Company Helps Homeowner Avoid Foreclosure*, TAMPA TRIBUNE, Nov. 25, 2008.

¹¹² Johnstone, *supra* note 10, at 783. This is *not* to say that regulations are *never* needed. The Homeowners Protection Act is an example of regulations that protect consumers by requiring that PMI be cancelable after a certain loan to value is reached. See also 47 C.J.S. *Interest & Usury* § 515 *supra* note 17.

Turning to the another problem with mortgage-related insurance products, reform measures aimed at predatory lending overall should also seek to eliminate SPCI rather than just to regulate it. If this is accomplished, then lenders who require an extra measure of security before lending to subprime borrowers are much more likely to turn to other options. One such option is likely to be PMI: lenders could, if a borrower was paying below 20 per cent down, require PMI. This would mean that instead of disguising SPCI in complicated loan documents to essentially dupe borrowers into financing a single large premium, they would need to find other options if credit protection was truly important to them. Whereas SPCI tends to be deceptive and is a product often offered by predatory brokers and lenders, PMI is a highly regulated industry at both the federal and usually the state levels. Furthermore, by reducing predatory lending, including the use of SPCI, not only are loans more safe for borrowers, they are safer, and by extension cheaper, for policy providers.¹¹³

B. PROPOSED PROTECTIONS

1. The Federal Reserve's Measures

a. *Proposed Measures.*

Creating effective reform measures is, of course, easier said than done. Competing interests, such as those described above,¹¹⁴ are indicative of the wide range of viewpoints about reform. However, the Federal Reserve has made an attempt at effective reform. On December 18, 2007, the Fed proposed and asked for public comment on changes to Regulation Z (Truth in Lending), which was to be adopted under HOEPA¹¹⁵. The board used primarily objective-based triggers, but drew largely from Federal Trade Commission notions about what is considered “unfair” or “deceptive.” The key points of the proposed reform included:

¹¹³ See Kathleen C. Engel & Patricia A. McCoy, *Turning a Blind Eye: Wall Street Finance of Predatory Lending*, 75 FORDHAM L. REV. 2039, 2041 (2007).

¹¹⁴ See *supra* § III.a.i.

¹¹⁵ Press Release, Board of Governors of the Federal Reserve System, (December 18, 2007) (*available at* <http://www.federalreserve.gov/newsevents/press/bcreg/20071218a.htm>).

- Prohibiting a lender from engaging in a pattern or practice of lending without considering borrowers' ability to repay the loans from sources other than the home's value
- Restrictions on prepayment penalties only to loans that met certain conditions, including a requirement that the penalty expire at least sixty days *before* any possible payment increase
- Requiring the lender to establish an escrow account for payment of property taxes and homeowners' insurance, with an opt out provision available only after one year
- Prohibiting abusive servicing practices, including failing to credit payments when the servicer receives it, failure to provide a payoff statement within a reasonable period of time, and "pyramiding" late fees.
- Prohibiting creditors or brokers from coercing or encouraging an appraiser to misrepresent the value of a home
- Prohibiting misleading and deceptive advertising practices for closed-end loans. For example, prohibiting use of the term "fixed" to describe a rate that will explode to an ARM within a few years; requiring applicable rates and payments to be disclosed in advertisements.
- Adjustment to the time frame for disclosures to permit borrowers to use the information to make comparisons to other mortgage products.¹¹⁶

b. *Public Reaction and Comments*

During the first part of 2008, various groups, including industry interest groups, consumer advocacy groups, local governments and others submitted comments about the proposal. Two of the most significant groups to offer comment included the consumer advocates and the industry interest groups.

¹¹⁶ *Id.*

Consumer interest groups offering comment included the American Association for Retired Persons (AARP), the National Consumer Law Center (NCLC), the Center for Responsible Lending (CRL), National Council of La Raza (NCLR), and several smaller, state-level groups. Overall, consumer advocates agreed with most of the proposal but felt that it did not go too far. More specifically, they wanted to see the new protections extend to *all* non-traditional loans. In support of this argument, consumer groups pointed to the narrow scope of the existing HOEPA regulations, which were only covering about five percent of all subprime loans.

Several of these groups, with the AARP leading the charge, called for a return to traditional underwriting standards¹¹⁷ that emphasized the three “C’s”: capacity, credit, and collateral. Capacity is a measure of whether the borrower is able to repay; credit is a measure of whether the borrower is likely to repay; and collateral requires an assessment of what assets the borrower has in case he does *not* repay.¹¹⁸ Along these lines, consumer groups favored requiring more specific standards in assessing ability to repay, including requiring the creditor or broker to consider W2 forms, payroll receipts, and other concrete measures of income verification.

The consumer groups differed radically from industry insiders on their treatment of prepayment penalties. Not all of the consumer groups recommended banning them outright, but most at least favored strict restrictions so that lenders could not penalize borrowers who wanted to refinance their mortgages once the rates exploded or if they were able to obtain a prime loan.

Among these groups, the escrow requirement was also popular.¹¹⁹ Many low- to moderate-income families may need cash more than their moderate- to high-income counterparts. This characteristic not only makes them more susceptible to advertising from predatory brokers and lenders, but it makes them less likely to be able to save for expenses such as property taxes and insurance. However, without paying these costs the borrowers are much more likely to fall victim to foreclosure.

Consumer groups also approved of the scope of the new disclosure requirements, although many noted that they, alone, were not sufficient. A few saw it as a trap to bait and switch because lenders are not required to disclose some applicable points and fees until the closing, at which time

¹¹⁷ See AARP, *supra* note 42, at 1, 4, 11.

¹¹⁸ *Id.*

¹¹⁹ *Id.* at 7-8.

many borrowers would be less likely to withdraw their application and shop for a new lender.¹²⁰

Additionally, in the proposed requirements, the Fed contemplated an outright ban on yield-spread premiums, which was applauded by consumer groups.¹²¹

These groups also were strongly opposed to the “pattern or practice” requirement, which allowed a borrower to pursue legal recourse if he could prove a pattern or practice of making loans without verifying income or assets.¹²² The criticism of this requirement was that it would essentially preclude borrowers from relief as it was difficult to establish this pattern. Furthermore, the consumer groups argued that it was counterintuitive to require would-be plaintiffs to prove a pattern or practice because by the time a pattern had been established, many borrowers would have been hurt, whereas if borrowers were allowed to sue based only on their own injuries from the lender

Consumer groups, and the NCLC in particular, were wary of a loophole that allowed lenders to continue to make no-documentation loans without liability as long as the originator’s loan decision would not have been different if the proper information *had* been available.¹²³ The group was concerned that this would serve as an incentive for originators to avoid proper underwriting techniques.

The Fed also received comments from industry interest groups, including the Mortgage Bankers Association, the American Securitization Forum (ASF), the American Community Bankers, and local real estate and lending groups. These groups, of course, viewed the proposed regulations differently than did consumer advocacy groups. They were primarily concerned with increased exposure to “extreme civil liability”¹²⁴ and an ensuing reluctance to extend credit to subprime borrowers.¹²⁵

¹²⁰ NATIONAL CONSUMER LAW CENTER, INC., THE FRB’S FINAL HOEPA RULE: A FIRST STEP, BUT REAL REFORMS ARE STILL NEEDED 2 (2008), *available at* http://www.consumerlaw.org/issues/predatory_mortgage/content/FRB-HOEPA-Rule-NCLCquickanalysis.pdf.

¹²¹ Center For Responsible Lending, *supra* note 106, at 4, 22.

¹²² *Id.* at 4.

¹²³ National Consumer Law Center, Inc., *supra* note 120, at 4.

¹²⁴ CONSUMER BANKERS’ ASSOCIATION, FED ISSUES TOUGH FINAL HOEPA RULES; DROPS BROKER DISCLOSURE 1 (2008), at 1. <http://www.cbanet.org/files/GRFiles/HOEpfinal.pdf>.

¹²⁵ Letter from the American Securitization Forum & Securities Industry & Financial Markets Association to the Board of Governors of the Federal Reserve

Like the consumer groups, industry groups generally offered little dissent regarding the disclosure requirements. This is likely because disclosure poses a relatively low burden to the lender, and as such there are very few legitimate arguments that brokers and lenders could make in opposition to these requirements.

Another concern was that the proposed regulations were too broad and would encompass prime as well as subprime loans, thereby harming an already fragile mortgage market by restricting lenders' willingness to lend.¹²⁶

Additionally, industry groups did not favor a prohibition on the yield-spread premium. Some defended these kickbacks on the basis that they a way for creditors and originators to compete for the best brokers to sell their loans.

Interestingly, from a general standpoint the consumer groups were less critical of the proposal than were the consumer advocate groups.¹²⁷

c. *The Fed's Final Rule*

After accepting comments, on July 30, 2008, the Fed published the final rule amending Regulation Z implementing HOEPA and TILA.¹²⁸ There were some notable changes from the proposed rule. First, at the behest of consumer groups like the CRL, the Fed changed the definition of "high-cost" in the context of mortgage loans.¹²⁹ In the proposed rules, whether a loan was defined as "high-cost" was determined by comparing that loan's rate to the yield on Treasury securities of comparable maturity. The CRL pointed out that a mortgage-based trigger was more appropriate, to account for economy-wide wide credit events. As a result, under the final rules the Fed will publish an "average prime offer rate based on a survey currently published by Freddie Mac. A loan is high-cost if it is a first loan and is at a rate that is 1.5 percentage points or more above the prime offer rate, or for subordinate-lien mortgages, if it is 3.5 percentage points above the prime offer rate."¹³⁰

System (Apr. 8, 2008), available at <http://www.americansecuritization.com/uploadedfiles/SIFMAASFRegZComments040808.pdf>

¹²⁶ *Id.*, at 2.

¹²⁷ *See id.*

¹²⁸ Truth in Lending (Regulation Z), 73 Fed. Reg. 44522 (Jan. 9, 2008) (to be codified at 12 C.F.R. 226).

¹²⁹ *Id.* at 44531.

¹³⁰ *Id.*

There were key protections passed specifically for higher-priced mortgage loans. The first requires lenders to verify repayment ability based on the highest scheduled payment in the first seven years, considering income, assets, and other debt obligations in the process. Lenders would be required to actually verify these facts as well. Because groups like Radian, Genworth, and MGIC have all been plagued by low- or no-documentation loans and have subsequently refused to insure them, the requirement that lenders verify a consumer's ability to repay is likely to qualify more subprime borrowers for loans under the new underwriting criteria.

Another protection for high-cost loans, representing a compromise between consumer groups and the industry groups, generally permits prepayment penalties, but *not* if the payment could change within the first four years.¹³¹ For other high-cost loans, the prepayment penalty period cannot last longer than two years.¹³² This also strikes a balance between paternalistic measures which would ban them altogether and promoting free choice by letting borrowers compare the rates on loans with penalties versus the ones on loans without in that it bans the most onerous penalties but permits others. It is however, substantially more restrictive than originally proposed.¹³³

Lastly, for high-cost mortgages the Fed kept the escrow requirement from the proposed rules largely intact. However, acknowledging the argument by the mortgage industry about the costs of implementation, instead of becoming effective in October 2009 like the rest of the provisions, the escrow requirement will take effect in 2010.¹³⁴

Provisions were included, such as those recommended by CRL that the regulations govern *all* mortgage loans, whether high-cost or not.. The Fed left intact the rules as proposed prohibiting abusive service practices¹³⁵ and coercion of appraisers to misstate home values.¹³⁶ The requirement to provide a good faith estimate of loan costs, including a payment schedule, within three days of receiving a consumer's application also remained

¹³¹ *Id.* at 44551.

¹³² *Id.*

¹³³ *See id.*

¹³⁴ Truth in Lending (Regulation Z), 73 Fed. Reg. 44522, 44523, 44595 (Jan 9, 2008) (to be codified at 12 C.F.R. 226).

¹³⁵ *Id.* at 44598.

¹³⁶ *Id.* at 44604.

intact.¹³⁷ Consumers can only be charged fees after receiving early disclosure. Furthermore, all mortgages are subject to more stringent standards regarding advertising: more information is required on advertisements, and lenders cannot advertise a “low fixed rate” if that rate in fact may suddenly explode to an adjustable rate after two or three years.¹³⁸

Notably, the Fed withdrew for further consideration its original proposal to prohibit brokers from receiving kickbacks in the form of yield-spread premiums. It also withdrew for further consideration more stringent disclosure requirements.¹³⁹ However, the Fed left intact a loophole allowing for the originators of no-documentation loans to escape liability if their decision would not have changed even *with* proper documentation.¹⁴⁰

2. H.R. 3915, or the Mortgage Reform and Anti-Predatory Lending Act of 2007

H.R. 3915, or the Mortgage Reform and Anti-Predatory Lending Act of 2007, would have filled most of the gaps left by the Fed’s proposed reform measures.¹⁴¹ However, although passed by the House, the Senate did not pass the Act which was thus cleared from the books at the end of 110th session of Congress. If it reintroduced and passed, the Act would represent a major step toward mortgage reform, even stronger than that taken by the Fed.

H.R. 3915 had a goal similar to that of the Fed’s reform measures, but added some key protections. One of the most important additional protections was a new duty of care imposed on all mortgage originators, including brokers and lenders.¹⁴² This would encompass the subjective triggers that were lacking in the Fed’s reform regulations and would help add much-needed teeth to reform measures. A bill like this, with its stricter provisions and more adequate remedies, would help ensure that lenders cannot avoid penalties for abusive practices by structuring the terms so as to fall just outside a definitional trigger. Furthermore, SPCI would have been prohibited on any residential mortgage, a move that acknowledged its

¹³⁷ *Id.* at 44600-01.

¹³⁸ *See id.* at 4457495.

¹³⁹ *Id.* at 44563.

¹⁴⁰ *See id.* at 44574.

¹⁴¹ *Id.* § 129A.

¹⁴² *Id.* at § 206(g).

inherent unfairness.¹⁴³ The bill would have also prohibited yield-spread premiums and other forms of steering incentives,¹⁴⁴ as well as unfair servicing practices such as refusing to credit a payment the day it was received and then applying a late fee.¹⁴⁵ Mortgage brokers and lenders would have been required to be licensed and registered.¹⁴⁶

The proposal also required creditors to verify ability to repay by considering several factors, including credit history, current and expected income. The scope of covered mortgages was also increased to cover open-end loans, whereas the Fed's regulations cover only closed-end loans. The triggers are based on the yield for Treasury securities of comparable maturities. There were special, additional protections for high-cost mortgages,¹⁴⁷ including a prohibition on balloon payments, recommending or encouraging default, excessive late fees, financing any points or fees abusive modification and deferral, and other abusive practices.¹⁴⁸ The borrower of such loans would also have been required to go through pre-loan counseling to ensure that he can actually understand and interpret all of the information that he would receive in required pre-closing disclosures.¹⁴⁹

IV. CONCLUSION: IS IT ENOUGH?

The recent subprime lending crisis finds its origins in predatory lending. Whereas legitimate subprime lenders help countless Americans with less-than-perfect credit get homes, predatory lending serves no function other than to pad the pockets of unscrupulous lenders. As the subprime crisis has spiraled out of control, two problems in mortgage insurance have emerged: the abuse of single-premium credit insurance and the issues plaguing the private mortgage insurance providers as a result of

¹⁴³ *Id.* at § 123(b).

¹⁴⁴ *Id.* at §§ 601-604. Steering incentives are defined as originator compensation that varies, directly or indirectly, or is based on the terms of any loan that is not a qualified mortgage as defined in § 129B(c)(3). *Id.*

¹⁴⁵ *Id.* at §§ 101-113.

¹⁴⁶ *Id.* at §§ 301(aa)(1).

¹⁴⁷ *Id.* at § 303. "Extremely" high-cost mortgages would be defined as those with very high points and fees, exceeding 5% of the total loan amount, using a comprehensive definition of points and fees that would include yield-spread premiums, and prepayment penalties. *Id.* at § 301(aa)(1).

¹⁴⁸ *Id.* at § 303(t).

¹⁴⁹ *Id.* at § 303(t).

the crisis. Both of these phenomena have proved a boon for predatory lenders.

Single premium credit insurance gets financed into the principal of the loan, which translates into earnings in interest for the loan originator as well as a large lump sum. The borrower's benefits are not nearly as appealing; usually, the insurance policy expires long before the premium is paid off. Any effective reforms must ban the use of this; disclosure alone is not enough.¹⁵⁰ At the very least, an intermediate step would be to prohibit financing the premium; that is, prohibit lenders from charging interest on the large premium. This will also require lenders to turn increasingly to PMI if they want extra security for loans on which borrowers are paying less than 20 per cent down, thereby shifting borrowers into a the much safer and better-regulated world of PMI.

As predatory lenders became increasingly aggressive, they set their sights on borrowers who wanted to avoid purchasing PMI when they had less than 20 per cent to put down on a home. Demand fell for PMI policies and as a result, predatory loan solutions for the problem of the borrower with little cash to put down rose. Addressing predatory lending and restoring PMI to a legitimate, feasible option for borrowers are symbiotic propositions. The best chance of accomplishing both of these objectives is for the federal government to use its authority to effectuate sweeping reforms.

Had the Senate passed H.R. 3915 and the bill had been signed into law and enacted, it would have represented a major step against predatory lending that encompassed nearly all abusive practices and also provided sufficient penalties and remedies to enforce them. As it is, however, the Fed's regulations alone are not enough to effectuate the changes needed in mortgage-related insurance products. What is needed are not strong disclosure requirements, because even if disclosed, the abusive practices tend to be complicated and hidden within an even more complicated framework of loan terms.¹⁵¹ Whereas disclosure may be sufficient for

¹⁵⁰ See generally Patricia A. McCoy, *Rethinking Disclosure in a World of Risk-Based Pricing*, 44 HARV. J. ON LEGIS. 123 at 147, 154 (2007).

¹⁵¹ See Elizabeth Renuart & Diane E. Thompson, *The Truth, the Whole Truth, and Nothing But the Truth: Fulfilling the Promise of Truth in Lending*, 25 YALE J. ON REG. 181 (2008) (explaining that evaluating the cost of credit and comparison shopping is very difficult even for sophisticated shoppers, but as lenders increasingly "unbundle the costs of their loans from the interest into an array of fees, outsource their overhead to third parties who add to consumers' costs, and unveil amazingly complex loan products that dazzle and confuse borrowers," even

comparison-shopping in the prime market, disclosure is no longer sufficient as the primary regulator of the subprime credit marketplace.¹⁵² Rather, the abusive practices must be banned outright. This was well-illustrated in *Cetto v. LaSalle Bank National Association*, debt obligations, and assets, among others.¹⁵³ Prepayment penalties would be prohibited on subprime loans, but permitted on other mortgages provided they expired three months before a loan resets. They would have been forbidden entirely on any subprime loan.¹⁵⁴ Because there is no pattern or practice requirement, a lender could have become liable based only on one mortgage. Therefore, there was an allowance for bona fide errors within thirty days of the loan closing, provided the lender corrected them.

The scope of covered mortgages was also increased to cover open-end loans, whereas the Fed's regulations cover only closed-end loans. The triggers are based on the yield for Treasury securities of comparable maturities. There were special, additional protections for high-cost mortgages,¹⁵⁵ including a prohibition on balloon payments, recommending or encouraging default, excessive late fees, financing any points or fees abusive modification and deferral, and other abusive practices.¹⁵⁶ The borrower of such loans would also have been required to go through pre-loan counseling to ensure that he can actually understand and interpret all of the information that he would receive in required pre-closing disclosures.¹⁵⁷

sophisticated shoppers would be hard-pressed to draw any sort of meaningful comparison).

¹⁵² *Id.*

¹⁵³ H.R. 3915, 110th Cong. at § 129B(a) (2007).

¹⁵⁴ *Id.* at 206(f).

¹⁵⁵ *Id.* at § 301(aa)(1). "Extremely" high-cost mortgages would be defined as those with very high points and fees, exceeding 5% of the total loan amount, using a comprehensive definition of points and fees that would include yield-spread premiums, and prepayment penalties. *Id.* at § 301(aa)(1).

¹⁵⁶ *Id.* at § 303.

¹⁵⁷ *Id.* at § 303(t).

**THE 2008 MENTAL HEALTH PARITY
AND ADDICTION EQUITY ACT:
AN OVERVIEW OF THE NEW LEGISLATION AND
WHY AN AMENDMENT SHOULD BE PASSED TO
SPECIFICALLY DEFINE MENTAL ILLNESS AND
SUBSTANCE USE DISORDERS**

*Sara Nadim**

This note examines the 2008 Mental Health Parity and Addiction Equity Act and argues that even though this Act represents a landmark improvement in mental illness parity coverage, an amendment should be passed to define what is specifically considered to be a mental illness or substance use disorder. The first part explores the history of federal mental parity law along with the efforts made to achieve parity. The second part discusses the specific provisions of the 2008 Act, specifically that it does not provide explicit definitions for mental health conditions or substance use disorders. Third, state definitions of mental illness are reviewed. Recent developments supporting the biological basis of mental illness are presented in the fourth part. Finally, the fifth part of the note evaluates the diminished societal costs that will come with the addition of parity in mental illness insurance coverage. This note argues that certain severe biologically based mental illnesses should be listed under the definition of mental illness that at minimum insurers should be required to cover. It supports its proposition by providing evidence that group health plan costs for employers will not increase greatly, and that societal costs, such as homelessness and loss of productivity in the workplace, will be greatly reduced when mental illness and substance use disorders are adequately treated.

* University of Connecticut School of Law, Juris Doctor Candidate for the Class of 2010.

I. INTRODUCTION

For decades mental health advocates have fought for insurance coverage parity for mental illnesses. On October 3, 2008 the Paul Wellstone and Pete Domenici Mental Health Parity and Addiction Equity Act of 2008 was signed in to law, bringing to fruition decades of advocacy work on behalf of the mentally ill. Over fifty-seven million, or one in four Americans, currently suffer from a mental illness.¹ In the words of the Speaker of the House, Nancy Pelosi:

This long-overdue legislation has brought mental illness and addiction out of the shadows and to the forefront of our work here in Congress. By requiring that illness in the brain be treated just like illness anywhere else in the body for insurance purposes, we are helping to end discrimination against those who seek treatment for mental illness and saving lives.²

While this bill provides positive advancement for mental illness coverage, the lack of a clear definition for what is considered a mental illness or substance use disorder will result in inequitable coverage for many individuals. There are many reasons why mental parity proponents were able to finally achieve the passage of this bill. The reasons have ranged from: (1) proof that group health plan costs would not become exorbitant; (2) a greater acceptance of and a reduction in stigma surrounding mental illness; (3) the recent evidence of a biological basis of mental illness; and (4) recognition of the enormous societal costs of not treating the mentally ill.

This note will examine the new bill and argue that while the bill provides significant advancements in mental illness parity coverage, an amendment should be passed to define what is specifically considered to be a mental illness or substance use disorder. Two reasons why this bill was able to garner the support necessary for passage in both the House and Senate was that it was proven that the cost of group health plans would not increase, as well as evidence that many mental illnesses have a biological

¹ NAT'L INST. OF MENTAL HEALTH, HEALTH & OUTREACH, STATISTICS (2009), <http://www.nimh.nih.gov/health/statistics/index.shtml>.

² Press Release, Nancy Pelosi, Speaker of the House, House of Representatives, Pelosi: Passage of Mental Health Parity Bill Gives Hope and Help to Millions of American Families (Sept. 23, 2008), <http://speaker.house.gov/newsroom/pressreleases?id=0831>.

basis.³ This note suggests that both of these reasons also support the advancement of an amendment that would provide a specific list of severe mental illnesses to be covered instead of leaving it up to the states and group health plans to determine.

While the House of Representatives and a handful of states have advocated or supported a definition of mental illness that includes all of the diseases listed in the Diagnostic and Statistical Manual of Mental Disorders (DSM), published by the American Psychiatric Association, this note does not advocate for such an approach. Rather, this note suggests that certain severe biologically based mental illnesses, specifically those based on the definition advocated by Paul Wellstone and Pete Domenic in 1999, should be listed under the definition of mental illness that, at a minimum, insurers should be required to cover. This proposition is supported by evidence that group health plan costs for employers will not increase greatly and that societal costs, such as homelessness and loss of productivity in the workplace, will be greatly reduced when mental illness and substance use disorders are adequately treated.

Part I will discuss the history of federal mental parity law and the efforts to achieve parity; Part II will discuss the specific provisions of the Paul Wellstone and Pete Domenici Mental Health Parity and Addiction Equity Act of 2008 and the fact that the bill does not provide explicit definitions for mental health conditions or substance use disorders; Part III will discuss how the states define mental illness; Part IV will cover recent developments supporting the biological basis of mental illness; and Part V will discuss the diminished societal costs that will come with the addition of parity in mental illness insurance coverage. The argument that coverage of specific illnesses will be cost prohibitive is not sound. Certain illnesses have been found to have a biological basis and, at a minimum, a specific list of illnesses should be included in the parity legislation in order to ensure that individuals throughout the country receive equal treatment and coverage.

³ See *infra* Parts IV, V.

II. THE HISTORY OF FEDERAL MENTAL HEALTH PARITY LAW

A. EARLY EFFORTS TO ACHIEVE PARITY

In the 1990's, the majority of employer-sponsored health plans that did include mental health services placed far greater restrictions on mental health services than for other medical services.⁴ In 1998, sixty-two percent of health plans imposed limits on inpatient treatment for mental health services and fifty-seven percent imposed limits on outpatient treatment.⁵ These limits were imposed purely on mental health services and typically were not placed on other medical services.⁶

In 1996, Paul Wellstone (D-MN) and Peter Domenici (R-MN) introduced the Mental Health Parity Act (MHPA) as an amendment to the Kassebaum-Kennedy bill for healthcare portability.⁷ The Health Insurance Portability and Accountability Act of 1996 (HIPAA) established standards for the privacy and security of health information, as well as standards for electronic data interchange (EDI) of health information.⁸ The proposed MHPA amendment was passed by the Senate; however, it was met with objections in the House.⁹ The concerns raised in the House included whether the enactment of such an amendment would result in an increase in premiums for private health plans and, if that were the case, whether it would be necessary to provide for an amendment which would only

⁴ Dana L. Kaplan, *Can Legislation Alone Solve America's Mental Health Dilemma? Current State Legislative Schemes Cannot Achieve Mental Health Parity*, 8 QUINNIPIAC HEALTH L.J. 325, 329 (2005). "Ninety-one percent of small firms and 99% of large firms offer mental health and substance abuse coverage in their most used medical plans." *Id.* at n.26. These medical plans restricted one or all of the following for mental illness treatments: inpatient day limitations, office visit limitations, annual and/or lifetime maximums, or higher deductibles and co-payment rates. *Id.*

⁵ *Id.* at 329.

⁶ *See id.*

⁷ Vallerie Propper & Ginger L. Pomiecko, *Parity for Mental Health: History and Consequences*, in PUBLIC HEALTH MANAGEMENT & POLICY (2009), <http://www.case.edu/med/epidbio/mphp439> (last visited Nov. 21, 2009).

⁸ Univ. of Miami, Miller Sch. of Med. *Health Insurance and Portability Act of 1996 (HIPAA)*, http://privacy.med.miami.edu/glossary/xd_hipaa.htm (last visited Jan. 25, 2008).

⁹ 142 CONG. REC. H9473-9564 (1996).

increase the coverage of mental health services under health plans without increasing such premiums.¹⁰ In the end, Senator Kassebaum and Kennedy decided to remove the amendment in order to pass their bill more swiftly.¹¹

Both of these men had personal experiences with mental illness, driving them to further advocate for this bill. Senator Domenici's daughter suffers from schizophrenia and Senator Wellstone's brother suffers from bipolar disease. Wellstone and Domenici once again attempted to gain the passage of the MHPA when they attached the amendment to the Employee Retirement Income Security Act of 1974 and to the Public Health Services Act.¹² This time the supporters of the amendment threatened to filibuster if the amendment was removed, as it had been in 1996.¹³

B. THE MENTAL HEALTH PARITY ACT OF 1998

The MHPA amendment went into effect on January 1, 1998.¹⁴ The amendment that was implemented reflected a compromise between the proponents and opponents of the bill, with the understanding and expectation that Congress would reach a more comprehensive agreement within six years.¹⁵ The MHPA included a sunset provision stating that "this section shall not apply to benefits for services furnished on or after September 30, 2001" by group health plans or health insurance coverage offered in connection with such a plan.¹⁶ However, Congress decided to extend the sunset provision every year since 2001, and a more comprehensive agreement was not reached until 2008 with the passage of the Mental Health Parity and Addiction Equity Act.¹⁷

¹⁰ *Id.*

¹¹ *See id.*

¹² Mental Health Parity Act of 1996, Pub. L. No. 104-204, § 702, 110 Stat. 2944 (current versions at 42 U.S.C. § 300gg-5 (2006); 29 U.S.C. § 1185a (2006)).

¹³ *See* Propper & Pomiecko, *supra* note 7.

¹⁴ 29 U.S.C. § 1185a (Supp. II 1996) (current version at 29 U.S.C. § 1185a (2006)).

¹⁵ Glen Cheng, Note, *Caring for New Jersey's Children with Autism: A Multifaceted Struggle for Parity*, 60 RUTGERS L. REV. 997, 1016 (2008).

¹⁶ 42 U.S.C. § 300gg-5(f) (Supp. II 1996) (repealed 2008).

¹⁷ *See* Cheng, *supra* note 15, at 1016 & n.135. "Public Law 107-116, § 701(a) extended the sunset to December 31, 2002. Public Law 107-313, § 2(a) extended the sunset to December 31, 2003. Public Law 108-197, § 2(a) extended the sunset to December 31, 2004. Public Law 108-311, § 302(c) extended the sunset to December 31, 2005." *Id.* at n.135.

The original MHPA legislation mandated generous mental health insurance benefits, although unfortunately the final legislation did not reflect this mandate.¹⁸ This was due to the fact that in order to satisfy the opposition groups and achieve some type of parity, several important measure that would have ensured comparable coverage between mental and health services and other medical services were abandoned.¹⁹ Due to the need for such compromises, the bill was viewed as only a step toward achieving full mental health parity.²⁰ Senator Domenici, a sponsor of the bill, stated that the legislation was “a compromise to begin down the path of parity and nondiscrimination for mentally ill people in this country who have health insurance.”²¹

The bill stated that if a group health plan does not include an aggregate lifetime limit on all medical and surgical benefits, the plan cannot impose any aggregate lifetime limit on mental health benefits.²² Additionally, if the plan does have an aggregate lifetime limit on medical and surgical benefits it must apply the limit to mental health benefits and not distinguish between the two.²³ The bill also contained provisions articulating the same standards for group health plans with regards to annual limits.²⁴ If a plan does not include an annual limit on all medical and surgical benefits, the plan may not impose any annual limits on mental health benefits.²⁵ One positive provision of the bill stated that for insurance plans that were subject to state laws, the MHPA did not preempt state laws that require more favorable treatment of mental health benefits.²⁶ However, this bill included many limitations that diminished the scope and force of the positive mandates.

The bill did not require a group health plan to provide any mental health benefits. If a group health plan did provide mental health benefits, the provisions of the bill did not affect the terms and conditions of the coverage such as limits on outpatient visits, in-patient day limits, days of

¹⁸ See Kaplan, *supra* note 4, at 330.

¹⁹ *Id.*

²⁰ *Id.* at 342.

²¹ *Id.*

²² 29 U.S.C. § 1185a(a)(1)(A) (Supp. II 1996) (current version at 29 U.S.C. § 1185a (2006)).

²³ *Id.* § 1185a(a)(1)(B).

²⁴ *Id.* § 1185a(a)(2)(B).

²⁵ *Id.* § 1185a(a)(2)(A).

²⁶ Michael J. Carroll, Note, *The Mental Health Parity Act of 1996: Let It Sunset if Real Changes Are Not Made*, 52 DRAKE L. REV. 553, 557-58 (2004).

coverage, deductibles, prior authorization requirements, requirements relating to medical necessity, or, in the case of a managed care plan, a primary care physician's referral requirement.²⁷ And lastly, if the application of the bill to the plan resulted in an increase in cost of at least one percent, the group health plan was exempt.²⁸ This last provision was included to alleviate fears that the bill would increase costs to a prohibitive level.

C. SUBSTANCE USE DISORDERS

The bill also did not include coverage for substance abuse treatment. Many found this to be illogical due to the fact that many people who suffer from mental illnesses also experience a co-occurring substance abuse problem.²⁹ The co-occurrence or dual diagnosis is often known as comorbidity.³⁰ Treatment for substance use disorders is a critical element in an individual's treatment for a mental disorder; similarly, treatment for a mental illness is a critical element in the recovery of a person with a substance use disorder.³¹ Often times the two are intertwined and recovery from either disorder is dependent upon the other.³² Treatment of each disorder separately has proven to be ineffective and research supports treatment that addresses both conditions.³³ A successful model of such treatment includes case management, group interventions, and assertive outreach to bring people into treatment.³⁴

The *Journal of the American Medical Association* has found that roughly fifty percent of individuals with severe mental disorders are affected by substance use disorders; thirty-seven percent of alcohol abusers and fifty-three percent of drug abusers also have at least one serious mental

²⁷ *Id.* at 557.

²⁸ *Id.* at 561.

²⁹ E.g., Charity Felts, Comment, *Dealing with a Depressed Workforce: Are American Employers Doing Enough to Support the Mental Health Challenges Affecting Today's Employees?* 9 SCHOLAR 119, 129 (2006).

³⁰ *Id.* at 129-30 & n. 64.

³¹ *Id.* at 129-30.

³² *Id.*

³³ *See id.*

³⁴ NAT'L ALLIANCE ON MENTAL ILLNESS, DUAL DIAGNOSIS AND INTEGRATED TREATMENT OF MENTAL ILLNESS AND SUBSTANCE ABUSE DISORDERS (2009), http://www.nami.org/Template.cfm?Section=By_Illness&Template=/TaggedPage/TaggedPageDisplay.cfm&TPLID=54&ContentID=23049.

illness; and of all people diagnosed as mentally ill, twenty-nine percent abuse either alcohol or drugs.³⁵ An Epidemiologic Catchment Area Survey found that individuals with severe mental disorders were at significant risk for developing a substance use disorder during their lifetime.³⁶ The research specifically found that forty-seven percent of individuals with schizophrenia had a substance use disorder, more than four times greater than the general population, and that sixty-one percent of individuals with bipolar disorder also had a substance use disorder, which is five times greater than the general population.³⁷

It is clear that in order to provide effective treatment to the mentally ill, substance use disorders must also be treated. However, mandated coverage for substance use disorders was not implemented until the Paul Wellstone and Pete Domenici Mental Health Parity and Addiction Equity Act of 2008 was signed in to law.

D. THE ROAD TO THE MENTAL HEALTH PARITY AND ADDICTION EQUITY ACT OF 2008

Due to the compromises that had to be made in the 1998 parity legislation, Senators Domenici and Wellstone came back in 2001 with a new bill in an effort to achieve full mental health parity. This time they were joined by Patrick Kennedy (R- RI) and Jim Ramstad (R-MN), both of whom had struggled with addiction issues.³⁸ The bill had begun to pick up momentum when Wellstone was killed in a plane crash in 2002.³⁹ After this occurred, his son David began to advocate for the measure.⁴⁰ The accumulation of the effort to pass a comprehensive mental parity bill was achieved this past year when the Mental Health Parity and Addiction Equity Act was passed under the 2008 economic stimulus package. One reason the bill was able to achieve passage was that just before the start of the current Congressional session, several mental health groups won the support of employers and health insurers by alleviating concerns over the

³⁵ *Id.*

³⁶ *Id.*

³⁷ *Id.*

³⁸ Julie Rovner, *Mental Health Parity Approved with Bailout Bill*, NAT'L PUB. RADIO (2008), <http://www.npr.org/templates/story/story.php?storyId=95435676>.

³⁹ *Id.*

⁴⁰ *Id.*

cost of mental parity.⁴¹ A 2006 study in the *New England Journal of Medicine* found that insurers' costs rose less than half a percentage point when full parity was required for federal workers starting in 2001.⁴² The Congressional Budget Office Cost Estimate also stated that if the more generous House bill were enacted, the costs for premiums would increase for group health insurance by an average of only about 0.4 percent.⁴³

Two bills were originally formulated in the House and Senate. Through a series of compromises, the two sides were able to arrive at an agreement. "A breakthrough occurred when sponsors of the House bill agreed to drop a provision that required insurers to cover treatment for any condition listed in the Diagnostic and Statistical Manual of Mental Disorders (DSM), published by the American Psychiatric Association."⁴⁴ The Senate also made concessions in the negotiation agreements and agreed to adopt some of the language in the House bill that required parity for out-of-network coverage.⁴⁵

Other reasons that such a comprehensive mental parity bill gained widespread support included the fact that new scientific research had revealed a biological basis and effective medical treatments for numerous mental illnesses, as well as evidence suggesting that providing mental health parity coverage will not be cost prohibitive. Additionally, employers have realized that productivity tends to decrease when workers are not treated for mental illnesses and substance use disorders, and that if they do receive treatment it can reduce the number of lost work days.⁴⁶ Furthermore, some argue that the stigma of mental illness may have faded as society has seen many members of the armed forces returning from the

⁴¹ *Id.*

⁴² Nancy Shute, *Paying a High Price for Mental Health*, U.S. NEWS, Oct. 25, 2007, available at <http://health.usnews.com/articles/health/health-plans/2007/10/25/paying-a-high-price-for-mental-health.html>.

⁴³ CONGRESSIONAL BUDGET OFFICE COST ESTIMATE (2007), available at <http://www.cbo.gov/ftpdocs/88xx/doc8837/hr1424ec.pdf>.

⁴⁴ Robert Pear, *Bailout Provides More Mental Health Coverage*, N.Y. TIMES, Oct. 5, 2008, available at <http://www.nytimes.com/2008/10/06/washington/06mental.html>.

⁴⁵ Frederic J. Frommer, *Wellstone Son Launches Push on Mental Health Bill*, BOSTON GLOBE, July 9, 2008, available at http://www.boston.com/news/local/rhode_island/articles/2008/07/09/wellstone_son_launches_push_on_mental_health_bill/.

⁴⁶ See Pear, *supra* note 44.

Middle East with serious mental health issues.⁴⁷ And lastly, the experimentation with parity at both the state level and in the health insurance program for federal employees, including members of Congress, has proved workable.⁴⁸

III. THE PAUL WELLSTONE AND PETE DOMENICI MENTAL HEALTH PARITY AND ADDICTION EQUITY ACT OF 2008

The Mental Health Parity and Addiction Equity Act (MHPAEA) was signed into law on October 3, 2008. The bill is an amendment to Section 712 of the Employee Retirement Income Security Act of 1974 (29 U.S.C. 1185a).⁴⁹ This bill will provide over one-third of Americans with improved mental health insurance coverage.⁵⁰ It will also eliminate a practice that has been in place for decades, where insurers have placed much higher co-payments and deductibles on treatment for mental illness. The new bill also eradicates the practice of restricting the number of inpatient hospital treatment and outpatient visits for mental health treatment.⁵¹ Federal officials have stated that the new law will improve coverage for 113 million people.⁵² This figure includes 82 million people who are in employer-sponsored plans that are not subject to state regulations.⁵³ The effective date, for most health plans, will be January 1, 2010.⁵⁴

A. SPECIFIC PROVISIONS OF THE MHPAEA

The new bill provides parity between medical and surgical benefits and mental health or substance use disorders for all “deductibles, copayments, coinsurance, and out-of-pocket expenses,” specifying that there are to be no separate cost sharing requirements that apply only to mental health or substance abuse disorder benefits.⁵⁵ The bill also states that there shall be no difference between treatment limitations for mental

⁴⁷ *Id.*

⁴⁸ *Id.*

⁴⁹ H.R. 1424, 110th Cong. (2008).

⁵⁰ Pear, *supra* note 44.

⁵¹ *See infra* Part II.A.

⁵² Pear, *supra* note 44.

⁵³ *Id.*

⁵⁴ *Id.*

⁵⁵ 42 U.S.C.A. § 300gg-5(a)(3) (West Supp. 2009).

health or substance use disorder benefits and medical and surgical benefits.⁵⁶ Treatment limitations include “limits on the frequency of treatment, number of visits, days of coverage, or other similar limits on the scope or duration of treatment.”⁵⁷ The bill also states that if a plan or coverage provides medical or surgical benefits by out-of-network providers, the plan must afford the same for mental health or substance use disorders.⁵⁸

The bill also has a few exceptions, including a cost exemption. The bill states that an employer is exempt from the parity requirements if the overall implementation of the bill would result in an increased cost of two percent or more during the first year after the legislation goes into effect and one percent in the following years.⁵⁹ Furthermore, employers with fifty employees or less are exempt from the parity requirements.⁶⁰ This affects the almost 113 million American employees who work at companies with fewer than fifty employees.⁶¹

The new bill also provides specific guidelines which articulate how an employer will qualify for the cost exemption. To qualify for the exemption, the plan must implement the new requirements for at least six months.⁶² If after six months the employer can show that the implementation of the bill results in a cost increase of one percent, the plan must give notification to the participants and beneficiaries of its decision to claim the exemption to the health benefit plan.⁶³ The exemption will not go into effect until thirty days after the notice requirements are fulfilled.⁶⁴ If the employer fails to follow these requirements they will be subject to a tax penalty and fined \$100 per day per individual who is affected by such a failure.⁶⁵

⁵⁶ *Id.* § 300gg-5(a)(3)(A)(ii).

⁵⁷ *Id.* § 300gg-5(a)(3)(B)(iii).

⁵⁸ *Id.* § 300gg-5(a)(5).

⁵⁹ *Id.* § 300gg-5(c)(2)(A)-(B).

⁶⁰ H.R. 1424, 110th Cong. (2008).

⁶¹ LONG-AWAITED BREAKTHROUGH ON FEDERAL MENTAL HEALTH PARITY (2007), <http://www.mhafc.org/pdf/FedParity2007.pdf>.

⁶² Peter M. Panken et al., *Employment and Labor Relations Law for the Corporate Counsel and the General Practitioner, Litigating Claims of Employee Benefits*, SN020 A.L.I.-A.B.A. 169, 179 (2008).

⁶³ *Id.*

⁶⁴ *Id.*

⁶⁵ *Id.*

It is not clear at this time how many companies will qualify for this exemption. The states that have experimented with parity have found that costs often did not rise to the one percent exemption level and frequently costs stayed the same or decreased.⁶⁶ For example, when Texas implemented parity for severe mental illnesses and substance use disorders, a study found that there was a decrease of fifty percent in per-member, per-person cost.⁶⁷ Managed care was also introduced at the same time.⁶⁸ Similar results were found in North Carolina, and a study on the impact of mental health parity in California revealed that costs did not increase after one year.⁶⁹

The bill further states that insurers must publish the criteria for medical necessity determinations.⁷⁰ The insurer must also provide an explanation for any denial of a claim made for mental health services.⁷¹ However, the bill does not provide an explicit definition for mental health conditions or substance use disorders. The bill states that “the term ‘mental health conditions’ and ‘substance use disorders’ are defined under the terms of the group plans and in accordance with applicable Federal and State law.”⁷²

The House of Representatives had urged the inclusion of a provision that would have required insurers to provide coverage for any condition listed in the DSM; however this was met with strong opposition.⁷³ Many opposed the use of the DSM due to the fact that it contains conditions such as caffeine intoxication, sleep disorders and jetlag. In the end, the House dropped its requirement that all DMS-IV disorders be covered equitably.

Without a clear definition of what the federal government considers to be a mental health condition or substance use disorder, such determinations will vary widely due to the fact that such definitions differ greatly from state to state. This will result in a great variation in individual coverage depending on where people reside. Additionally, with no clear

⁶⁶ See Kate Mulligan, *More Data Confirm Affordability of Parity*, 37 PSYCHIATRIC NEWS 18 (2002), <http://pn.psychiatryonline.org/cgi/content/full/37/12/18>.

⁶⁷ *Id.*

⁶⁸ *Id.*

⁶⁹ *Id.*

⁷⁰ See Pear, *supra* note 44.

⁷¹ *Id.*

⁷² 42 U.S.C.A. § 300gg-5(e)(4) (West Supp. 2009).

⁷³ See Pear, *supra* note 44.

definitions of what is considered to be a mental illness or substance use disorder, the issue of whether something is medically necessary will continue to come into debate. Furthermore, the lack of categorization or definitions for mental health conditions and substance use disorders potentially leaves the door open for insurance providers to exclude illnesses more readily than if there were more stringent guidelines.

IV. HOW STATES DEFINE MENTAL ILLNESS

In the article *An Analysis of the Definitions of Mental Illness Used in State Parity Laws*, Marcia C. Peck, M.D., M.P.H., and Richard M. Scheffler, Ph.D. analyzed the mental health parity laws in thirty-four states.⁷⁴ The study analyzed the different definitions used by the states in defining mental illness and the effect these variations have on the coverage that individuals with a mental illness receive.⁷⁵ They found that three statutory terms were used to define mental illness in state parity legislation: (1) “broad-based mental illness,” (2) “serious mental illness,” or (3) “biologically based mental illness.”⁷⁶ They further found that,

States rarely, if ever, considered disease prevalence, needs-based studies, and clinical judgment. In our opinion the definitions that states use result from a political and economic process involving mental health advocates and providers, pro- and antiparity legislators, insurers, and employers.⁷⁷

Generally, the majority of states have laws specifically enumerating what is a mental illness. These states base their mental illness definitions on “biologically based” mental illnesses and “serious” mental illnesses.⁷⁸ These definitions are most frequently included in parity legislation that outline the specific mental illnesses which must be covered by insurers on an equal basis with physical illnesses. The states that use the term “biologically based mental illness” as the statutory term used to define mental illness in their parity legislation are: Alabama, Colorado, Iowa,

⁷⁴ Marcia C. Peck, M.D., M.P.H. & Richard M. Scheffler, Ph.D., *An Analysis of the Definitions of Mental Illness Used in State Parity Laws*, 53 PSYCHIATRIC SERVICES 1089, 1089 (2002), <http://psychservices.psychiatryonline.org/cgi/content/full/53/9/1089>.

⁷⁵ *Id.*

⁷⁶ *Id.*

⁷⁷ *Id.* at 1091.

⁷⁸ *Id.* at 1089.

Massachusetts, Missouri, New Hampshire, New Jersey, South Carolina, South Dakota and Virginia.⁷⁹ The theoretical basis for this type of definition is biological psychiatry.⁸⁰ The mental illnesses that these states list have a scientifically demonstrable effect on the brain.⁸¹ The argument for covering these illnesses is that the brain is being damaged by the mental illness and therefore should be treated and covered like any other damaged or injured organ.⁸²

The states that use the term “serious mental illness” as the statutory term used to define mental illness in their parity legislation are: California, Kansas, West Virginia, Louisiana, Illinois, Pennsylvania, Delaware, Texas, Montana, Maine, Nebraska, Nevada, Oklahoma, and Hawaii.⁸³ The theoretical basis for this type of definition is public policy, not clinical.⁸⁴ The aim of this definition is to identify severe and persistent mental illnesses based on “functional disability and duration of [the] illness.”⁸⁵ While these states all define mental illness on the same theoretical basis, either as “biologically based mental illness” or “serious mental illness,” it is important to note that there is still a great variety among what mental illnesses are covered.

The states that define mental illness as “biologically based mental illness” all provide coverage for schizophrenia, major depressive disorder, obsessive-compulsive disorder, and bipolar disorder.⁸⁶ However, out of the ten states that define mental illness in that manner, only two (Virginia and New Jersey) define autism as a mental illness.⁸⁷ Only Massachusetts, Missouri, Virginia, and South Carolina define childhood depression as a

⁷⁹ JERRY CONNOLLY & JILL STRASSER, PH.D., DEFINING “MENTAL OR NERVOUS CONDITIONS,” 1st Sess., at 5 (Feb. 2006), http://www.cbs.state.or.us/ins/public_meetings/sb_minutes/attachments/SB1_defining-presentation.pdf.

⁸⁰ *Id.*

⁸¹ *Id.*

⁸² *See id.*

⁸³ *See* Peck & Scheffler, *supra* note 74, at 1092 tbl. 2.

⁸⁴ CONNOLLY, *supra* note 29, at 5.

⁸⁵ *Id.*

⁸⁶ *See* Peck & Scheffler, *supra* note 74, at 1093 tbl. 3.

⁸⁷ *Id.*; *see also*, IOWA CODE ANN. § 514c.22 (West 2007); N.J. STAT. ANN. § 17B:26-2.1s (West 2006); VA. CODE ANN. § 38.2-3412.1:01 (2007). On July 1, 2009, Massachusetts began to include autism in its definition of mental illness. MASS. GEN. LAWS ANN. ch. 176B § 4A (2007).

mental illness⁸⁸ and only Virginia defines attention deficient/hyperactive disorder (ADHD) as a mental illness.⁸⁹ Massachusetts, Missouri and New Hampshire are the only states in the group that consider post-traumatic stress disorder (PTSD) to be a mental illness.⁹⁰

Similar to the variation of definitions that exist between states that define mental illness as “biologically based,” there is great variation in the illnesses covered in states which define it as “serious mental illness.” For example, California’s definition includes five disorders: schizophrenia, schizo-affective disorder, bipolar disorders and delusional depressions, and pervasive developmental disorder.⁹¹ Maine’s definition, on the other hand, includes fourteen disorders: psychotic disorders, including schizophrenia, dissociative disorders, mood disorders, anxiety disorders, personality disorders, paraphilias, attention deficit and disruptive behavior disorders, pervasive developmental disorders, tic disorders, eating disorders, including bulimia and anorexia, and substance abuse-related disorders.⁹² These examples illustrate that simply because states have drafted their statutory definitions based on the same theoretical basis, it does not mean that the same illnesses will be covered.

As Doctors. Peck and Scheffler noted, in addition to “biologically based” definitions, some states use “broad-based” definitions of mental illness in their statutes.⁹³ “Broad-based” definitions are based upon the

⁸⁸ MASS. GEN. LAWS ANN. ch. 176B, § 4A(a) (West 2009); MO. ANN. STAT. § 376.826(4)(b),(d) (West 2002); S.C. CODE ANN. § 38-71-290 (1976); VA. CODE ANN. § 38.2-3412.1:01(E) (LexisNexis 2007).

⁸⁹ VA. CODE ANN. § 38.2-3412.1:01(E) (LexisNexis 2007). *See* MASS. GEN. LAWS ANN. ch. 176B, § 4A(c) (West 2009) which states that “[A]ny such subscription certificate shall also provide benefits on a non-discriminatory basis for children and adolescents under the age of 19 for the diagnosis and treatment of non-biologically-based mental, behavioral or emotional disorders, as described in the most recent edition of the DSM, which substantially interfere with or substantially limit the functioning and social interactions of such a child or adolescent.” This would include attention deficient/hyperactive disorder; however, it is not defined as a “biologically based” mental illness by Massachusetts.

⁹⁰ MASS. GEN. LAWS ANN. ch. 176B, § 4A(a) (West 2009); MO. ANN. STAT. § 376.826(4)(c) (West 2002); N.H. Rev. Stat. Ann. § 417-E:1(III)(i) (LexisNexis 2009).

⁹¹ CAL. INS. CODE § 10123.15 (West 2005).

⁹² ME. REV. STAT. ANN. tit. 24-A § 2843(5-C)(A), (5-D)(A) (2000).

⁹³ *See* Peck & Scheffler, *supra* note 74, at 1090, 1091 tbl.1.

DSM.⁹⁴ The use of a “broad-based” definition of mental illness was the position advocated by the House of Representatives during the debates over the current parity bill. An example of states that use “broad-based” definitions include: Connecticut, Kentucky, Utah, Rhode Island, and Washington.⁹⁵ These states undoubtedly provide greater coverage than do states that define mental illness according to a short list of “biologically based” or “serious” mental illnesses.

There are two major mental health advocacy groups in the United States, the National Association for the Mentally Ill (NAMI) and the National Mental Health Associations (NMHA). Each of these groups define mental illness differently. NAMI promotes ending discrimination and demands fair legislative policies for “priority populations with serious mental illness.” Priority populations include those with schizophrenia, schizoaffective disorder, major depressive disorder, obsessive-compulsive disorder, panic disorder and other severe anxiety disorders, and ADHD.⁹⁶ The NMHA defines mental illness broadly, addressing a person's ability to function rather than his or her diagnosis.⁹⁷

These different definitions can lead to different conclusions about what is a mental illness and therefore what is “medically necessary.” The MHPAEA states that the definition of mental disorders and substance use disorders should be in accordance with state and federal laws.⁹⁸ From the survey of many state statutes, it is clear that there is not a consensus on the state level of what is a mental illness.⁹⁹ It should be noted that this

⁹⁴ *Id.* at 1090.

⁹⁵ JERRY CONNOLLY & JILL STRASSER, DEFINING “MENTAL OR NERVOUS CONDITIONS,” Address before the Senate Advisory Committee, (Feb. 27, 2006).

⁹⁶ NAT'L ALLIANCE ON MENTAL ILLNESS, PRIORITY AND SPECIAL POPULATION (2009), http://www.nami.org/Content/NavigationMenu/Inform_Yourself/About_Public_Policy/NAMI_Policy_Platform/2_Priority_and_Special_Populations.htm (last visited January 26, 2009).

⁹⁷ Peck & Scheffler, *supra* note 74, at 1091.

⁹⁸ 42 U.S.C.A. § 300gg-5(e)(4) (West. Supp. 2009).

⁹⁹ In addition to a lack of consensus on the state level for mental illness, Medicaid coverage that individuals receive also varies greatly from state to state. The states have wide latitude within the confines of the federal guidelines and therefore the number of people covered and the amounts that each states spends on services varies across the states. Similar to this note, this has led to a discussion of solutions that would result in more equitable coverage across the states. *See* John Holahan & David Liska, *Variations in Medicaid Spending Among States Series A*,

variation does not appear with regard to what the states view to be a physical illness. In fact the states do not even deem it necessary to statutorily define physical illness. Therefore, with such a wide variation on the definition of mental illness in state laws, insurers will likely try to adopt the most stringent and narrow definitions of mental illness. As evidenced by the varying state laws, depending on the state in which one resides, the illnesses covered vary greatly.

V. BIOLOGICAL BASIS OF MENTAL ILLNESSES

The way society views mental illness has greatly evolved over the last two centuries.¹⁰⁰ The most recent report from the Surgeon General found that the stigma attached to mental illness dates back to the 19th century separation between mental health treatment and mainstream treatment in the United States.¹⁰¹ National surveys have tracked the public's perception of mental illness since the 1950's.¹⁰² The Surgeon General's Report states that "[i]n the 1950s, the public viewed mental illness as a stigmatized condition and displayed an unscientific understanding of mental illness."¹⁰³ In contrast, a 1996 survey found that people had a greater scientific understanding of mental illness; however, there was still a large amount of social stigma and people were more inclined to consider an individual with schizophrenia to have a mental illness in comparison to an individual with depression.¹⁰⁴

Mental illness is not diagnosed in the same way that physical illnesses may be diagnosed. One cannot test for mental illness by doing a

No. A-3 URB. INST. 1, 1-3 (1997), <http://www.urban.org/publications/307035.html>; Nicole Huberfeld, *Bizarre Love Triangle: The Spending Clause, Section 1983, and Medicaid Entitlements*, 42 U.C. DAVIS L. REV. 413, 469-70 (2008) stating "Though Medicaid was created to provide a statutory entitlement to states, providers, and enrollees, it has failed to ensure that enrollees receive promised benefits, both by lack of agency action and lack of statutory enforcement provisions." *Id.* at 469.

¹⁰⁰ See MENTAL HEALTH: A REPORT OF THE SURGEON GENERAL 6-7 (1999), <http://www.surgeongeneral.gov/library/mentalhealth/chapter1/sec1.html#overarching>.

¹⁰¹ *Id.* at 6.

¹⁰² *Id.* at 7.

¹⁰³ *Id.*

¹⁰⁴ *Id.*

blood test, x-rays, or throat swab.¹⁰⁵ Mental health professionals often meet with their patients to discuss the patients' symptoms and to have them describe how they have been feeling and how these feelings have affected their daily lives.¹⁰⁶ They may ask the patient about the length of time these symptoms have persisted as well as their severity.¹⁰⁷ After conducting this type of consultation, the mental health professionals will consult the DSM-IV.¹⁰⁸

In addition to this traditional method of diagnosing mental illness, there have been many advances in science that have permitted mental health professionals to analyze and diagnose their patients. For years scientists have been trying to investigate the biological and chemical processes of the brain. There have been many technological advances since the 1970's that now permit scientists and researchers to more closely examine and study the living brain.¹⁰⁹ One such technological advance includes magnetic resonance imaging, commonly known as MRI. This technology has allowed researchers to compare normal brain functions with those of individuals suffering from mental illness.¹¹⁰

Today it is clear that many mental disorders have a biological basis.¹¹¹ Mental illness is associated with changes in the brain's structure, chemistry and function.¹¹² Scientists already have the knowledge of how the brain typically functions and this new and ongoing research, which has revealed how the biological processes change when a person has a mental illness, has caused "scientists to minimize the distinctions between mental illnesses and these other brain disorders."¹¹³

¹⁰⁵ NAT'L INST. OF HEALTH, THE SCIENCE OF MENTAL ILLNESS 23 (2005), <http://science-education.nih.gov/supplements/nih5/Mental/guide/info-mental-a.htm> (last visited March 5, 2009).

¹⁰⁶ *Id.*

¹⁰⁷ *Id.*

¹⁰⁸ *Id.* at 23-24.

¹⁰⁹ Okianer Christian Dark, *Tort Liability and the "Unquiet Mind": A Proposal to Incorporate Mental Disabilities into the Standard of Care*, 30 T. MARSHALL L. REV. 169, 202 n.176 (2004).

¹¹⁰ *Id.*

¹¹¹ See MENTAL HEALTH: A REPORT OF THE SURGEON GENERAL, *supra* note 100, at 5, 15.

¹¹² See NAT'L INST. OF HEALTH, *supra* note 105, at 21.

¹¹³ *Id.*

The brain's basic functional unit is the neuron.¹¹⁴ The neuron possesses dendrites which receive signals and an axon that transmits the signals to other neurons.¹¹⁵ The area where an axon terminal ends near a receiving dendrite is called the synapse.¹¹⁶ In order for a neuron to relay information it uses both electrical signals and chemical messages called neurotransmission.¹¹⁷ People with mental illness have been found to have brain scans that do not reflect a normal functioning brain.

In the United States over the last five years, research studies examining the link between physical brain abnormalities and disorders like severe depression and schizophrenia have begun to make a strong case that the disorders are not scary tales of minds gone mad but manifestations of actual, and often fatal, problems in brain circuitry.¹¹⁸

In 1990 Congress and the President declared the 1990s to be the "Decade of the Brain,"¹¹⁹ and many studies over the last twenty years have made the biological connection between mental illness and the brain which has helped to reshape the way people look at mental illness.¹²⁰ This new research has refuted "the nineteenth century distinction between the organic mental illnesses (dementias and toxic psychoses) and the functional mental illnesses (including the neuroses and various affective or depressive disorders and the schizophrenic syndromes.)"¹²¹ Physical brain abnormalities can be viewed as biological instead of "mental" illness.¹²² In addition to being able to view the brain and its neural and electrical responses, scientists have found that certain mental illnesses also have a genetic basis. For example, Dr. Steven E. Hyman, a former director of the National Institute of Mental Health, has stated that "[g]enetic mutations and

¹¹⁴ *Id.* at 24.

¹¹⁵ *Id.* at 25.

¹¹⁶ *Id.*

¹¹⁷ *Id.* at 25-26.

¹¹⁸ Sarah Kershaw, *Insure Me, Please, The Murky Politics of Mind-Body*, N.Y. TIMES, Mar. 30, 2008, at WK, available at <http://www.nytimes.com/2008/03/30/weekinreview/30kers.html>.

¹¹⁹ Proclamation No. 6158, 55 Fed. Reg. 29,553-54 (July 20, 1990).

¹²⁰ See, e.g., Richard E. Gardner, III, Comment, *Mind Over Matter?: The Historical Search for Meaningful Parity Between Mental and Physical Health Care Coverage*, 49 EMORY L.J. 675, 682 (2000); Kershaw, *supra* note 118.

¹²¹ Gardner, *supra* note 120, at 682.

¹²² *Id.* at 682-83.

unlucky combinations of normal genes contribute to the risk of autism and schizophrenia.¹²³

Additional research has proven that there is a biological basis for schizophrenia, bipolar disorder and major depression.¹²⁴ Schizophrenia is one of the most debilitating types of mental illness.¹²⁵ It can often interfere with an individual's ability to think clearly, manage emotions, make decisions, interact and relate, and differentiate between reality and fantasy.¹²⁶ Research on the biological basis of schizophrenia has shown that there is a possible genetic disposition to the illness.¹²⁷

Research has shown that there are possible abnormalities in certain genes or in certain areas of the genome at a specific point on a specific chromosome.¹²⁸ Hundreds of studies have also proven that schizophrenic patients have less grey matter than non-schizophrenic patients in addition to enlarged ventricles and fluid-filled spaces in the brain.¹²⁹ Additionally, electrical transmissions in schizophrenic patients have been found to be abnormal.¹³⁰ This conclusion was reached in a study where electrodes were placed on the heads of schizophrenic patients and the electrical events were recorded and analyzed.¹³¹ A schizophrenic's brain's neurons do not function normally in the frontal lobe, revealing that a schizophrenic patient has fewer neurons than a normal patient, that their neurons are more randomly organized, and that their frontal lobes are also smaller.¹³² "The National Institute of Mental Health likens the search for better treatments

¹²³ See Pear, *supra* note 44.

¹²⁴ See Gardner, *supra* note 120, at 683-85.

¹²⁵ *Id.* at 683.

¹²⁶ NAT'L ALLIANCE ON MENTAL ILLNESS, SCHIZOPHRENIA (2009), http://www.nami.org/Content/ContentGroups/Helpline1/Schizophrenia_Fact_Sheet.htm.

¹²⁷ Subhagata Chattopadhyay, *Tracking Genetic and Biological Basis of Schizophrenia*, 2 INTERNET J. MENTAL HEALTH (2004).

¹²⁸ See generally, Richard E. Straub & Daniel R. Weinberger, *Schizophrenia Genes-Famine to Feast*, 60 BIOLOGICAL PSYCHIATRY 81, 82 (2006).

¹²⁹ Martha E. Shenton et al., *A Review of MRI Findings in Schizophrenia*, 49 SCHIZOPHRENIA RES. 1, 23, 34-35 (2001).

¹³⁰ See *id.* at 35.

¹³¹ Jyrki Ahveninen et al., *Inherited Auditory-Cortical Dysfunction in Twin Pairs Discordant for Schizophrenia*, 60 BIOLOGICAL PSYCHIATRY 612, 613-14 (2006).

¹³² See Gardner, *supra* note 120, at 683.

for those suffering from schizophrenia . . . to those involving heart disease or diabetes.”¹³³

Research over the last few decades has also revealed a biological basis in depression. The National Alliance on Mental Health has stated that scientific research “has firmly established that major depression is a biological, medical illness.”¹³⁴ Major depression is a condition which is continually persistent and can drastically interfere with an individual’s behavior, mood, level of activity, and physical health.¹³⁵ Research has shown that patients with major depression have a decreased level of neural activity in a specific area of the brain.¹³⁶ Additionally, major depression responds well to biologically based therapy which suggests that there is an organic nature to the illness.¹³⁷ Among medical illnesses, depression is the leading cause of disability in the United States and many other countries.¹³⁸

In addition to schizophrenia and major depression, research has shown that there is also a biological basis for bipolar disorder. Bipolar disorder, which is also known as manic depression, causes extreme shifts in one’s mood, energy and overall functioning.¹³⁹ Recurring episodes of mania and depression can last from a few days to months and usually begin in adolescence or early adulthood.¹⁴⁰ Studies have suggested that there may be a genetic basis for bipolar disorder due in part to the fact that the disorder often runs in families.¹⁴¹ The exact cause of bipolar disorder is unknown, although scientists believe that it is caused by multiple factors which produce a chemical imbalance in certain areas of the brain.¹⁴² While schizophrenia, major depression and bipolar disorder are included in most state definitions of mental illness, many other severe disorders “for which a

¹³³ See Dark, *supra* note 109, at 202.

¹³⁴ NAT’L ALLIANCE ON MENTAL ILLNESS, MAJOR DEPRESSION (2009), http://www.nami.org/Template.cfm?Section=By_Illness&template=/ContentManagement/ContentDisplay.cfm&ContentID=7725 .

¹³⁵ *Id.*

¹³⁶ See Gardner, *supra* note 120, at 684.

¹³⁷ *Id.* at 685.

¹³⁸ *Id.*

¹³⁹ NAT’L ALLIANCE ON MENTAL ILLNESS, BIPOLAR DISORDER (2009), http://www.nami.org/Template.cfm?Section=By_Illness&Template=/TaggedPage/TaggedPageDisplay.cfm&TPLID=54&ContentID=23037.

¹⁴⁰ *Id.*

¹⁴¹ *Id.*

¹⁴² *Id.*

biological basis has emerged through recent research” are excluded from the definition of mental illness by many states.¹⁴³

The advocates of mental health parity have tried to incorporate specific definitions of mental illness in past legislation. In 1999, Senators Domenici and Wellstone introduced the Mental Health Equitable Treatment Act of 1999.¹⁴⁴ This bill provided for specific illnesses to be covered called “severe biologically based mental illness” which included “schizophrenia, bipolar disorder, major depression, obsessive compulsive and panic disorders, posttraumatic stress disorder, autism, and other severe and disabling mental disorders such as severe anorexia nervosa and attention-deficit/hyper activity disorder.”¹⁴⁵ This bill did not pass and three years later they introduced the Mental Health Equitable Treatment Act of 2002.¹⁴⁶ This bill advocated for a broader definition of mental illness and stated that mental health benefits should include “all categories of mental health conditions listed in the Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition, or the most recent edition.”¹⁴⁷ This was the position advocated by the House of Representatives during the debate over the Paul Wellstone and Pete Domenici Mental Health Parity and Addiction Equity Act of 2008, but it was ultimately rejected.¹⁴⁸

With the growing acceptance of mental illnesses and their biological basis, there is no reason not to have a specific list of severe mental illness and substance use disorders in mental parity legislation. The new legislation states that mental health conditions or substance use disorders can be defined under the terms of the group health plan in accordance with applicable federal and state law.¹⁴⁹ Certain states do not even define mental illness at all and leave it solely to the health plan provider to define.¹⁵⁰ The evidence that many mental illnesses are

¹⁴³ Gardner, *supra* note 120, at 685. For example, many states exclude obsessive-compulsive disorder, panic disorder, eating disorders, and post-traumatic stress disorder. *See supra* Part III.

¹⁴⁴ Mental Health Equitable Treatment Act of 1999, S. 796, 106th Cong. (1999).

¹⁴⁵ *Id.* § 2(b)(5).

¹⁴⁶ Mental Health Equitable Treatment Act of 2001, S. 543, 107th Cong. (2001).

¹⁴⁷ *Id.* § 2(e)(3).

¹⁴⁸ *See* Pear, *supra* note 44.

¹⁴⁹ Emergency Economic Stabilization Act of 2008, Pub. L. No. 110-343, §512, 122 Stat. 3881-3892 (2008).

¹⁵⁰ *See* Kaplan, *supra* note 4, at 353.

biologically based and treatable is stronger today than it ever has been. Therefore, moving forward, there should be an attempt to pass an amendment that would more specifically state which mental illnesses should be covered instead of leaving it to the states and group plans, which results in a great variation in coverage that individuals may receive.

A good starting point for defining mental illness in parity legislation would be the definition of mental illness advocated for in 1999 that listed only severe mental illnesses. This definition was much narrower than the position advocated in 2002 and advocated by the House of Representatives in 2008, which was to include all illnesses in the DSM. While many states provide coverage for schizophrenia, bipolar disorder and major depression, many states do not cover obsessive-compulsive and panic disorders, PTSD, autism, ADHD and severe eating disorders, all of which were included in the list of severe mental illness in 1999. An amendment defining mental illness under these terms¹⁵¹ would provide the minimum list of mental illnesses that insurers would be required to cover. States should of course be allowed to provide additional coverage at their discretion.

VI. THE DIMINISHED SOCIETAL COSTS ASSOCIATED WITH PROVIDING MENTAL HEALTH PARITY COVERAGE

One of the reasons that MHPAEA was able to gain passage is that concerns over the cost implications of providing parity for mental health coverage were alleviated. Many feared that the cost of health plans for employers and insurers would skyrocket, however that was proven to be an unwarranted concern. As previously mentioned, a 2006 study in the *New England Journal of Medicine* found that insurers' costs rose less than half a percentage point when full parity was required for federal workers starting in 2001.¹⁵² Additionally, the Congressional Budget Office Cost Estimate, prepared in 2007, found that if the more generous House bill were enacted,

¹⁵¹ The definition would include "severe biologically based mental illness" defined as schizophrenia, bipolar disorder, major depression, obsessive-compulsive and panic disorders, posttraumatic stress disorder, autism, the eating disorders anorexia and bulimia, and severe attention-deficit/hyperactivity disorder.

¹⁵² See Shute, *supra* note 42.

the costs for premiums would increase for group health insurance by an average of only about 0.4 percent.¹⁵³

In addition to the fact that employer costs regarding health plans will not increase, many factors point towards proving that providing mental health parity will reduce the costs associated with treating mental illness.¹⁵⁴ For example, many studies have shown that individuals who receive psychiatric care incur fewer medical costs over the long term and less morbidity than those individuals who do not receive psychiatric treatment.¹⁵⁵

Psychiatrist Stephen M. Stahl, Director of the Clinical Neuroscience Research Center, has found numerous hidden costs that are associated with not treating major depression. Examples of these hidden costs include but are not limited to: fatal accidents resulting from impaired concentration; patient morbidity such as suicide attempts, accidents, resultant illnesses, lost jobs, failure to advance in career and school; and social costs such as dysfunctional families, absenteeism, decreased productivity, job-related illnesses, and adverse effects on quality control in the workplace.¹⁵⁶ Employers have also found that productivity tends to increase after workers are treated for mental illnesses and substance abuse problems and that such treatments can reduce the number of lost workdays.¹⁵⁷ The American Psychiatric Association has reported that untreated mental illness costs employers \$70 billion each year, which is primarily due to lost productivity.¹⁵⁸

In addition to the impact that mental illness has on work productivity, the mentally ill account for a large portion of both the homeless and the incarcerated. The Federal Task Force on Homelessness and Severe Mental Illness found that approximately one-third of the

¹⁵³ CONGRESSIONAL BUDGET OFFICE COST ESTIMATE (2007) at 4, <http://www.cbo.gov/ftpdocs/88xx/doc8837/hr1424ec.pdf>.

¹⁵⁴ Maria A. Morrison, *Changing Perceptions of Mental Illness and the Emergence of Expansive Mental Health Parity Legislation*, 45 S.D. L. REV. 8, 22-26 (2000).

¹⁵⁵ *Id.* at 26.

¹⁵⁶ *Id.*

¹⁵⁷ See Carroll, *supra* note 26, at 582.

¹⁵⁸ Sara Noel, Comment, *Parity in Mental Health Coverage: The Goal of Equal Access to Mental Health Treatment Under the Mental Health Parity Act of 1996 and the Mental Health Equitable Treatment Act of 2001*, 26 HAMLINE L. REV. 377, 398-99 (2003).

estimated 600,000 homeless people suffer from a severe mental illness.¹⁵⁹ Additionally, “[a]mong all jail inmates, twenty-four percent reported at least one symptom of psychotic disorder, and sixty-four percent reported some degree of mental health problems.”¹⁶⁰ The trend in the last several decades has been a shift from institutionalization of the mentally ill to the incarceration of the mentally ill.¹⁶¹ The reasons for this trend often have to do with a lack of funding from both the federal and state level, disallowing many mental hospitals and community-based mental health services the ability to provide adequate treatment.¹⁶² This has led to an increase, among other things, to the already increasing incarceration rate.¹⁶³ Studies have shown that incarcerating one individual costs more than \$23,000 per year, which is undoubtedly a huge cost on society.¹⁶⁴ While it is clear that there have been grave concerns over the cost of implementing mental health parity in the past,¹⁶⁵ numerous studies and data reports have alleviated those fears.¹⁶⁶ Without a specific definition of mental illness, many severe disorders will continue to go untreated and society will continue to suffer the consequences of inadequate treatment of the mentally ill.

VII. CONCLUSION

The Paul Wellstone and Pete Domenici Mental Health Parity and Addiction Equity Act of 2008 is a significant advancement in providing parity for mental illnesses and substance use disorders in that it finally provided complete parity between medical and surgical benefits and mental health and substance use disorders.¹⁶⁷ Although it is a significant step up

¹⁵⁹ *Id.* at 399 n.128.

¹⁶⁰ John D. King, *Candor, Zeal, and the Substitution of Judgment: Ethics and the Mentally Ill Criminal Defendant*, 58 AM. U. L. REV. 207, 211 n.12 (2008).

¹⁶¹ *See id.* at 212 n.15.

¹⁶² *Id.*

¹⁶³ *Id.*

¹⁶⁴ Maureen Carroll, Comment, *Educating Expelled Students After No Child Left Behind: Mending an Incentive Structure That Discourages Alternative Education and Reinstatement*, 55 UCLA L. REV. 1909, 1914 n.23 (2008).

¹⁶⁵ “Cost was the second most influential factor cited by interviewees as shaping the definition of mental illness.” Peck & Scheffler, *supra* note 74, at 1092.

¹⁶⁶ *See* Shute, *supra* note 42; CONGRESSIONAL BUDGET OFFICE COST ESTIMATE (2007), <http://www.cbo.gov/ftpdocs/88xx/doc8837/hr1424ec.pdf>.

¹⁶⁷ *See* 42 U.S.C.A. § 300gg-5(a)(3)(A) (West Supp. 2009).

from all previous parity bills, the new bill still has a glaring deficiency which, in order to be remedied, requires the passage of an amendment which would aid in achieving full parity for all citizens across all states.

The bill currently does not specifically define mental illness or substance use disorders. It states that the definition of mental illness and substance use disorders should be in accordance with state and federal laws.¹⁶⁸ As evidenced by the varying state laws, the illnesses covered, and thus how much coverage individuals are afforded, varies greatly depending on the state in which one resides. With the growing acceptance of mental illnesses and the evidence that mental illnesses are biologically based, it is clear that mental illnesses are serious disorders that can be and must be treated. Monetary costs have been a central part of the parity debates for years and have often been a major factor in how states shape their definitions of mental illness. Now that concerns over cost have been alleviated, it is no longer a valid argument against a list of specific mental illnesses to be covered. Furthermore, society as a whole will greatly benefit from the adequate treatment of the mentally ill.¹⁶⁹

An amendment should be passed in order to provide a specific definition of what constitutes a mental illness or substance use disorder so that the coverage people receive is not varied. The amendment should be based on the definition that was advocated in the Mental Health Equitable Treatment Act of 1999.¹⁷⁰ That bill provided for specific illnesses to be covered called “severe biologically-based mental illness” and included “schizophrenia, bipolar disorder, major depression, obsessive compulsive and panic disorders, posttraumatic stress disorder, autism” as well as “anorexia nervosa and attention-deficit/hyper activity disorder.”¹⁷¹ An amendment should be passed based on this previously proposed definition and should identify each of the illnesses listed above. The MHPAEA has done a great deal in providing parity between mental illness, substance use coverage and physical illness coverage. However, in order to truly provide parity for mental illness and substance use disorders, an amendment should be passed so that individuals throughout the country receive the same level of coverage and can receive treatment for the most debilitating mental disorders.

¹⁶⁸ *Id.* § 300gg-5(e)(4).

¹⁶⁹ *See supra* Part V.

¹⁷⁰ Mental Health Equitable Treatment Act of 1999, S. 796, 106th Cong. (1999).

¹⁷¹ *Id.*

EXAMINING CURRENT PROPOSALS FOR INCREASING THE FEDERAL ROLE IN DEALING WITH COASTAL HURRICANE RISK

Louis Cruz*

This note distinguishes predatory from subprime lending, while focusing on the insurance consequences of predatory lending. It considers how single premium credit insurance (SPCI) and private mortgage insurance (PMI), two mortgage-related insurance products, have affected the current predatory lending crisis. This note argues for reform that eliminates SPCI and makes PMI a more feasible option for insureds. Such reform would allow subprime lenders to offer mortgages to qualified borrowers, while reducing the amount of predatory lending and foreclosures. The introduction of this note presents some background information regarding subprime lending and predatory lending. The second part examines several issues concerning the role of insurance in the subprime mortgage market. Third, reform measures necessary to alleviate the issues with mortgage insurance are discussed. Finally, the fourth section studies recent actions by the Federal Reserve Board and analyzes whether they can be expected to bring meaningful change. It concludes that, although the Fed's new regulations are a step in the right direction, there needs to be an outright ban of SPCI and predatory must be stopped completely.

I. INTRODUCTION

The recent flurry of coastal risk proposals vying for federal adoption mirrors the sudden onslaught of a severe coastal storm. The clamor to develop a framework to effectively manage coastal risk is understandable. The public and private sector response to Hurricane Katrina show that the current system for handling coastal risk is becoming less feasible by the day. Katrina has led leading observers to note that,

* University of Connecticut School of Law, Juris Doctor Candidate for the Class of 2010.

“[i]nsurers – and society as a whole – need to reconsider the potential for mega-catastrophes” and in turn “will need to adjust to this new reality.”¹

Ideally, the proposal that emerges as successful should exhibit three characteristics. First, the plan the federal government ultimately adopts should expand the capacity of the existent private market to field coastal risk so as to better deal with future storms. Second, the plan should avoid completely crowding out the existent private sector players that already seek to insure coastal risk. Finally, the successful plan should minimize the degree to which non-coastal property owners subsidize the risk undertaken by those who choose to live in storm-prone coastal areas.

It may be that no current proposal fully satisfies all three criteria. Indeed, upon analysis, none of them seem to. However, some of the current coastal risk plans contain aspects that show promise in light of these three criteria. The proposals that show promise lack federal reinsurance mechanisms, thereby not crowding out private market reinsurers and sidestepping large scale cross-subsidization. Additionally, the most favorable of the available plans contain coastal risk transferring mechanisms designed to increase the use of the private capital market’s capacity to deal with future catastrophic coastal risk. Also, plans that contemplate homeowner catastrophe savings accounts may provide a way to enable coastal residents to afford actuarially sound private market insurance rates. Finally, plans that envision private insurer catastrophe reserving may present ways to increase private market capacity to deal with coastal catastrophe risk, thereby keeping private insurers in the business of insuring coastal risk.

The following examination of the current proposals begins with a brief survey of the meteorological and geographical settlement trends that have worked to necessitate an immediate rethinking of coastal risk policy. Next, this study will move into a review of the current coastal risk framework. Following this, the current proposals will each be discussed. Finally, each proposal will be evaluated against the capacity, displacement and cross-subsidization criteria discussed above.

¹ TOWERS PERRIN, HURRICANE KATRINA: ANALYSIS OF THE IMPACT ON THE INSURANCE INDUSTRY 4 (2005).

II. BACKGROUND

A. EXTREME WEATHER TRENDS

On a world-wide scale, extreme weather events are on the rise. The number of major weather-caused natural catastrophes has increased from an average of 1.5 per year since the 1950s to 4.5 in recent years.² As part of this trend, catastrophic storms in the Atlantic and Gulf Coasts of the U.S. have also become more prevalent.³ While climatologists disagree as to whether the available data is adequate to determine the magnitude of this current extreme weather event uptick,⁴ what cannot be debated is that the number of storms per hurricane season in the Atlantic and Gulf coasts has increased since as recently as 1995.⁵ In comparison to the period of 1970-1994, the amount of hurricane activity on the Atlantic and Gulf coasts has increased by over 60% in the ten years that followed.⁶ Further examination of the historical extreme weather data reveals that compared to an Atlantic basin annual average of 1.5 major storms a year in the period from 1970 to 1994, the region has experienced an average of 3.9 major storms per year since 1995.⁷ When measuring only the subset of storms that actually make

² COMITÉ EUROPÉEN DES ASSURANCES, REDUCING THE SOCIAL AND ECONOMIC IMPACT OF CLIMATE CHANGE AND NATURAL CATASTROPHES 10 (2007).

³ Rick Jervis, *Data Show U.S. Riding Out Worst Storms on Record*, USA TODAY, Oct. 22, 2008, at 3a.

⁴ See Eliot Kleinberg, *Hurricanes May Trace to Cycles, Not Warming*, PALM BEACH POST, June 3, 2007, at 1C (noting that there are problems with researchers basing their conclusions on a small statistical sample of scientifically significant research); *but cf.* Cathy Zollo, *Experts Spar Over Warming's Impact*, SARASOTA HERALD-TRIBUNE, Apr. 15, 2006, at A1 (many climatologists argue there is enough data, and some describe the recent uptick in storms as a natural cyclical occurrence, while others attribute it to global warming).

⁵ See generally Stanley B. Goldenberg et al., *The Recent Increase in Atlantic Hurricane Activity: Causes and Implications*, 293 SCIENCE 474 (2001) (discussing the years 1995 to 2000 as experiencing the "highest level of North Atlantic hurricane activity in the reliable record.").

⁶ HOWARD KUNREUTHER, REFLECTIONS ON U.S. DISASTER INSURANCE POLICY FOR THE 21ST CENTURY 2 (Am. Enter. Inst.-Brookings Joint Ctr. For Regulatory Studies, 2007).

⁷ PHILIP J. KLOTZBACH & WILLIAM M. GRAY, EXTENDED RANGE FORECAST OF ATLANTIC SEASONAL HURRICANE ACTIVITY AND LANDFALL STRIKE PROBABILITY FOR 2009 26 (2009).

landfall the data is just as striking. From the 1930s to the 1960s, there was an average of 1.8 storms per year making landfall, while over the last twelve years that figure has risen to 2.2.⁸

In spite of the fact that the causes of this trend are beyond the scope of this study, what is clear is that there has been a pronounced increase in the number of coastal storms in the southeast Atlantic and Gulf coasts over the past decade. While this meteorological trend is startling when considered on its own, what makes it all the more attention worthy is that it coincides with a trend of increasing coastal populations in the United States.

B. COASTAL POPULATION TRENDS

At present, coastal counties comprise 17% of the total land area of the United States, but claim 53% of the country's total population.⁹ While the country's coastal-to-non-coastal population ratio has remained relatively stable over the last forty years, the limited area of coastal geography has contributed to significantly higher population densities in these coastal areas.¹⁰ Data from the 2000 census show that while the country at large experienced a density increase in the range of 38% from 1970 to 2000, the Southeast Atlantic coast area had an increase in population density of nearly 66% over the same period.¹¹ This study focuses particularly on those coastal regions that have historically been most susceptible to coastal storm damage: the Southeast Atlantic and Gulf coasts.¹²

Currently, 9% of the nation's coastal population resides in the Southeast Atlantic region.¹³ Florida in particular has most of its total

⁸ Manuel Lonfat et al., *Atlantic Basin, U.S. and Caribbean Landfall Activity Rates over the 2006-2010 Period: An Insurance Industry Perspective*, 59A TELLUS 499, 500 (2007).

⁹ KRISTEN M. CROSSETT ET AL., NAT'L OCEANIC AND ATMOSPHERIC ADMIN., POPULATION TRENDS ALONG THE COASTAL UNITED STATES: 1980-2008 6 (2004).

¹⁰ *Id.*

¹¹ KUNREUTHER, *supra* note 6, at 1.

¹² See Stanley A. Changnon, *Characteristics of Severe Atlantic Hurricanes in the United States: 1949-2006*, 48 NAT. HAZARDS 329, 333 (2009) (noting that in the 57 year period examined, the Southeast and South ranked first and second in the United States for the number of losses from storm events).

¹³ CROSSETT ET AL., *supra* note 9, at 16.

population distributed throughout its coastal counties.¹⁴ Specifically noteworthy is the fact that it had been anticipated that the southeast Atlantic region's coastal population would increase by 1.1 million people, or 8%, between 2003 and 2008.¹⁵ This would be the largest percentage increase of all U.S. coastal regions within that five year period.¹⁶ Concomitant with this population growth is an increase in insured coastal property. More than 80% of Florida's total insured property exposure can be classified as coastal.¹⁷ The estimated total value of insured coastal exposure in Florida is over \$2 trillion dollars.¹⁸

Extending from the Florida Keys to southern Texas and including the coastline of six states, the Gulf region's coastal population of 19.1 million residents claims just over 13% of the nation's total coastal population.¹⁹ The majority of the population in both Louisiana and the western coast of Florida are in Gulf coastal counties.²⁰ A total of 23% of the region's total land area and 32% of the region's population are distributed throughout the Gulf Region's 144 coastal counties.²¹ The combined commercial and residential insured exposure in the Gulf region is estimated to be over one trillion dollars.²²

C. IMPLICATIONS AND IMPACT

On the whole, the total insured value of property in U.S. coastal regions is increasing at an astonishing rate. Recently it was determined that from 2004 through 2007, the insured value of properties in coastal areas of the United States grew at a compound annual growth rate of just

¹⁴ *Id.*

¹⁵ *Id.*

¹⁶ *Id.*

¹⁷ Press Release, Prop. Cas. Insurers Ass'n of Am., Coastal Insurance Concerns Demonstrate Need for New Catastrophe Solutions (Mar. 15, 2007), <http://www.pciaa.net/> (follow "Media Center: News Releases: By Date" menu; then scroll to Mar. 15, 2007).

¹⁸ *Id.*

¹⁹ CROSSETT ET AL., *supra* note 9, at 18.

²⁰ *Id.*

²¹ *Id.*

²² See Ieva M. Augustums, *Nearly \$1T of Insured Property in Ike's Path*, USA TODAY, Sept. 12, 2008, http://www.usatoday.com/money/economy/2008-09-11-2141571223_x.htm.

over 7%.²³ This annual growth rate will lead to a doubling of the total insured coastal value every decade.²⁴ In light of the recent increases in extreme weather events in the Atlantic and Gulf coasts, this insured growth takes on a number of serious implications. Due to the increasing frequency of coastal storms, the growing size of coastal communities and the rapid growth rate of insured exposure concentrated in coastal areas, extreme weather related events are becoming increasingly financially destructive. Five of the ten most expensive storms in U.S. history have occurred since 1990.²⁵ Of particular import, the 2004/05 hurricane season was unprecedented. In terms of U.S. insured loss, the six hurricanes that hit the southeast Atlantic and Gulf Coasts make up half of the list of the twelve largest disasters in the last forty years.²⁶ Katrina alone is the most costly event the U.S. insurance industry has ever experienced, resulting in more than \$61 billion in insured losses and \$125 billion in total losses.²⁷

Another important implication for these current trends in extreme weather and coastal population growth is the need for a reevaluation of current policy. The federal role in coastal disaster policy has been the subject of recent intense debate. To date, a flurry of legislative and private sector proposals have followed in the wake of the 2004/05 hurricane season. Each of these proposals envisions a reoriented role for federal government in coastal catastrophe policy in an attempt to overhaul the current coastal risk policy framework.

III. CURRENT COASTAL RISK POLICY FRAMEWORK

In the context of coastal hurricanes, the current policy landscape is a patchwork of private insurance, government assistance and government subsidized insurance programs. Presently, private insurance plays a relatively limited role. Due to the fact that many private insurers perceive the risks associated with coastal catastrophic loss as unacceptably high, many have curtailed writing such policies.²⁸ Some major insurers, such as

²³ AIR WORLDWIDE CORP., *THE COASTLINE AT RISK: 2008 UPDATE TO THE ESTIMATED INSURED VALUE OF U.S. COASTAL PROPERTIES 1* (2008).

²⁴ *Id.*

²⁵ ALEXANDER BOLONKIN, *CONTROL OF REGIONAL AND GLOBAL WEATHER 3* (2007), <http://arxiv.org/pdf/physics/0701097>.

²⁶ KUNREUTHER, *supra* note 6, at 2.

²⁷ COMITÉ EUROPÉEN DES ASSURANCES, *supra* note 2, at 10.

²⁸ *See, e.g.,* Sandra Fleishman, *Sea Change in Insurers' Coastal Coverage: Many Firms Opt to End or*

Allstate, no longer write new policies in Florida, Louisiana, Mississippi, parts of New York, and coastal Texas.²⁹ Similarly, State Farm has stopped offering insurance that would protect against storm damage within one mile of the ocean and announced in 2006 that it would sell no new policies in Mississippi.³⁰ Moreover, where available, private coverage often protects against wind damage but, excludes flood coverage.³¹ As a result, property owners who want flood coverage must purchase it separately. This is done through either the National Flood Insurance Program or state sponsored Fair Access to Insurance Requirement plans.³²

A. NATIONAL FLOOD INSURANCE PROGRAM

The National Flood Insurance Program was established by Congress in 1968.³³ When enacted the goals of the National Flood Insurance Program were twofold: minimize flood damage through floodplain management, and provide property owners with flood insurance.³⁴ Corresponding to these goals are the two main initiatives that currently characterize the National Flood Insurance Program. First, National Flood Insurance Program actively generates flood maps indicating 100-year floodplains.³⁵ These maps anchor the National Flood Insurance

Limit New Policies, WASH. POST, Dec. 30, 2006, at F1.

²⁹ Spencer S. Hsu, *Insurers Retreat from Coasts; Katrina Losses May Force More Costs on Taxpayers*, WASH. POST, Apr. 30, 2006, at A1; see also *Examining the Terrorism Risk Insurance Program, Before S. Comm. on Banking, Housing and Urban Affairs*, 107th Cong. 24 (2007) (Statement of J. Robert Hunter, Director of Insurance, Consumer Federation of America).

³⁰ Fleishman, *supra* note 28; see also *An Examination of the Availability and Affordability of Property and Casualty Insurance in the Gulf Coast and Other Coastal Regions: Hearing Before the S. Committee on Banking, Housing and Urban Affairs*, 109th Cong. 7 (2007) (statement of Mel Martinez, Member, S. Comm. on Banking, Housing, and Urban Affairs).

³¹ JUSTIN R. PIDOT, *COASTAL DISASTER INSURANCE IN THE ERA OF GLOBAL WARMING: THE CASE FOR RELYING ON THE PRIVATE MARKET* 37 (Georgetown Envtl. Law & Policy Inst., 2007).

³² *Id.* at 12, 21.

³³ U.S. GOV'T ACCOUNTABILITY OFFICE, GAO-07-285, *CLIMATE CHANGE: FINANCIAL RISKS TO FEDERAL AND PRIVATE INSURERS IN COMING DECADES ARE POTENTIALLY SIGNIFICANT* 46 (2007).

³⁴ See PIDOT, *supra* note 31, at 12-13.

³⁵ *Id.* at 13.

Program's premium rates and mitigation requirements.³⁶ Second, the program offers up to \$250,000 in insurance against flood damage to homeowners in communities that have adopted floodplain regulations meeting the minimum standards set by the program.³⁷ Although participation was initially voluntary the National Flood Insurance Program now requires homeowner's within mapped floodplains who have federally insured mortgages to purchase and maintain flood insurance.³⁸

Perhaps the most important aspect of the National Flood Insurance Program is the *Write-Your-Own* program. In 1983 the federal government permitted private companies to write National Flood Insurance Program coverage under the *Write-Your-Own* program.³⁹ Under this program private insurers wrote policies and handled claims and in exchange received roughly a third of the premiums collected as sales commission and sustained 3.3% of incurred losses.⁴⁰ Currently, of the nearly 5.3 million policies written by the National Flood Insurance Program, upwards of 95% of the policies are written by private companies under the *Write-Your-Own* program.⁴¹

In spite of its long history, the National Flood Insurance Program has significant shortcomings. Due to the heavy subsidization of the premiums by the federal government the program is often criticized as not being actuarially sound and therefore ultimately not self-supporting.⁴² The effect of subsidized premiums is two-fold. First, the program does not generate enough reserves to protect against catastrophic loss.⁴³ This can be seen by the fact that of the \$23 billion in claims paid out by the National Flood Insurance Program following Hurricane Katrina much of this money

³⁶ *Id.*

³⁷ 42 U.S.C. §§ 4012, 4013 (2006).

³⁸ 42 U.S.C. § 4012a (b)(1) (2006).

³⁹ DEP'T OF HOMELAND SEC., PRIVACY IMPACT ASSESSMENT FOR THE NATIONAL FLOOD INSURANCE PROGRAM APPEALS PROCEDURE 2, 5 (2006).

⁴⁰ WILLIAM O. JENKINS, JR., U.S. GOV'T ACCOUNTABILITY OFFICE GAO-05-532T, NATIONAL FLOOD INSURANCE PROGRAM: OVERSIGHT OF POLICY ISSUANCE AND CLAIMS 8 (2005); Robert J. Rhee, *Catastrophic Risk and Governance After Hurricane Katrina: A Postscript to Terrorism Risk in a Post-9/11 Economy*, 38 ARIZ. ST. L.J. 581, 610 (2006).

⁴¹ JENKINS, *supra* note 40, at 7.

⁴² U.S. GOV'T ACCOUNTABILITY OFFICE GAO-06-119, FEDERAL EMERGENCY MANAGEMENT AGENCY: IMPROVEMENTS NEEDED TO ENHANCE OVERSIGHT AND MANAGEMENT OF THE NATIONAL FLOOD INSURANCE PROGRAM 9 (2005).

⁴³ *Id.*

came from loans taken directly from the federal treasury.⁴⁴ Second, because the premiums are subsidized and therefore artificially low, property owners have no incentive to avoid moral hazard in or to invest mitigation measures.⁴⁵ This dynamic is demonstrated by the fact that property owners that have repeatedly suffered damage from floods have rebuilt their property in the same locations and continue to receive subsidized rates.⁴⁶ This is in stark contrast to the ideally analogous scenario in which private insurers have the ability to cancel coverage.

B. STATE DISASTER INSURANCE PLANS

Many states have in operation, Fair Access to Insurance Requirement or “FAIR” plans. Ostensibly, these plans provide coverage to property owners unable to secure insurance in the private market. Although FAIR plans were not originally designed to cope with coastal natural catastrophe loss, these plans have expanded considerably and today many states operate plans that primarily provide coastal storm coverage.⁴⁷ FAIR plans are state-run insurance pools in which all property insurers licensed in a state are required to participate.⁴⁸ Participating insurers share in the profits and losses of the high risk coverage.⁴⁹ In 2004, a total of \$400 billion of coastal property was insured by state run FAIR plans.⁵⁰

Instead of FAIR plans, other states have enacted Beach and Windstorm Insurance Plans that provide coverage to property-owners in coastal communities.⁵¹ Much like FAIR plans, most states that operate Windstorm plans require participation by private property insurance carriers in the state. Again, as in the case of FAIR plans, here the insurers

⁴⁴ See RAWLE O. KING, CONG. RESEARCH SERV., NATIONAL FLOOD INSURANCE PROGRAM: TREASURY BORROWING IN THE AFTERMATH OF HURRICANE KATRINA 1, 3-4 (2006).

⁴⁵ See PIDOT, *supra* note 31, at 34.

⁴⁶ *Id.*; Spencer M. Taylor, *Insuring Against the Natural Catastrophe after Hurricane Katrina*, 20 NAT. RESOURCES & ENV'T. 26, 28 (2006).

⁴⁷ See PIDOT, *supra* note 31, at 21.

⁴⁸ See Ins. Info. Inst., *Residual Markets*, Aug. 2009, <http://www.iii.org/media/hottopics/insurance/residual/>.

⁴⁹ *Id.*

⁵⁰ ROBERT P. HARTWIG & CLAIRE WILKINSON, INS. INFO. INST., PUBLIC/PRIVATE MECHANISMS FOR HANDLING CATASTROPHIC RISKS IN THE UNITED STATES 33 (2005).

⁵¹ See Ins. Info. Inst., *supra* note 48.

share profits and losses. In 2004 upwards of \$30 million of coastal property was insured under some form of state Beach or Windstorm plan.⁵²

Still other states have crafted even more unique programs. Florida created the “Florida Citizens”, not-for-profit “insurer of last resort.”⁵³ In contrast to FAIR and Windstorm plans, Florida Citizens is less an insurance pool, than a state run insurer.⁵⁴ At present, Florida Citizens is the largest insurer in the state with more than 1.3 million policyholders.⁵⁵ Unfortunately, Florida Citizens’ premium rates are set by statute at levels that are not actuarially sound.⁵⁶ In 2007 Florida Citizens had more than \$400 billion in exposure, but received only \$3 billion in yearly premiums.⁵⁷ When Florida Citizens’ losses exceed its claims-paying capacity in a single year, it is required by statute to post-fund itself by imposing a statewide assessment on every other line of insurance sold in the state.⁵⁸

IV. THE PROPOSED PLANS

A. COALITION OF AMERICANS FOR SMART NATURAL CATASTROPHE POLICY: MITIGATION

Formed in late 2005, the Coalition of Americans for Smart Natural Catastrophe Policy is a grouping of environmentalists, academics, and consumer rights groups.⁵⁹ At present, the Coalition is one of the most high

⁵² HARTWIG & WILKINSON, *supra* note 50, at 39.

⁵³ *Id.* at 32, 34.

⁵⁴ *Id.* at 40.

⁵⁵ INS. INFO. INST., AN EXCERPT FROM: RESIDUAL MARKET PROPERTY PLANS -- FROM MARKETS OF LAST RESORT TO MARKETS OF FIRST CHOICE 2 (2007).

⁵⁶ *See* PIDOT, *supra* note 31 at 18.

⁵⁷ John W. Rollins, Florida Property Insurance – The “Citizens” View (June 19, 2007), <http://www.casact.org/education/spring/2007/handouts/rollins.pdf>.

⁵⁸ INS. INFO. INST., *supra* note 55, at 4.

⁵⁹ Presently the Coalition consists of: Consumer Federation of America, Defenders of Wildlife, Environmental Defense, Friends of the Earth, National Wildlife Federation, Republicans for Environmental Protection, Association of State Floodplain Managers, Americans for Prosperity, Council for Citizens Against Government Waste, Competitive Enterprise Institute, FreedomWorks, Taxpayers for Common Sense, Association of Bermuda Insurers and Reinsurers, Reinsurance Association of America, and the National Association of Professional Insurance Agents. *See* Smarter Safer, About Us, <http://www.smartersafer.org/about-us/> (last visited Sept. 30, 2009)(for more information on the coalition’s expanding roster).

profile groups advocating their own national catastrophe proposal. Unlike many of the other proposals discussed later, the Coalition's plan focuses on policy encouraging homeowner mitigation methods rather than federal reinsurance or federal lending to state insurance pools. According to the Coalition's current mission statement they believe:

[T]he Federal government has a role in encouraging and helping homeowners to undertake mitigation efforts to safeguard their homes against hurricanes, [but] ... the coalition oppose proposals being considered in Congress that would create moral hazards by providing direct or indirect subsidies for coastal homeowners' insurance policies, thereby giving people incentives to build homes in hurricane-prone, environmentally sensitive areas.⁶⁰

Specifically, the Coalition proposes that federal intervention be limited to existing programs modified to increase emphasis on preparedness and mitigation as opposed to any sort of federally aided insurance coverage expansion.⁶¹ The full breadth of the Coalition's platform can be seen in its recent promotion of the Flood Insurance Reform Modernization Act of 2007 and the Property Mitigation Assistance Act; as well as its opposition to the Homeowners Defense Act of 2007.

The Coalition supported the version of the Flood Insurance Reform and Modernization Act of 2007 that appeared before the Senate.⁶² While described by the Coalition as modest legislation reauthorizing the National Flood Insurance Program, the Flood Insurance Reform and Modernization Act of 2007 would have made fairly significant changes to the program.⁶³ First, the act would have attempted to make the National Flood Insurance Program satisfy traditional criteria for actuarial soundness by phasing out discounted premiums previously available for structures built prior to the mapping and implementation of the program's floodplain management requirements.⁶⁴ At present, these prorate map structures pay heavily discounted rates on the first \$35,000 of their structure's insured value, and full risk-based premium rates for the remaining insured value.⁶⁵ Currently,

⁶⁰ *Id.*

⁶¹ *See id.*

⁶² Smarter Safer, Senate Legislation, <http://www.smartersafer.org/about-the-legislation/senate-legislation>, (last visited Sept. 30, 2009).

⁶³ *Id.*

⁶⁴ *See* Flood Insurance Reform and Modernization Act of 2007, S. 2284, 110th Cong. §§ 8(g), 14 (2007).

⁶⁵ *An Examination of the National Flood Insurance Program: Hearing Before The United States Senate Comm. On Banking, Housing, and Urban Affairs*, 110th

nearly a quarter of all policies written by the National Flood Insurance Program are subsidized under this aspect of the program.⁶⁶

Further, the Senate version of the Flood Insurance Reform and Modernization Act of 2007 would have allowed the increase of National Flood Insurance policy rates by 15% a year, up from the previous 10% cap.⁶⁷ The bill's drafters arrived at this figure because the previous 10% ceiling was shown not to be enough to ensure the program would have sufficient funds to cover future obligations for policyholder claims, operating expenses, and interest on debt stemming from the 2005 hurricane season.⁶⁸

In addition to supporting the Flood Insurance Reform and Modernization Act of 2007 the Coalition also successfully opposed a substantial amendment to the bill that would have made wind coverage available to National Flood Insurance policyholders.⁶⁹ In support of their opposition to the wind coverage amendment, the Coalition cited estimates that such an amendment would result in as much as \$161 billion in new taxpayer liabilities in 2009 alone if the U.S. Gulf Coast suffered a hurricane season comparable to that of 2005.⁷⁰ Further still, the Coalition urged that the expanded program would also threaten public safety by encouraging further development in hurricane prone coastal areas.⁷¹ In a press release lauding the Senate's subsequent rejection of the amended bill, the Coalition suggested that instead of expanding the National Flood Insurance Program, Congress should look to, "safety-oriented reform solutions that would help homeowners better prepare for storms and reduce destruction caused by

Cong. 2 (2007) (testimony of David I. Maurstad Assistant Admin. for Mitigation and Fed. Ins. Admin. of Fed. Emergency Mgmt. Agency).

⁶⁶ *Id.*

⁶⁷ See Flood Insurance Reform and Modernization Act of 2007, S. 2284, 110th Cong. § 6(b) (2007).

⁶⁸ KING, *supra* note 44, at 2-3.

⁶⁹ Press Release, Americans for Smart Natural Catastrophe Policy, Americans for Smart Natural Catastrophe Policy Applauds Congress' Extension of National Flood Program Without Irresponsible Expansion (Sept. 29, 2008), www.smartersafer.org/newsroom/press-releases (follow "Americans for Smart Natural Catastrophe Policy Applauds Congress' Extension of National Flood Program Without Irresponsible Expansion" hyperlink).

⁷⁰ *Id.*

⁷¹ *Id.*

natural catastrophes.”⁷² The 2007 Property Mitigation Act is an example of a safety-oriented solution can that was supported by the Coalition.

As part of the Coalition’s overall program, the group championed the unsuccessful 2007 Property Mitigation Assistance Act.⁷³ The Act sought to “authorize grants and loans to homeowners to harden their homes against hurricanes and other disasters.”⁷⁴ Pursuant to the Act, states that met the terms in the bill would have received grants of at least \$500,000 to be distributed to residents through loan and grant programs to help residents take such measures as adding storm shutters, hurricane clips, safe rooms or any other activity that would mitigate the risks of future hazards and natural disasters.⁷⁵ The idea was to help at-risk homeowners protect their homes from damage instead of expanding federal involvement in insurance coverage or aid.⁷⁶

Finally, the Coalition’s overall view of national coastal natural catastrophe policy can be seen in their opposition to the Senate version of the Homeowners’ Defense Act of 2007. The Coalition described the proposed legislation as irresponsible.⁷⁷ Essentially, the Homeowners’ Defense Act of 2007 consisted of three bundled programs designed to work with existing state insurance pools. The first program would have established an interstate federal consortium that would have attempted to aid multiple states running coastal risk insurance pools in combining and transferring this risk to the capital markets.⁷⁸ The second program was an insurance stabilization plan that would have allowed the Federal Treasury to make loans to state pools in order to ensure their continued, liquidity in the aftermath of a natural catastrophe.⁷⁹ The third and final component of the Act would have established a Federal reinsurance program designed to sell reinsurance to state and interstate pools.⁸⁰

⁷² *Id.*

⁷³ Smarter Safer, *supra* note 62.

⁷⁴ *Id.*; Property Mitigation Assistance Act of 2007, S. 2328, 110th Cong. § 2(m) (2007).

⁷⁵ Property Mitigation Assistance Act of 2007, S. 2328, 110th Cong. § 2(m) (2007).

⁷⁶ See Smarter Safer, *The Senate Should Support The Property Mitigation Assistance Act of 2007*, <http://www.smartersafer.org/uploads/S2328.pdf> (last visited Sep.15, 2009).

⁷⁷ Smarter Safer, *supra* note 62.

⁷⁸ Homeowners’ Defense Act of 2007, S.2310 §§ 101-108 (2007).

⁷⁹ See *id.* § 201.

⁸⁰ See *id.* § 301.

The Coalition's criticism of the Senate version of the Homeowners' Defense Act of 2007 was that each aspect of the Act would simply expand federal liability by introducing a federal role in state insurance pools.⁸¹ As to the Federal Risk Consortium, the Coalition pointed out that the States were already free to associate, and address catastrophe risk.⁸² Therefore a permanent federal role in this process would be unnecessary.⁸³ Similarly, in their opposition to the aspects of the Senate version of the Homeowners' Defense Act of 2007 that provided for Federal reinsurance and Federal loans to state programs, the Coalition adopted the position that the Act would both displace existing private reinsurers and promote the continuation of financially unsound state insurance programs.⁸⁴

B. ALLSTATE'S PROPOSAL: FEDERAL REINSURANCE BACKSTOP TO STATE POOLS

Supporting a bill that would have created a federal reinsurance backstop even prior to the monumental 2005 hurricane season, Allstate was among the first private sector supporters of the concept of an increased federal role in coastal natural catastrophe insurance coverage.⁸⁵ As Allstate's was the first of the major insurer backed plans, the proposal was vaguer than some subsequent proposals. However, a more fleshed out version of the original Allstate proposal was soon incorporated into the stalled Homeowners Insurance Protection Act of 2005. Allstate subsequently endorsed the Act.⁸⁶ With this in mind, the following description of the Allstate proposal includes details from the Homeowners Insurance Protection Act of 2005 to fill in some of the gaps that this early private sector proposal exhibited.

⁸¹ See Smarter Safer, *The Senate Should Not Adopt S. 2310* (2007), <http://www.abir.bm/downloads/SmartNatCatBriefingPapers.pdf> (last visited Sep. 15, 2009).

⁸² *See id.*

⁸³ *See id.*

⁸⁴ *See id.*

⁸⁵ R.J. Lehmann, *A Catastrophic Battle*, BEST'S REVIEW, Oct. 2006, at 38.

⁸⁶ *Id.*; Business Wire, *Allstate CEO to Call for National Catastrophe Plan at National Press Club on Friday, Jan. 13; Devastation of 2005 Makes Reform an Immediate National Priority*, Jan. 9, 2006, available at http://findarticles.com/p/articles/mi_m0EIN/is_2006_Jan_9/ai_n15989710/ (last visited Sept. 30, 2009).

Allstate's initial position was to create a federal government sponsored catastrophe program that would provide reinsurance to a system of state and regional catastrophe funds.⁸⁷ This required states that did not already have them to create catastrophe pools funded by, "all entities that benefit from a robust local economy such as the banking and real estate sectors."⁸⁸ In turn, the Department of the Treasury would make reinsurance contracts available to any of the state programs that met minimum requirements.⁸⁹

As to rates, the text of the Homeowners Insurance Protection Act of 2005 required that rates be risk-based.⁹⁰ States would pay premiums directly to the federal government.⁹¹ The federal reinsurance would cover 90% of losses in excess of either the capacity of each state program or the projected losses from a 200-year event, whichever is greater.⁹² This translates into the federal reinsurance coverage being triggered by events causing roughly \$50 billion or more in homeowners' losses.⁹³

C. TRAVELERS AND NATIONWIDE: FEDERAL REINSURANCE AND INCREASED FEDERAL REGULATION

Two other leading property-casualty insurance carriers, Travelers and Nationwide, soon followed Allstate in calling for an increased federal role in coastal catastrophe policy. The Travelers/Nationwide proposal had two significant aspects. First, the proposal sought to spread big windstorm insurance risk across state borders by having it federally regulated.⁹⁴

⁸⁷ Lehmann, *supra* note 85.

⁸⁸ Press Release, Ins. Info. Inst., Reforms Needed to Prepare for Major Catastrophes, Insurance CEOs Tell Forum, <http://www.iii.org/media/updates/archive/press.748575/> (Jan.10, 2006).

⁸⁹ Homeowners Insurance Protection Act of 2005, H.R. 4366, 109th Cong. § 4 (2005).

⁹⁰ *Id.* § 7(b)(6)(B).

⁹¹ *Id.* § 7(a).

⁹² *Id.* § 7(b)(3).

⁹³ David Dankwa, *Allstate Takes Hot Seat as Industry Debates Catastrophe-Fund Options*, BESTWIRE Jan. 16, 2006, available at <http://www3.ambest.com/Frames/FrameServer.asp?AltSrc=23&Tab=1&Site=news&refnum=80532>.

⁹⁴ The Travelers Cos., *Insurance Agents, Travelers, Nationwide Urge Coastal Wind Policy*, INS. J., July 16, 2008, <http://www.insurancejournal.com/news/national/2008/07/16/91909.html>.

Second, similar to the Allstate proposal, the Travelers/Nationwide plan called for a federal reinsurance mechanism for extreme weather events.⁹⁵

The first notable aspect of the Travelers/Nationwide plan was that it called for a federal role in promoting a plan to spread coastal windstorm risks across state borders.⁹⁶ As opposed to setting up individual catastrophe funds state by state, the Travelers/Nationwide plan encouraged a system of inter-state zones: the Gulf, Florida, Southeast and Northeast.⁹⁷ Pursuant to the Travelers/Nationwide plan the federal government would not have a financial role in maintaining these pools, but would oversee the wind underwriting by private insurers.⁹⁸ To do this, the Travelers/Nationwide plan would require the creation of an independent federal agency to regulate rates and set uniform rules from Texas to Maine for the wind coverage portion of a homeowner's policy, while the states would continue to regulate the other portions of a homeowner's policy.⁹⁹

With regard to premiums, the Travelers/Nationwide proposal requires setting "risk-based and actuarially sound rates using approved standards and certified windstorm risk models approved by the federal commission."¹⁰⁰ However, the drafters of the Travelers/Nationwide proposal recognized that risk-based premiums may be out of reach for many coastal residents.¹⁰¹ In response, the proposal provided for temporary transitional subsidies lasting between 10 to 15 years.¹⁰² To this end, coastal residents who were unable to afford coverage would receive tax credits to help them afford premiums.¹⁰³ The credits would be funded by tax surcharges imposed on those coastal residents able to afford coverage.¹⁰⁴

The Travelers/Nationwide federal reinsurance mechanism functioned much like the Allstate backstop. The federal backstop required regions set up catastrophe funds and pay premiums to a federal program in

⁹⁵ *Id.*

⁹⁶ *See id.*

⁹⁷ THE TRAVELERS INST., TRAVELERS COASTAL WIND ZONE PLAN 5 (2009), http://www.travelers.com/iwcm/Trv/docs/TRV_Coastal_Wind_Zone_Web.pdf.

⁹⁸ *Id.* at 6.

⁹⁹ *Id.* at 5.

¹⁰⁰ *Id.*

¹⁰¹ Jay S. Fishman, *Before the Next 'Big One' Hits*, WALL ST. J., Aug. 27, 2007, at A10.

¹⁰² *Id.*

¹⁰³ THE TRAVELERS INST., *supra* note 97, at 6.

¹⁰⁴ *Id.*

return for reinsurance that would be used to support the regional fund.¹⁰⁵ The threshold for the Travelers/Nationwide federal reinsurance mechanism has yet to be finalized as it is stated to be for “extreme events (such as hurricanes causing losses several times greater than those arising out of Hurricane Katrina).”¹⁰⁶

D. THE HARTFORD’S PLAN: HOMEOWNER CATASTROPHE SAVINGS ACCOUNTS, MANDATORY FLOOD INSURANCE, AND FEDERAL BACKSTOPPING

The Hartford announced its public-private natural catastrophe plan in the summer of 2008.¹⁰⁷ The Hartford’s Coastal Catastrophe Partnership Plan was “designed to deal with the looming economic crisis posed by a major hurricane.”¹⁰⁸ The plan was to occupy a middle ground of government involvement. Hartford CEO, Romani Ayer, said the Coastal Catastrophe Plan was designed to navigate between the “‘socialistic’ efforts of some in Congress, and reinsurers who believe all catastrophe protection should be left to the private market.”¹⁰⁹ It was the last of the plans put forth by a major private insurer and in many ways it was the most detailed. The plan had three facets that merit discussion. First, it called for tax-deferred catastrophe savings accounts.¹¹⁰ Second, the Hartford proposal required

¹⁰⁵ See *supra* Part III.B.

¹⁰⁶ The Travelers Cos., *supra* note 94.

¹⁰⁷ See The Hartford Fin. Serv. Group Inc., *The Hartford Unveils National Coastal Home Insurance Plan*, INS. J., July 31, 2008, <http://www.insurancejournal.com/news/national/2008/07/31/92393.htm> (noting that the plan will include subsidized premiums for homeowners and a federal reinsurance program).

¹⁰⁸ Daniel Hays, *Hartford Proposes National Cat Plan Featuring Federal Backstop, IRA-Type Fund*, NAT’L UNDERWRITER – PROP. & CAS. INS., Aug. 4, 2008, available at 2008 WLNR 14506223 (quotation omitted).

¹⁰⁹ *Id.*

¹¹⁰ See RAMANI AYER, THE HARTFORD FIN. SERV. GROUP, BUILDING A NATURAL CATASTROPHE SOLUTION 5 (2007), http://www.coastalpartnership.org/documents/CCP_Plan_Outline.pdf (noting that the Hartford proposes “the creation of IRA-like savings vehicles – perhaps called ‘supplemental catastrophic security accounts’”).

mandatory homeowner's flood insurance.¹¹¹ Finally, the proposal called for a Federal reinsurance backstop.¹¹²

Perhaps the most striking aspect of the Hartford's plan was that it encouraged the creation of IRA-like savings vehicles dubbed "supplemental catastrophic security accounts."¹¹³ The accounts would have permitted property owners in coastal regions to establish tax-deferred reserves to pay for insurance and other expenses related to disasters.¹¹⁴ Specifically, the idea is that allowing homeowners to use tax-deferred dollars to pay for catastrophe insurance could induce more people to buy it.¹¹⁵

The second aspect of the Hartford plan required coastal homeowners to buy flood insurance.¹¹⁶ The Hartford's plan would achieve this by allowing for the alternative policy options of either requiring certification of flood insurance coverage through current programs such as the National Flood Insurance Program or by inserting flood coverage as part of the standard homeowner's policy.¹¹⁷

Finally, much like the Allstate and Travelers/Nationwide plans, the Hartford proposal also included a federal reinsurance backstop for existing state and regional funds.¹¹⁸ The Hartford backstop would be triggered upon the occurrence of a 1-in-100 year loss.¹¹⁹ Thus, the key difference between the previous proposed federal reinsurance backstops and the Hartford's plan lies in the fact that the latter seems to anticipate the most amount of government intervention.¹²⁰

¹¹¹ *Id.*

¹¹² *Id.* at 6.

¹¹³ Hays, *supra* note 108.

¹¹⁴ *Id.*

¹¹⁵ U.S. GOV'T ACCOUNTABILITY OFFICE, GAO-08-7 PUBLIC POLICY OPTIONS FOR CHANGING THE FEDERAL ROLE IN NATURAL CATASTROPHE INSURANCE 35 (2007).

¹¹⁶ Ayer, *supra* note 110, at 5.

¹¹⁷ *Id.*

¹¹⁸ *Id.* at 6.

¹¹⁹ *Id.*

¹²⁰ *Id.*

E. THE HOUSE VERSION OF THE HOMEOWNERS' DEFENSE ACT OF 2007: BONDS, LOANS, REINSURANCE

In response to the 2005 hurricane season U.S. House Reps. Tim Mahoney and Ron Klein co-authored a bill designed “to bring relief to property owners struggling with the affordability and availability of homeowners insurance.”¹²¹ The resultant Homeowners' Defense Act of 2007 sought to, “assist state-sponsored insurance programs on covering losses from natural disasters.”¹²² Essentially, the proposal offered by the Homeowners' Defense Act of 2007 had three major features.

First, the proposal would essentially securitize coastal risk. This aspect of the proposal would have established a federally overseen mechanism that permitted multiple states to join together to help pay for each others' disaster costs and then transfer those costs to the private markets through catastrophe bonds and reinsurance contracts.¹²³ The program would have accomplished the private market risk transfer via the establishment of a central National Catastrophe Risk Consortium tasked to, work with states to create an inventory of catastrophe risk obligations held by state reinsurance funds and issue securities linked to the catastrophe risk insured in capital markets.¹²⁴

Second, the proposal would have created a federal reinsurance program similar to the type championed by both the Allstate and Travelers/Nationwide plans to backstop existing state insurance pools.¹²⁵ The actual reinsurance mechanism would be fiscally backed by the Federal Natural Catastrophe Reinsurance Fund.¹²⁶ This fund would be directed towards covering contract payments for losses resulting from the federal reinsurance program.¹²⁷ The fund would be financially supported by, amounts received annually from the sale of reinsurance contracts, amounts earned on investments, and appropriations.¹²⁸

¹²¹ See Amie Parnes, *Bill Seeks to Provide Homeowners Insurance Relief*, NAPLES NEWS, Aug. 3, 2007, available at http://www.naplesnews.com/news/2007/aug/03/bill_seeks_provide_homeowners_insurance_relief/.

¹²² *Id.*

¹²³ See Homeowners' Defense Act of 2007, H.R. 3355, 110th Cong. §§ 2(b), 101(d), 102 (2007).

¹²⁴ See *id.* § 102.

¹²⁵ See *id.* §§ 301-303.

¹²⁶ *Id.* § 305.

¹²⁷ *Id.*

¹²⁸ H.R. 3355, § 305.

Finally, the proposal would have provided loans to state insurance programs.¹²⁹ This aspect of the proposal called for the creation of a National Homeowners' Insurance Stabilization Program, which would have provided low-interest federal loans directly to qualified state insurance programs in order to, "ensure the solvency of such programs, to improve the availability and affordability of homeowners' insurance, to incent risk transfer to the private capital and reinsurance markets and to spread the risk of catastrophic financial loss resulting from natural disasters."¹³⁰

F. NAIC'S COMPREHENSIVE NATIONAL CATASTROPHE PLAN:
CATASTROPHE RESERVING, ALL PERILS POLICIES AND
TENTATIVE SUPPORT FOR FEDERAL REINSURANCE

The National Association of Insurance Commissioners' Catastrophe Insurance Working Group developed its national natural disaster plan in 2005.¹³¹ The NAIC plan addresses natural catastrophe coastal risk in terms of three main "layers."¹³² The first layer corresponds to private market solutions and calls for a mandatory all-perils residential policy and insurance company catastrophe reserving.¹³³ Dealing primarily with state level issues, the second "layer" calls for the establishment of a uniform system of state catastrophe funds.¹³⁴ While corresponding to federal level policy, the third layer tentatively supports further examination of creating a federally overseen catastrophe reinsurance mechanism.¹³⁵

Under the all-perils component of the NAIC plan, the idea would be to create a homeowner insurance policy that would provide coverage against all types of natural catastrophes.¹³⁶ The NAIC posits that, private insurers would provide coverage against all perils in a single standard homeowner's policy that would reflect each property owner's risk of loss

¹²⁹ *See id.* § 202.

¹³⁰ *Id.* § 201.

¹³¹ NAT'L ASS'N OF INS. COMM'RS, NATURAL CATASTROPHE RISK: CREATING A COMPREHENSIVE NATIONAL PLAN 2 (Version 15a 2009), http://www.naic.org/documents/committees_c_090615_nat_catastrophe_paper_draft.pdf.

¹³² *See id.* at 7.

¹³³ *Id.* at 7-8.

¹³⁴ *Id.* at 9.

¹³⁵ *Id.* at 10.

¹³⁶ NAT'L ASS'N OF INS. COMM'RS, *supra* note 131, at 8.

due to natural disaster.¹³⁷ At this point coastal the NAIC plan would not require coastal residents purchase an all perils policy, however, the availability of such a policy would be mandatory.¹³⁸

Pursuant to the catastrophe reserving aspect of the NAIC plan, private insurers would be permitted to establish tax-deferred reserves to pay expenses related to natural coastal disasters in order to further expand the financial base available for underwriting catastrophe risk.¹³⁹ The idea underlying catastrophe reserving is that with an additional means of building cash reserves insurance companies would be more able and willing to underwrite policies. At present, tax-exempt reserves would necessarily require amending the US tax code.¹⁴⁰

Within the second layer of the NAIC's plan, each state would be asked to decide whether its exposure to coastal natural catastrophes warrants creating a catastrophe fund or whether the private market has the capacity to provide adequate coverage without additional funding.¹⁴¹ If established the state catastrophe funds would be responsible for creating and managing the insurance capacity of their respective regions.¹⁴² The funds would have the discretion to implement their own operating structures to best fit their particular needs.¹⁴³ This would entail defining catastrophic loss thresholds, determining appropriate private insurers/state fund retention amounts, and ensuring that premium rates are actuarially sound.¹⁴⁴

The third and final layer of the NAIC plan consists of cautious support for a federal reinsurance role.¹⁴⁵ Although the NAIC acknowledges that a federal reinsurance program seems to be a potential solution, the Commission also acknowledges that the debate over whether federal reinsurance involvement is necessary is still ongoing.¹⁴⁶ Beyond this, the specific character of the federal reinsurance mechanism the NAIC would support is still not clear.

¹³⁷ *Id.*

¹³⁸ *Id.* at 9.

¹³⁹ NAT'L ASS'N OF INS. COMM'RS, *supra* note 131, at 6.

¹⁴⁰ *Id.*

¹⁴¹ *Id.* at 9.

¹⁴² *Id.*

¹⁴³ *Id.*

¹⁴⁴ *Id.*

¹⁴⁵ *See id.* at 10-11.

¹⁴⁶ *Id.* at 11.

V. ANALYSIS: CAPACITY, PRIVATE MARKET DISPLACEMENT, CROSS-SUBSIDIZATION

The plans proposed offer a myriad of policy options. Keeping track of them is a task in itself. One useful way of evaluating these current coastal risk proposals is to look at how well they deal with certain of evaluative criteria: private insurance industry capacity, private market displacement, and cross subsidization.

Capacity means “the ability of an insurance company... to pay claims in the event of a loss.”¹⁴⁷ In this context, capacity specifically refers to the private insurance industry’s ability to cover the insured costs of property damage wrought by a coastal storm. With respect to capacity, the analysis of each plan concerns two issues. First, the extent each plan suggests the private market presently has adequate capacity to handle coastal risk. Second, how each plan seeks to expand private market capacity to handle growing future coastal risk.

Similarly, private market displacement refers to the concern that federal government activity in coastal natural catastrophe insurance will hinder the private market actors already involved.¹⁴⁸ The focus of this inquiry is on how each plan seeks to minimize the displacement of the private coastal risk insurance markets. That is, whether any of the plans utilize policy options or forms of government intervention that will either encourage private market participation or, at least, not completely crowd out private insurers from covering coastal risk.

Finally, in this context cross subsidization refers to the situation whereby the government covers the exposure property owners who choose to live in high risk coastal locations take on by spreading the cost of that risk to taxpayers who do not reside in these areas.¹⁴⁹ The issue here is whether and to what extent each plan contains policy tools that could help localize the costs of owning property in storm-prone coastal areas.

¹⁴⁷ NAT’L ASS’N OF INS. COMM’RS, *supra* note 131, at 11.

¹⁴⁸ See Press Release, Prop. Cas. Insurers of Am., Windstorm Coverage Widely Available, Federal Involvement Unnecessary (April 21, 2008) (<http://www.pciaa.org> (follow “MediaCenter: News Releases: By Date:” then scroll to “April 21, 2008”)).

¹⁴⁹ See CHARLES PERROW, *THE NEXT CATASTROPHE: REDUCING OUR VULNERABILITIES TO NATURAL, INDUSTRIAL, AND TERRORIST DISASTERS* 65 (Princeton 2007).

A. CAPACITY

The nation's two largest homeowners' insurers, State Farm and Allstate, argue that some natural disasters are too large for the private insurance market to handle.¹⁵⁰ State Farm spokesman Jeff McCollum argues that "[c]ommercially available insurance is designed for localized disasters, not 'mega-catastrophes'."¹⁵¹ He goes on to note that, "[i]nsurance was never designed to cover a catastrophe as big or bigger than Katrina."¹⁵² Similarly, a spokesperson for the Allstate remarked that, "[o]ur view is that there are some events that have the potential to be so large as to exceed the capabilities of the insurance industry, as well as the funding and financing capability of individual states."¹⁵³

Similarly, some state insurance commissioners recognize the capacity issue as a potential stumbling block to completely private market solutions to coastal risk. Florida Insurance Commissioner Kevin McCarty told the House Financial Services Committee that in his view, "large natural catastrophes are a national economic problem, not simply a local insurance problem."¹⁵⁴ McCarty went on to argue that, "[b]ecause of the absolute size of the economic losses that are possible due to hurricanes in Florida, the private market, public mechanisms, and even the states themselves simply do not have sufficient capacity to provide recovery from a truly mega-catastrophic hurricane event."¹⁵⁵ As a result, McCarty concluded that, "Congress and the states need to work together to develop a comprehensive plan to better manage and mitigate the natural catastrophic events of tomorrow."¹⁵⁶

¹⁵⁰ Larry Lipman, *Catastrophe Insurance Spawns Storm of Debate*, PALM BEACH POST, April 8, 2007 at A13.

¹⁵¹ *Id.*

¹⁵² *Id.*

¹⁵³ Joel Garreau, *A Dream Blown Away: Climate Change Already Has a Chilling Effect on Where Americans Can Build Their Homes*, WASH. POST, Dec. 2, 2006 at C01.

¹⁵⁴ Lipman, *supra* note 150.

¹⁵⁵ Associated Press, *Citizens Property Insurance Increases Upset Policyholders in Florida*, INSURANCE JOURNAL, <http://www.insurancejournal.com/news/southeast/2006/05/15/68332.htm> (last visited Oct. 12, 2009).

¹⁵⁶ *Id.*

1. The Coalition of Americans for Smart Natural Catastrophe Policy

Not all think private market capacity is such a problem. The proposal put forth by the Coalition of Americans for Smart Natural Catastrophe Policy is concentrated almost entirely on policy options aimed at risk mitigation as opposed to augmenting private market capacity.¹⁵⁷ At base, the Coalition's position consists of supporting the homeowner mitigation grants and the gradual phasing out of subsidized National Flood Insurance Program premiums.¹⁵⁸ Implicit in this combination of policy choices is the assumption that, if implemented, an increased government role in coastal catastrophe insurance is unnecessary because the private market has the capacity to cope with such risks.

As a result of this assumption the Coalition's plan lacks policy aimed at expanding private market capacity. While policy that encourages mitigation can be viewed as an indirect means of decreasing coastal risk relative to market capacity, it does not directly address the capacity issue. At this point the Coalition's plan offers no policy innovations aimed at aiding private insurers increase their own ability to meet the needs of the growing coastal catastrophe insurance market. This sort of position is problematic insofar as the underlying assumption that present market capacity is sufficient is debatable.¹⁵⁹ Further still, with global climate changes in mind market capacity will need to keep pace with coastal catastrophes of increasing severity. In this way the Coalition's plan does not offer an adequate response to the private market capacity issue.

2. Allstate, Travelers', and the Hartford

In contrast to the position taken in the plan put forth by the Coalition of Americans for Smart Natural Catastrophe Policy, the stances taken by Allstate, Travelers', and the Hartford all suggest significantly less confidence in private industry capacity to handle coastal catastrophic risk. Common to each of these three private insurer proposals are federal reinsurance programs.¹⁶⁰ Acknowledgement that a federal reinsurance backstop is necessary clearly serves as an avowal that the private

¹⁵⁷ See *supra* notes 61-83 and accompanying text.

¹⁵⁸ *Id.*

¹⁵⁹ See *supra* notes 150-56 and accompanying text.

¹⁶⁰ See *supra* notes 104-6, 118-20 and accompanying text.

reinsurance industry does not have the capacity to handle major coastal catastrophe.

Similarly, the plans put forth by all three of these private industry leaders signal a perceived lack of private market capacity to cover catastrophic coastal loss in that all three share common support for a continued role for state and regional government overseen insurance pools.¹⁶¹ Although insurance pools essentially spread risk, in part, to prevent insolvency,¹⁶² the history of government organized coastal risk insurance pools is such that after a catastrophe has struck they are often funded in part by the public sector.¹⁶³ In this sense, it is important to note that each of the private industry backed proposals requires the erection of some form of a system of government backed insurance pools.

Interestingly, because of their reliance on federal reinsurance mechanisms, the private industry response to the capacity issue, as seen by the proposals put forth by Allstate, Travelers', The Hartford and State Farm, have less to do with growing private industry capacity than enlisting the federal government as a reinsurance safety net. Instead of approaches that would augment private industry ability to cover catastrophic risk, such as catastrophe reserving, the plans set forth by the three leading insurers look to the federal government as a potential backstop as a primary means of shifting the exposure of coastal loss. This approach is undesirable because over-reliance on federal support implicates creating large liabilities for the federal government and, as will be discussed shortly, unfairly cross-subsidizing coastal residents at the expense of non-coastal residents.

3. House Version of the Homeowners Defense Act of 2007

The House version of the Homeowners' Defense Act of 2007 proposal evidences significant reservation as to the ability of private markets to handle catastrophic coastal risk. The Act would both create a

¹⁶¹ See *supra* notes 87-88, 118 and accompanying text.

¹⁶² See generally, John Pollner, *Catastrophe Risk Management: Using Alternative Risk Financing and Insurance Pooling Mechanisms* 87 (The World Bank, Working Paper No. WPS2560, 2001), available at <http://www.worldbank.org/reference/> (follow "Documents and Reports" hyperlink; then search "WPS2560"; then follow "Catastrophe Risk Management..." hyperlink).

¹⁶³ *Id.* at 85-86.

federal reinsurance presence in already established state catastrophe pools and provide low-interest federal loans directly to existing state reinsurance programs.¹⁶⁴ However, through its National Catastrophe Risk Consortium program,¹⁶⁵ the Act provides a more complicated overall take on the private market capacity expansion issue.

The Homeowners' Defense Act is the first of the plans examined to offer a policy option directly aimed at expanding the private market's ability to handle catastrophic coastal risk. The Consortium component of the Act seeks to transfer coastal risk to the private financial market.¹⁶⁶ Through the creation of financial instruments linked to catastrophe risks insured or reinsured by members of the Consortium, this aspect of the Homeowners' Defense Act sought to create and encourage a securities market linked to coastal risk.¹⁶⁷ This approach essentially uses a federally overseen mechanism to funnel coastal risk from state and regional pools to private securities markets.¹⁶⁸ In theory, by tapping into the private capital markets this approach has the potential to drastically increase private sector capacity to handle coastal risk. As a result, this approach offers an innovative policy option geared towards expanding private sector capacity.

4. The National Association of Insurance Commissioners

Due to its limited tentative support for some sort of federal reinsurance mechanism,¹⁶⁹ NAIC's Comprehensive National Catastrophe Plan is marked by a relatively limited direct federal intervention that assumes the private markets' have significant ability to handle coastal catastrophic storm related loss. Moreover, the NAIC proposal contains a policy tool in catastrophe reserving that may actually expand the private market's capacity to handle coastal risk.

The catastrophe reserving mechanism contained in the NAIC proposal seeks to expand the ability of private market insurers to underwrite coastal risk by allowing them to build limited tax-deferred pre-

¹⁶⁴ See *supra* notes 123-30 and accompanying text.

¹⁶⁵ See *supra* notes 123-24 and accompanying text.

¹⁶⁶ *Id.*

¹⁶⁷ *Id.*

¹⁶⁸ See Homeowners' Defense Act of 2007, H.R. 3355, 110th Cong. §102(6) (2007).

¹⁶⁹ See *supra* notes 145-46 and accompanying text.

catastrophe capital bases.¹⁷⁰ This sort of option represents the most direct means of increasing private market capacity to deal with coastal storms.

B. PRIVATE MARKET DISPLACEMENT

There is a line of reasoning with considerable traction within the coastal insurance industry that the greatest threat to the property insurance market, "is not the force of hurricane winds but legislation and regulations that displace available private capital or make it economically unfeasible for private companies to operate in coastal markets."¹⁷¹ Similarly, there is a movement within the industry that would prefer to see a halt to federal expansion into the coastal catastrophe market.¹⁷² Frank Nutter, president of the Reinsurance Association of America argues that expanding the current federal role "would displace a vibrant [private] reinsurance market to the detriment and cost of the U.S. taxpayers."¹⁷³

1. The Coalition of Americans for Smart Natural Catastrophe Policy

Clearly, the mitigation based proposal put forth by the Coalition of Americans for Smart Natural Catastrophe Policy would require the least amount of private market displacement. Perhaps overestimating the capacity of the private market, the Coalition proposal requires no federal reinsurance program and does not require any other new federal intervention beyond modest homeowner mitigation grants and loans.¹⁷⁴ In this way, the Coalition plan is at the far end of the spectrum occupied by proposals that would cause little private market displacement.

¹⁷⁰ See *supra* note 139 and accompanying text.

¹⁷¹ Larry Lipman, *Catastrophe Insurance Spawns Storm of Debate*, PALM BEACH POST, Apr. 8, 2007, at A13.

¹⁷² See *id.*

¹⁷³ Press Release, Reinsurance Ass'n of Am., RAA Testifies Before Senate Banking, Housing, and Urban Affairs Committee (Apr. 11, 2007), www.reinsurance.org/files/public/pr41107.pdf.¹⁷⁴ See *supra* notes 61-83 and accompanying text.

¹⁷⁴ See *supra* notes 61-83 and accompanying text.

2. Allstate, Travelers', and the Hartford

In contrast to the Coalition plan, the proposals put forth by Allstate, Travelers', and the Hartford would all anticipate significant private market displacement. All three proposals require a federal reinsurance mechanism specifically designed to support state insurance pools.¹⁷⁵ The displacement implications for this type of federal reinsurance mechanism are clear. If even eligible, private reinsurers would be forced to compete with a federal reinsurance program in providing reinsurance to the pools.¹⁷⁶ This scenario has the potential to cause significant private market disruption.

However, this reinsurance displacement should be viewed in broader context. A federal reinsurance option could lead to greater participation from private primary insurers.¹⁷⁷ In this sense, any displacement of private reinsurers may be offset or even surpassed by greater primary insurer willingness to underwrite coastal risk.

3. Homeowners Defense Act of 2007

The response to the private market displacement question offered by the proposal embodied in the House version of the Homeowners Defense Act of 2007 is unique. Although the plan advocates government-run insurance pools and a federal reinsurance fund,¹⁷⁸ the role of the Consortium in transferring risk to the private securities market complicates the private market displacement analysis. In the narrow sense, the federal reinsurance fund would certainly displace private reinsurers.¹⁷⁹ However, the net effect of the Consortium's risk transferring scheme may potentially offset some of the private market reinsurance displacement by facilitating risk transfer to the private securities market.

4. The National Association of Insurance Commissioners

The NAIC program, consisting of private industry catastrophe reserving and mandatory offer of an all-perils policy¹⁸⁰ suggests expanding

¹⁷⁵ See *supra* notes 87-89, 118 and accompanying text.

¹⁷⁶ U.S. GOV'T ACCOUNTABILITY OFFICE, *supra* note 115, at 34-35.

¹⁷⁷ *Id.*

¹⁷⁸ See *supra* notes 123-30 and accompanying text.

¹⁷⁹ U.S. GOV'T ACCOUNTABILITY OFFICE, *supra* note 115, at 33.

¹⁸⁰ See *supra* notes 136-39.

the role of private insurers as opposed to displacing them. The catastrophe reserving policy option would potentially increase private insurer capacity to handle large coastal risk and thereby increase its ability to underwrite risk.¹⁸¹ In this way, catastrophe reserving may prevent private market displacement by fortifying its ability to handle coastal risk. Similarly, the all-perils portion of the NAIC plan explicitly seeks to have private insurers provide comprehensive storm coverage in a single standard homeowners' policy.¹⁸² Here, the effort to expand private activity in the context of coastal risk is quite clear. An all-perils policy program such as the one endorsed by the NAIC may have the opposite impact of actually further entrenching the private market in fielding coastal risk.¹⁸³

C. SUBSIDIZING COASTAL PROPERTY OWNERS

Federal intervention in coastal risk could result in significant public subsidies to property owners who make the choice to construct or maintain a home or business in a coastal area. The current National Flood Insurance Program is a prime example of a program that facilitates this dynamic.¹⁸⁴ Ultimately the subsidized rates that many coastal property owners pay are made possible by tax dollars of many non-coastal residents.¹⁸⁵ The term that describes this occurrence is "cross-subsidization."

1. The Coalition of Americans for Smart Natural Catastrophe Policy

As for current proposals that attempt to minimize such subsidization, the Coalition of Americans for Smart Natural Catastrophe Policy's mitigation based plan confronts the issue in blunt terms. In effect, the Coalition's platform is built on the fear that if its policyholders are not adopted citizens in regions who face little coastal storm risk, they would be

¹⁸¹ NAT'L ASS'N OF INS. COMM'RS, *supra* note 131, at 7; U.S. GOV'T ACCOUNTABILITY OFFICE, *supra* note 111, at 34-35.

¹⁸² See NAT'L ASS'N OF INS. COMM'RS, *supra* note 131, at 8.

¹⁸³ See U.S. GOV'T ACCOUNTABILITY OFFICE, *supra* note 115, at 34-35.

¹⁸⁴ See *The Nat'l Flood Ins. and Repetitive Loss Properties: Hearing Before the Subcomm. on Housing and Cmty. Opportunity of the H. Comm. on Fin. Serv.*, 107th Cong. 3 (2001) (statement of Rep. Bereuter, Member House Subcomm. on Housing and Cmty. Opportunity).

¹⁸⁵ See *id.* at 3-4.

subsidizing residents of areas where hurricanes are a constant threat.¹⁸⁶ The Coalition urges no federal involvement in insuring coastal risk beyond the existing National Flood Insurance Program, and even then would have the subsidized rates that characterize much of the program phased out within a matter of years.¹⁸⁷ In addition, the Coalition successfully fought adding windstorm coverage to the National Flood Insurance Program because it would have the impact of transferring more coastal risk from coastal property owners to non-coastal homeowners by way of the federal government.¹⁸⁸ For these reasons, the Coalition's plan deals most effectively with curbing cross-subsidization.

2. Allstate, Travelers', and the Hartford

At the other end of the spectrum the Allstate, Travelers/Nationwide, and Hartford plans would all have non-coastal residents subsidize coastal property owners. The federal reinsurance program that is common to three plans would likely rely on post-funding appropriations in the event of a coastal catastrophe.¹⁸⁹ In the context of a federal reinsurance mechanism, post-funding appropriations would likely be drawn from federal tax dollars and would therefore represent a significant form of cross-subsidies.¹⁹⁰

In theory, a federal reinsurance mechanism does not necessarily have to lead to federal subsidies. If the proposed federal reinsurance programs were operated according to actuarially sound principles it should be self-sufficient and not require tax derived post-funding in the event of a catastrophe. However, given the history of programs such as the National Flood Insurance Program it is difficult to see a government devised insurance scheme not relying on post-funding measures.¹⁹¹ If the National Flood Insurance Program is any indication, a federal reinsurance mechanism will not be self-sufficient and the program will likely rely on the liberal use of cross-subsidized post-event appropriations.¹⁹² For this

¹⁸⁶ See Smarter Safer, Statement of Principles, <http://www.smartersafer.org/about-us/statement-of-principles> (last visited Oct. 8, 2009).

¹⁸⁷ *Id.*; see *supra* notes 62-72 and accompanying text.

¹⁸⁸ See *supra* notes 62-72 and accompanying text.

¹⁸⁹ The history of other government managed coastal risk programs supports this assertion. See *supra* notes 45, 57-58 and accompanying text.

¹⁹⁰ See, e.g., *supra* note 45 and accompanying text.

¹⁹¹ See *supra* notes 42-46 and accompanying text.

¹⁹² See *supra* notes 42-46 and accompanying text.

reason the Allstate/Nationwide, Travelers', and Hartford plans do not successfully minimize cross-subsidization.

3. Homeowners Defense Act of 2007

On balance the three prongs of the program offered in the Homeowners' Defense Act of 2007 favor significant cross-subsidization. Clearly, the plan's federal reinsurance backstop and related reinsurance fund could lead to non-coastal taxpayers subsidizing their coastal counterparts.¹⁹³ Similarly, the federal stabilization program designed to provide low-interest federal loans directly to qualified state reinsurance programs would also lead to government facilitated cross-subsidization.¹⁹⁴ However, on the other side of the scale, the risk transferring National Catastrophe Risk Consortium has the potential to offset some of the other two programs insofar as it can tap into private sector capacity to take on coastal risk and thereby eliminate some need for utilization of the other two programs.¹⁹⁵ Since the potential success of the Consortium is difficult to gauge it is ultimately hard to tell how effective it will be in curbing the need to utilize the proposals other cross-subsidy funded programs.

4. The National Association of Insurance Commissioners

The NAIC proposal contains policy measures that localize the risk of catastrophic coastal storms, but also some that may lead to cross-subsidization of coastal risk. By withholding unreserved support for a federal reinsurance system, the NAIC plan, in its present form would essentially only allow for catastrophe reserving and all-perils homeowners' policies to be provided by private companies.¹⁹⁶ These two remaining aspects of the NAIC proposal would have very different impacts in terms of cross-subsidization. While catastrophe reserving would allow private insurers to build up capital reserves to better deal with future catastrophic events, it would not directly involve having noncoastal tax-payers subsidize coastal residents. However, the all-perils homeowners' policy aspect of the NAIC program may indirectly lead to moderate cross-subsidization. While

¹⁹³ See *supra* notes 125-28, 189-92 and accompanying text.

¹⁹⁴ See KING, *supra* note 44 at 5; see also *supra* notes 129-31 and accompanying text.

¹⁹⁵ See *supra* notes 128-29 and accompanying text.

¹⁹⁶ See *supra* notes 139-46 and accompanying text.

the NAIC would require these policies be written through private insurers, the all-perils option could lead to government subsidies for low-income property owners to afford such coverage.¹⁹⁷

VI. CONCLUSION

Some estimate that the destructive potential of tropical storms in the North Atlantic has increased dramatically since the 1970s.¹⁹⁸ Alarming enough on its own, this weather trend coincides with a time in which many parts of the coastal U.S. are becoming more densely populated than ever.¹⁹⁹ Add to this an average annual growth rate that will lead to the doubling of the total covered coastal value every decade,²⁰⁰ and it becomes clear that the coastal catastrophe policy framework needs reevaluation.

Even before the wakeup call of the monumental 2005 hurricane season policymakers and private industry leaders were beginning to rethink the coastal risk landscape. The proposals examined in this paper represent a fairly wide range of thought on the federal role in insuring coastal risk. Evaluated by way of the capacity, private market displacement, cross subsidization criteria laid out here, none of the plans are perfect. Instead, each proposal is flawed while a few exhibit promising aspects.

In advocating systems of federally reinsured state run catastrophe pools, the proposals made by three of the largest property and casualty insurance companies, could result in a considerable expansion of federal involvement. Federal reinsurance would not directly expand private market capacity, but would instead shift coastal risk exposure to the federal government. Similarly, this expansion would come at the cost of private market displacement and significant cross-subsidization.

In contrast, the plan put forth by the Coalition of Americans for Smart Natural Catastrophe Policy is characterized almost exclusively by risk mitigation efforts and would seemingly halt any additional federal involvement in managing coastal risk. While the absence of a federal reinsurance scheme is laudable insofar as it would curb private market displacement and significant cross-subsidization, the Coalition approach is

¹⁹⁷ U.S. GOV'T ACCOUNTABILITY OFFICE, *supra* note 115, at 34-35.

¹⁹⁸ See Kerry Emanuel, *Increasing Destructiveness of Tropical Cyclones Over the Past 30 Years*, 436 NATURE 686 (2005).

¹⁹⁹ See *supra* notes 9-22 and accompanying text

²⁰⁰ See *supra* note 24 and accompanying text.

problematic because places unwarranted confidence in the capacity of the private market and existing coastal risk framework to handle future storms.

Containing measures aimed at expanding private insurer capacity to take on coastal risk and only reserved support for a federal reinsurance mechanism, the proposal made by the National Association of Insurance Commissioners offers a possible approach with promise. The NAIC plan's support of insurance company catastrophe reserving may potentially be effective in increasing private market capacity to handle coastal catastrophe risk. This capacity expanding upside might offset the private market displacement and cross-subsidization the NAIC plan's potential reinsurance mechanism would cause.

Finally, the House version of the Homeowners' Defense Act of 2007 would have created an innovative mechanism to transfer coastal risk to private capital markets, but would also have provided for federal loans to state insurance pools and erected a federal reinsurance backstop. The innovative capacity expansion upside that the Defense Act's risk transferring mechanism provides is undercut by the amount of cross-subsidization and private market displacement the proposal's federal reinsurance and loan programs could lead to.

In the months ahead other proposals will be made and eventually a comprehensive program will be settled upon. It was the goal of this modest examination to add to the discussion that will lead to an ultimate decision by examining and comparing some of the high profile options with particular attention to certain concerns: private market capacity to field catastrophic coastal risk, continuing private market role in coastal catastrophic risk, and cross-subsidization of coastal risk.

**CONNECTICUT INSURANCE
LAW JOURNAL**

The Connecticut Insurance Law Journal is published at least twice per year. Domestic subscriptions are \$28.00 *per year*, international subscriptions are \$32.00 *per year*. Back issues and back volumes are available (for any volume up to and including Volume 12) for \$15.00 per issue or \$25.00 *per volume*; international orders add \$3.00 *per issue* or \$5.00 *per volume*.

Absent timely notice of termination, subscriptions are renewed automatically. To subscribe to the Insurance Journal or order back issues online, please visit our website located at www.insurancejournal.org, or you may fill out the form below and mail to:

Pat Carbray,
Connecticut Insurance Law Journal,
65 Elizabeth Street,
Hartford, CT 06105;
subscriptions@insurancejournal.org.

Please Print below and complete reverse:

Contact Name: _____

Contact Title: _____

Organization: _____

Department: _____

Address: _____

City: _____ State: _____ Zip Code: _____

Phone: _____ Fax: _____

E-Mail: _____

Additional Comments: _____

Check all that apply:

- I would like to order (number) _____ subscriptions at \$28.00 per subscription (\$32.00 for international orders).
- I would like to order (number) _____ complete set(s) of back issues (Volumes 1-12) at \$150.00 per set (\$165.00 for international orders).
- I would like to order the following back volumes or back issues at \$15.00 per issue, \$25.00 per volume (international orders are \$18.00 per issue and \$30.00 per volume). Please list volume, issue and quantity for each.

Examples: 4 copies, Volume 3, Issue 2
2 copies - Volume 3
