ISSN 1936-5349 (print)
ISSN 1936-5357 (online)

HARVARD

JOHN M. OLIN CENTER FOR LAW, ECONOMICS, AND BUSINESS

TESTING FOR ADVERSE SELECTION IN INSURANCE MARKETS

Alma Cohen and Peter Siegelman

Discussion Paper No. 651

11/2009

Harvard Law School
Cambridge, MA 02138

Testing for Adverse Selection in Insurance Markets

Alma Cohen[*] and Peter Siegelman[**]

Abstract

This paper reviews and evaluates the empirical literature on adverse selection in insurance markets. We focus on empirical work that seeks to test the basic coverage–risk prediction of adverse selection theory—that is, that policyholders who purchase more insurance coverage tend to be riskier. The analysis of this body of work, we argue, indicates that whether such a correlation exists varies across insurance markets and pools of insurance policies. We discuss various reasons why a coverage–risk correlation may be found in some pools of insurance policies but not in others. We also review the work on the disentangling of adverse selection and moral hazard and on learning by policyholders and insurers.

Keywords: Insurance, adverse selection, moral hazard, accidents, claims, annuities, automobile insurance, health insurance, life insurance, risk-aversion, risk, coverage.

JEL classification: D82, G22.

[*] Senior Lecturer, Tel-Aviv University Eitan Berglass School of Economics; William K. Jacobs Visiting Professor of Law and Economics, Harvard Law School; Faculty Research Fellow, National Bureau of Economic Research.

[**] Roger Sherman Professor, University of Connecticut School of Law. We would like to thank Georges Dionne and two anonymous referees for their valuable comments. For financial support, Cohen is grateful to the Israeli Science Foundation.

TABLE OF CONTENTS

I. INTRODUCTION 1

II. PREDICTING A COVERAGE–RISK CORRELATION 5

A. PREDICTION 5
B. TESTING 6
C. ISSUES 7
1. The Need for Full Access to Insurer's Information 7
2. Claims vs. Accidents 8
3. Additional Unobservable Differences among Policyholders 8

III. EVIDENCE OF THE COVERAGE–RISK CORRELATION 9

A. AUTOMOBILE INSURANCE 10
B. ANNUITIES, LIFE INSURANCE, AND REVERSE MORTGAGES 12
1. Annuities 12
2. Life Insurance 14
C. LONG TERM CARE 15
D. CROP INSURANCE 16
E. HEALTH INSURANCE 17

IV. FACTORS EXPLAINING THE ABSENCE OF COVERAGE–RISK CORRELATION 18

A. POLICYHOLDERS' LACK OF INFORMATIONAL ADVANTAGE, OR THEIR INABILITY TO USE IT 18
1. Absence of (Useful) Private Information 18
2. Private Information in the Possession of Some but Not All Policyholders 20
3. Failure by Policyholders to Use Private Information They Have 21
4. Superior Information or Predictive Ability by Insurers? 22
B. OFFSETTING FACTORS: INTERACTION BETWEEN RISK AND CHARACTERISTICS ASSOCIATED WITH INCREASED PURCHASE OF INSURANCE 23
1. Interaction between Risk and Risk Aversion: Propitious Selection 23
2. Interaction between Risk and Other Characteristics Associated with the Tendency to Purchase Insurance 25
C. INSTITUTIONAL AND REGULATORY FACTORS 26

V. DISENTANGLING ADVERSE SELECTION AND MORAL HAZARD 28

A. THE PROBLEM 28
B. RANDOMIZED EXPERIMENTS AND NATURAL EXPERIMENTS 29
C. DYNAMIC PROPERTIES 30
D. INTERACTION OF THE COVERAGE–RISK CORRELATION WITH POLICYHOLDER CHARACTERISTICS AND BEHAVIOR 31

VI. LEARNING OVER TIME 32

VII. CONCLUSION 35

I. INTRODUCTION

Adverse selection exists in an insurance market when buyers of insurance have information about their risk that the insurers who underwrite their policies lack and use this information in making their insurance purchases. The policyholder may be better informed about either the probability of a loss, the distribution of the size of the loss in the event that a loss occurs, or both. This paper offers a survey and an evaluation of the vast empirical literature on adverse selection in insurance markets.

Although substantial work has been done on adverse selection outside insurance markets, we focus on the insurance context for several reasons. First, the term "adverse selection" itself originated in the context of insurance and the insurance market has been the locale for some of the earliest economic theorizing about it (Arrow (1963), Pauly (1974), Rothschild and Stiglitz (1976)). In presenting the Nobel Prize to Joseph Stiglitz, Professor Jörgen W. Weibull cited the concept of adverse selection by noting that "[a] prime example can be found in insurance, where companies usually offer alternative contracts, where higher deductibles may be traded off against lower premiums. In this way, their clients are, by their own choice of contract, effectively divided into distinct risk classes."[1]

Second, insurance markets offer a relatively good setting for the empirical testing of adverse selection theory, thanks to the quality of the data available to researchers who study such markets. In other contexts in which adverse selection may take place, the quality of one party to a contract, and the information available about this quality to her counterparty, is often "soft" and unverifiable. Consider, for example, a labor market in which the ability of an employee may be better known to the employee herself than to her employer. In this case, even a researcher who has full access to the employer's written records may be unable to observe the quality of the employee's work because this work may be combined with that of other employees to produce output. The researcher may also be unable to observe what the employer knows about the employee's quality, as this information is often not fully reported in the employer's written records. In contrast, a researcher with full access to an insurer's records has all the information that the insurer has about the customer's risk when selling the insurance policy, as well as the ex post realization of policyholders' risks.[2]

[1] See http://nobelprize.org/nobel_prizes/economics/laureates/2001/presentation-speech.html.

[2] Of course, the insurer in turn does not have all the information that is available to policyholders, such as driving speed and miles driven in the case of auto insurance. Such information as the insurer does have, however, often exists in a form that can be, and in some cases has been, made available to researchers.

Third, as we show below, the literature on adverse selection in insurance markets is quite large,[3] and the existence and magnitude of adverse selection in insurance markets is of practical and policy significance. All this makes adverse selection in insurance markets a worthy topic for a survey in and of itself.

The basic prediction of adverse selection theory concerns the correlation between insurance coverage and risk. Under this prediction, policyholders who are known to themselves (but not to their insurer) to be high-risk will tend to choose higher insurance coverage (lower deductibles); thus, coverage and risk are expected to be positively correlated. This coverage–risk correlation has been the major focus of empirical work in this area.

Below we review a large number of studies that test for the presence of a coverage–risk correlation. We group the studies by the insurance market on which they focus: (i) automobile insurance, (ii) annuities, life insurance, and reverse mortgages (all products that center on mortality risk), (iii) long-term care, (iv) crop insurance, and (v) health insurance. A coverage–risk correlation has been found for some markets but not for others, and for some pools of insurance policies in a given market but not for others.

Facing a corpus of studies of which only some find the predicted coverage–risk correlation, researchers who comment on empirical work in this area have pronounced the evidence "mixed," "inconclusive," or "ambiguous."[4] We arrive at a different assessment. We argue that one should not expect the question of whether a coverage–risk correlation exists to be answered identically in all insurance markets or even in all pools within a market. Thus, one should not regard studies that reach opposite conclusions about the existence of a coverage–risk correlation as necessarily in conflict with each other.

Rather, we argue, the existing body of empirical evidence, informed by theoretical reasoning, provides good reasons to expect the existence of adverse selection to vary across markets and, indeed, even across segments of the same market. Whether and to what extent adverse selection exists should be expected to depend on the type of insurance product involved, the buyers' characteristics, and institutional and regulatory factors. A main focus of our survey is on distinguishing and discussing factors that vary across insurance markets and policy pools in their existence and magnitude, possibly leading to the absence of coverage–risk correlation in the data. The factors that we discuss are (i) the absence of useful private information, (ii) the existence of private information for some but not all policyholders in a market,

[3] In 2005, for example, insurance carriers and related activities accounted for 2.4 percent of US GDP (see http://www.iii.org/economics/national/gdp/).

[4] See, for example, Löfgren et al. (2002)), which is partly based on materials accompanying the award of the Nobel Prize for work on asymmetric information.

(iii) policyholders' inability or failure to use the private information that they have, (iv) the presence of superior information or predictive power on the part of the insurer, (v) propitious selection resulting from interaction between risk and risk aversion or other policyholder characteristics associated with an increased tendency to purchase insurance, and (vi) institutional arrangements.

In our view, researchers who do empirical work in this area should not think of themselves as participating in an effort to resolve once and for all the question of whether adverse selection and the coverage–risk correlation exist. We are on solid ground in believing that such a correlation does exist in some markets and policy pools but not in others. Future work in this area would do best to address itself to the question—on which some progress has already been made—of how to identify the circumstances under which one may expect the coverage–risk correlation and adverse selection to arise. We hope that our survey will provide a conceptual framework for this kind of thinking about empirical work on adverse selection in insurance markets.

Although we devote substantial attention to empirical work on the coverage–risk correlation, we also review work on two additional issues. While a coverage–risk correlation may *not* arise despite the presence of adverse selection due to factors that we discuss below, it is also true that a coverage–risk correlation may arise due to moral hazard even in the absence of adverse selection. Therefore, we review the empirical work on the disentangling of moral hazard and adverse selection. In addition, while many studies focus on informational asymmetries at a particular point in time, the information about policyholders' risk types that policyholders and insurers possess may be at least partly produced by learning over time. We also discuss the empirical work that relates to such learning.

Before proceeding, we wish to stress that adverse selection has important implications for policy, and is not only of interest to economists. The theory of adverse selection has had an important effect on insurers, government regulators, and courts.[5] To illustrate, as of August 2007, there were more than 130 state and federal opinions in U.S. courts that discussed adverse selection, in all types of insurance markets, from pension guarantees[6] to long-term disability insurance.[7] Concern about adverse selection in health insurance has prompted courts to permit marketing practices—such as paying downstream firms a bonus not to carry a rival's product—

[5] For more detailed discussion of the policy significance of adverse selection arguments, see Siegelman (2004).

[6] *Borntrager v. Cent. States, S.E. & S.W. Areas Pens. Fund*, 425 F.3d 1087 (8th Cir. 2005).

[7] *Currie v. Group Insurance Commission*, 290 F.3d 1 (1st Cir. 2002).

that would otherwise constitute clear antitrust violations.[8] Similar concerns led the U.S. Equal Employment Opportunity Commission to allow employers to exclude persons with disabilities from employer-provided health insurance if the inclusion of such persons would result in "unacceptable adverse selection."

Given public officials' receptivity to arguments based on the existence of adverse selection, it is unsurprising that policy advocates have made substantial use of such arguments. Priest (1987) argued that the U.S. insurance "crisis" of the mid-1980s, in which certain kinds of liability insurance commanded sharp premium surges or were withdrawn altogether, was produced by an adverse selection death spiral that resulted from "judicial compulsion of greater and greater levels of provider third-party insurance for victims …"[9] Romano (1989) proposed a similar adverse selection story to explain the "crisis" that befell the market for Directors and Officers liability insurance at roughly the same time.

Importantly, policymakers and policy advocates have in the past relied primarily on theoretical models or predictions of adverse selection. Our thesis—that the basic prediction of stylized adverse selection models will be manifested in some circumstances but not others—indicates that future policy discussions on this subject should avoid relying on the general possibility of adverse selection. Policy analysis for a given market should try to rely on an empirical study of that market, or at least, absent such a study, should attempt to analyze, based on the lessons of the existing body of empirical work, whether the market has the characteristics that have been found to give rise to adverse selection and a coverage–risk correlation.

Earlier surveys on which we build are Dionne and Doherty (1992), Cutler and Zeckhauser (2000), Chiappori (1999, 2000), Dionne, et al. (2000), Chiappori and Salanié (2003), Dionne, Doherty and Fombaron (2000), and Hall (2006), but we stress work done in recent years. In addition, we differ from earlier surveys in our focus on the heterogeneity of insurance markets and on the identification and discussion of factors that can explain the absence of adverse selection or a coverage–risk correlation in some insurance markets or to subsets of policies within them.

The rest of this paper is organized as follows: Section II discusses the coverage–risk correlation prediction. Section III reviews empirical tests of the prediction in various insurance markets. Section IV discusses factors that may lead to the lack of such a correlation in data for a given insurance market or set of insurance policies. Section V focuses on efforts to disentangle

[8] *Ocean State Physicians Health Plan, Inc. v. Blue Cross and Blue Shield of Rhode Island*, 692 F. Supp. 52 (D.R.I, 1988).

[9] Priest (1987, at 1524). For critiques of this argument, see sources cited in Siegelman (2004, n. 29).

empirically moral hazard and adverse selection. Section VI discusses empirical work on learning by policyholders and insurers about policyholders' risk types. Section VI concludes.

II. Predicting a Coverage–Risk Correlation

A. Prediction

The theory of adverse selection in insurance markets was introduced by Rothschild and Stiglitz (1976) and has since been developed and extended in many ways. In the basic adverse selection model, agents have private information about expected claims that insurers do not have, and insurers compete with each other. In the standard setting (which corresponds to reality in most lines other than life insurance), agents must purchase insurance exclusively from one insurer among those competing in the marketplace.

Although the theoretical literature is quite rich and offers adverse selection models that differ in significant ways,[10] one prediction that the models make—what we call the coverage–risk correlation—appears to be fairly robust and arises in a wide range of circumstances (Chiappori, Jullien, Salanié, and Salanié (2006)).[11] Since insurers cannot distinguish between high-risk and low-risk agents, the two groups must be offered the same prices for insurance. Given that the two groups face the same prices, their different risks will lead them to act differently. In particular, high-risk agents can be expected to purchase more insurance. When insurers offers menus of insurance contracts (policies), the coverage–risk correlation can be expected to manifest itself in a tendency among high-risk agents to choose contracts that offer more comprehensive coverage (e.g., lower deductibles). When insurers offer a single product, the coverage–risk correlation can be expected to manifest itself in a greater tendency among high-risk agents to purchase insurance.

Having clarified how higher coverage may manifest itself, we should also explain the ways in which higher risk may manifest itself. For the purposes of the coverage–risk correlation, we refer to an agent as having higher risk if she or he generates higher expected insurance payouts due to a larger number of expected claims, a higher expected payout in the event of a

[10] For a survey and discussion of adverse selection models, see, for example, Dionne, Doherty, and Fombaron (2000).

[11] In a recent theoretical contribution, Koufopolous (2007) identifies some circumstances in which this correlation does not arise. In his model, "if some agents choose zero coverage, there can exist separating equilibria that exhibit negative or zero correlation between coverage and the accident probability." And "if there are multiple loss levels, . . . the positive relationship between coverage and the accident probability . . . may not hold true even if all equilibrium contracts offer strictly positive coverage and administrative costs are zero."

claim, or both. Agents may, of course, have private information about either of these components.

It is also important to emphasize that the prediction of a coverage–risk correlation implies that coverage will be correlated with risk, *controlling* for all relevant policyholder characteristics that are *observable* to the insurer. These observable characteristics allow the insurer to place policyholders in different risk classes. What the insurer cannot do in the presence of information asymmetry is distinguish between higher-risk and lower-risk agents who belong to the same risk class on the basis of on their observable characteristics.[12]

B. Testing

Insofar as the researcher has access to insurers' information about policyholders, a natural way to test the coverage–risk correlation is to run a regression of the following type:

$$Risk_i = \alpha + \beta \cdot Coverage_i + \gamma \cdot X_i + \varepsilon_i ,$$

where $Risk_i$ is a variable representing the ex post realization of policyholder i's risk, $Coverage_i$ is a variable representing the policyholder's *i*'s choice of coverage, and X_i is a vector of all policyholder's characteristics that are known to the insurer and potentially relevant for classifying his or her risk. Dionne, Doherty, and Fombaron (2001) note that it would be preferable to use expected coverage rather than actual coverage in the foregoing specification in order to address problems of non-linearities or misspecifications.

[12] In the standard adverse selection model, the insurer uses all relevant policyholder characteristics that it can observe in making its pricing decision. Accordingly, policyholders who are riskier, according to their observable characteristics, will have no reason to buy more coverage since they will be charged for such coverage at an appropriate rate. There is evidence, however, that insurers do not always incorporate all relevant information into their pricing decisions. Finkelstein and Poterba (2006) show that British insurers do not take policyholders' residential addresses into account in pricing, even though this information helps to explain mortality risk after controlling for other observables. They suggest that the failure to use this information is attributable to social or political pressure by the public against basing annuity prices on such "extraneous" information. In a study of farm insurance, Makki and Somwaru (2001) show that information about policyholders' yield and revenue in the preceding ten years is not fully used in pricing even though it is available to the insurer and predicts future risks. In such situations, one may predict not only (i) that coverage will be correlated with risk, controlling for observables, but also (ii) that coverage will be correlated with observables associated with risk and not used in pricing. Finkelstein and Poterba (2006) find evidence consistent with the latter prediction.

The left-hand side of the equation may be a continuous variable such as the total cost of insurance payouts or the cost of insurance payouts in the event of a claim; it can also be a dichotomous variable, such as the number of claims, or a dummy indicating whether a claim was submitted. This approach has been long and extensively used in the literature.

An alternative approach, introduced by Chiappori and Salanié (1997, 2000) and subsequently used by others, is the bivariate model. The bivariate model uses two equations that are determined either simultaneously or sequentially. In the insurance context, one equation would be for the choice of coverage and the other for risk.

$$(i)\ \ Coverage_i = f\left(X_i\right) + \varepsilon_i$$
$$(ii)\ Risk = g\left(X_i\right) + \eta_i$$

Since the two equations are determined together, the correlation between the residuals of each of the regressions can be estimated. If the correlation is statistically significantly different from zero, then the two regressions are determined simultaneously. If the correlation is found to be zero, then each of the regressions may be estimated separately. Therefore, finding a positive correlation between the two residuals is consistent with a coverage–risk correlation. It may be shown that these two parametric models are equivalent under general conditions. The major differences between them depend on the distributional assumptions made conditional on the covariates.

Chiappori and Salanié (2000) also suggest a non-parametric test that is meant to overcome the relatively restrictive functional forms used in the two approaches presented above. Their fully nonparametric procedure is based on χ^2 tests for independence. They create 2^m "cells" constructed from m exogenous dummy variables. For each cell they compute a 2 x 2 table generated by two dummies variables—coverage, which is equal to 1 if high and 0 otherwise, and risk, which is equal to 1 if the policyholder had a least one accident and 0 otherwise—and then use several methods to test for the independence of coverage from risk (conditional on being in a given cell).

C. Issues

1. The Need for Full Access to Insurer's Information

In assessing the extent to which the coverage–risk correlation may be tested adequately by a study, it is important to know whether the researcher had full access to all information that the insurer used in classifying applicants and setting prices. In the automobile insurance context, for example, early studies such as Dahlby (1983), Dahlby (1992), and Puelz and Snow (1994)

had less information than the insurer did, whereas recent works such as Cohen (2005) and Saito (2006) are based on access to all insurer data. Insofar as a study uses only some of the policyholder information that is available to the insurer, a coverage–risk correlation may be found due to characteristics that are observed by the insurer but not by the researcher and are correlated with both higher risk and the tendency to choose higher coverage.

2. Claims vs. Accidents

Another problem that researchers need to take into account is that not all accidents/losses lead to the submission of claims. In particular, choosing a higher level of a deductible may prevent the submission of a claim that would be submitted if a lower deductible were chosen. Thus, to test whether higher deductibles are correlated with lower risks, it would not make sense to count all claims submitted by high-deductible and low-deductible policyholders. Since low-deductible policyholders are able to file claims for accidents that cause too little damage to claim under high-deductible policies, a researcher who counts *all* claims reported by low-deductible policyholders would find that this group submits more claims even if both groups are identical in their risk type. One approach in such a case, used by Cohen (2005), is to look only at claims that can be submitted by both groups of policyholders, i.e., claims exceeding the high deducible level.

Furthermore, policyholders may sometimes be reluctant to submit claims for accidents in amounts that barely exceed their deductible level. They may elect not to submit such claims in order to avoid the transaction costs involved in submitting a claim and/or to avoid having a record of a claim that may lead to an increased premium in subsequent years (Hosios and Peters (1989)). Therefore, to study whether high-deductible policyholders are riskier than low-deductible policyholders, Cohen (2005) also examines the correlation between deductible choices and claims that exceed the high deductible level by a considerable margin (100%).

Dionne and Gagné (2001) develop an econometric model to account for the possibility of policyholder reluctance to submit claims exceeding the deductible. They stress that the threshold above which a policyholder will report a loss may be specific to each individual—e.g., the cost of filing a report may depend on the value of each individual's time—and should be considered a personal deductible. This deductible is observable neither to the insurer nor to the researcher.

3. Additional Unobservable Differences among Policyholders

Although we discuss these issues in detail below, we should flag a series of problems that surface once we relax some key assumptions of the basic adverse selection model. In this basic model, the only significant information about policyholders that is unobservable to insurers

concerns policyholders' risk. In reality, however, there may be additional and important differences among policyholders that are unobservable to the insurer.

First, there may be unobservable differences in policyholder characteristics, such as the level of risk aversion, that affect choices of insurance coverage. Insofar as there are characteristics that are associated both with choices to buy higher insurance coverage and with lower riskiness, they may neutralize or reverse the positive coverage–risk correlation that the pure adverse selection model predicts. We discuss this issue in detail in Section IV.

Second, there may be unobservable differences in policyholders' precaution levels, i.e., differences in "hidden actions." In particular, policyholders who have more insurance coverage have less incentive to take precautions that can reduce risk. Thus, moral hazard may itself produce the positive coverage–risk correlation that adverse selection may be expected to produce. Accordingly, while the finding of such a correlation is consistent with adverse selection, it is also consistent with, and may be fully due to, moral hazard, as Dionne et al. (forthcoming) stress. We discuss empirical attempts to disentangle adverse section and moral hazard in Section VI.

III. Evidence of the Coverage–Risk Correlation

This section reviews the empirical evidence that relates to whether a coverage–risk correlation exists. While the theoretical analysis of contracts under asymmetric information began in the 1970s, by which time the adverse consequences of hidden knowledge were already widely accepted, the empirical testing of the models began only in the mid-1980s. Since then, however, as we describe below, much work has been done on a variety of insurance markets.

We divide our review by type of insurance market. We consider, in turn, automobile insurance (Subsection A), annuities, life insurance, and reverse mortgages, all of which center on mortality risk (B), long-term care (C), crop insurance (D), and health insurance (E).

Table 1 lists all the studies we discuss in this section, grouped by type of insurance. For each study, the first column indicates whether the study found evidence of a correlation between the level of coverage purchased by the policyholder and the risk posed by the policyholder. Additional descriptive information (type of insurance studied, data used, whether the authors had access to all information available to the insurer) is also provided, along with a brief summary of the conclusions.

A. Automobile Insurance

In the automobile insurance market, three initial studies suggested the existence of a coverage–risk correlation, but their findings were challenged by subsequent research. Dahlby (1983) and Dahlby (1992), the first two studies on the subject, did not have individual data on coverage. Puelz and Snow (1994) used individual data and also found adverse selection. Dionne, Gouriéroux, and Vanasse (2001) criticized Puelz and Snow for failing to take nonlinear effects into account and reported (using different data) that the insurer's risk classification was sufficient in the sense that there was no residual adverse selection in each risk class in the insurer's portfolio, once nonlinear effects were accounted for.

In a well known study of the French auto insurance market, Chiappori and Salanié (2000) found no correlation between risk and coverage. This study focused on a relatively homogeneous group of about 6,000 "beginning drivers" with one to three years' experience. The researchers had almost all the information that the insurer used to set premiums, a complete profile of the types of insurance contracts chosen by policyholders,[13] and characteristics of the accident(s) for which the policyholder claimed coverage from the insurer, if any. They tested for adverse selection using five variants of the coverage–risk correlation test, including (i) estimating independent probit equations for (a) the type of contract purchased and (b) the probability of an accident, and testing whether the residuals are correlated; (ii) estimating the two equations above as a bivariate probit and directly testing whether the estimated correlation parameter is zero; and (iii) a variety of discrete non-parametric methods, based on testing whether, conditional on the values of the most-important explanatory variables, loss probability and type of coverage are independent of each other in contingency tables. All of these procedures yielded the same conclusion: those who are more likely to submit claims do not buy more insurance.

Richaudeau (1999) examined the choice between basic third-party coverage and comprehensive insurance (which covers third-party liability plus damage to the policyholder's own vehicle in at-fault situations). His data, culled from a survey of French drivers, include observations on the total number of accidents experienced by each policyholder (whether or not these accidents were reported to the insurer) and a great deal of information about individual policyholders and their cars, including total miles driven per year. Positive correlation was first tested by running a probit regression on the decision to purchase comprehensive vs. basic

[13] In France, auto insurance is sold in two broad categories: a mandatory contract covering only third-party liability and a broader optional contract that also covers first-party losses to the policyholder herself and her vehicle, even if she is at fault. Although optional contracts come with a variety of deductibles, Chiappori and Salanié simplified the analysis by looking only at the choice between the minimum mandatory coverage and any type of expanded coverage.

insurance. The error term in this equation measures the policyholder's riskiness after controlling for observed variables that explain the choice of insurance plans. The results were used as an explanatory variable in a second equation that explains the number of accidents in which the policyholder was involved, fitted by a negative binomial model. The test for asymmetric information is whether the residual from the insurance equation significantly explains the number of accidents, and Richaudeau finds that it does not. He also finds, however, that without controlling for the total miles driven by the policyholder (which insurers do not know), the coefficient *is* close to statistically significant. He concludes that a modified version of adverse selection may be at work: those who drive more are more likely to purchase comprehensive insurance even though they are not at the higher risk per mile driven. This is not the intrinsic risk that is typical of most adverse selection models, but neither is it insurance-induced risky behavior (moral hazard).

Cohen (2005) obtained results suggesting that Chiappori and Salanié's finding of no coverage–risk correlation may have been due to their focus on beginning drivers. Cohen studied the Israeli insurance market, focusing on all new customers of a single insurer, and enjoying full access to all insurer data about the customers. Using methodologies essentially similar to those of Chiappori and Salanié, Cohen found no correlation between coverage and accident risk for beginning drivers (those with fewer than three years' experience) but did find a sizeable and statistically significant correlation for drivers with more than three years' experience. Among policyholders who were recent customers of the insurer (but had more than three years of experience on the road), the average number of claims was 36 percent higher for those who chose the low deductible than for those who chose the standard deductible. This finding was robust to controls for the insurer's entire information set. For example, the probability of having submitted at least one claim (in an amount exceeding the larger deductible) was about 4 percent higher for low-deductible policyholders, even after controlling for all policyholder and vehicle characteristics.

Consistent with the possibility that customers with significant driving experience may have private information, Cohen (2005) found that customers with a bad record with the insurer that she studied were more likely to "flee their record" by switching to another insurer.[14] Since at that time Israeli insurance companies did not have access to verifiable information about policyholders' claims records at other companies and official accident records, the ability to switch insurers provided policyholders who had bad records with a potential information advantage. Indeed, the study found that policyholders who switched insurers were disproportionately likely to have poor claims histories and, presumably, hoped to select against

[14] Boyer and Dionne (1989) established that a policyholder's prior claims record is a good predictor of future risk.

their new insurer by pooling with other new customers who had better claims histories than theirs.[15]

Saito (2006) studied the Japanese auto insurance market in a period immediately after extensive deregulation. Using the bivariate probit technique, Saito concluded that there was only a very weak and insignificant positive correlation (for both beginning and experienced drivers) between crash risk and the purchase of own-vehicle coverage, even when controlling for all variables observed by the insurer. There was a *negative* and statistically significant relationship between crash risk and the purchase of a zero-deductible policy. Although the insurer did not use geographic data in setting prices, Saito found that drivers in high-risk prefectures were not more likely to purchase zero-deductible or own-loss policies than those in low-risk prefectures, which again suggests little information advantage for policyholders.

B. Annuities, Life Insurance, and Reverse Mortgages

We group these three insurance products together because all of them cover risks closely connected with mortality or longevity, although the last—reverse mortgages—involves more than pure mortality risk.

1. Annuities

Friedman and Warshawsky (1990) were among the first to investigate selection in the market for annuities, concluding that annuitants outlived otherwise-similar non-annuitants. Two recent studies by Finkelstein and Poterba (2002, 2004) and one by McCarthy and Mitchell (forthcoming) find evidence of an information asymmetry in favor of annuitants. Finkelstein and Poterba (2002), who focus on the UK, and McCarthy and Mitchell (forthcoming), who examine the US, the UK, and Japan, use a roughly similar methodology that compares the aggregate mortality risk of annuitants with that of non-annuitants using standardized life tables. Annuities are typically priced on the basis of age and gender only; annuitants are not asked whether they smoke and are not required to undergo a medical exam. Thus, if annuitants outlive age- and gender-comparable non-annuitants in the aggregate, this result will not be an artifact of the insurer's ability to make finer classifications than econometricians on the basis of variables that the latter do not observe.

[15] The findings of this study with respect to experienced drivers do not indicate that one should expect to find a coverage–risk correlation in the French automobile insurance market studied by Chiappori and Salanié. Whereas Israel did not have systems for information sharing among insurers, French insurers do share information about the policyholder risks. The French system is explained in detail in Dionne, Michaud, and Dahchour (2007).

The presence of adverse selection is inferred from the fact that, after controlling for age, annuitants on average do outlive the same-age population at large. For example, McCarthy and Mitchell (forthcoming) found that the death rates of male voluntary annuitants were about two-thirds as high as those of the same-age general population in the US and the UK, and about four-fifths as high in Japan. Finkelstein and Poterba (2002), using somewhat richer data, detected selection not just on the purchase/no purchase margin but also by the type of annuity selected. For example, those who purchased annuities that included a payment to the annuitant's estate in the event of early death did, in fact, tend to die sooner.

Finkelstein and Poterba (2004) had more refined data, and could therefore examine several additional aspects of contract design. Using a single insurer's complete records for 42,000 individual annuity purchasers, they estimated hazard rates for mortality, conditional on various parameters of the annuity purchased and all other covariates known to the insurer (age, gender, etc.). They observed only a slight tendency for longer-lived individuals to purchase annuities with a higher initial annual payout, contrary to the prediction that a standard selection test would yield. (This conclusion obtained most strongly for those who purchased annuities voluntarily; some retirees are required by law to annuitize their savings, and there was stronger evidence of selection on payout size among these compulsory purchasers.)

However, Finkelstein and Poterba (2004) also find evidence that longer-lived individuals are more likely to purchase "back-loaded" policies—those in which a greater share of total payouts are made in later years (assuming that the annuitant survives) due to escalating nominal payouts over time. Such policies are obviously more advantageous to those who believe they will live longer than their insurer predicts. Another aspect of contract design on which selection may occur is the extent to which payments survive the annuitant's death. Some contracts are structured so as to provide for guarantee periods, so that if the annuitant dies while the guarantee remains in force, the annuity continues to make payments to the annuitant's estate for the remainder of the guarantee. As would be expected if annuitants self-select on this margin, Finkelstein and Poterba (2004) find that longer-lived individuals are less likely to purchase annuities that come with guaranteed survivor payments, since such guarantees are most valuable to those who die earlier.

Reverse mortgages are annuity-like instruments that allow homeowners to remain in their homes and to borrow against the cash realized from the future sale of their homes (which often, but not always, happens due to their death). Although there are many home-rich but cash-poor elderly homeowners for whom a reverse mortgage should be attractive, the market has failed to develop. Davidoff and Welke (2005) consider the possibility that adverse selection has forestalled the development of this product. Such selection might occur if consumers have private information about their long life-spans, low mobility, or low appreciation rates for their

homes, and use this information to speculate against the holder of the mortgage (who benefits from rapid departure in much the same way that an annuity issuer benefits if the annuitant dies earlier). However, using calibrated numerical simulations as well as the positive–correlation test, Davidoff and Welke decisively reject the possibility of adverse selection.[16]

In the context of reverse mortgages, the positive–correlation test requires homeowners with reverse mortgages to stay in their homes longer than those without such mortgages. Davidoff and Welke, using from 77,000 reverse mortgages covered by the US government's Home Equity Conversion Mortgage (HECM) program as well as the American Housing Survey panel (for comparison purposes), find that HECM participants did not have longer life spans than the general population of the same age and gender, after controlling for observables. An alternative to the positive–correlation test asks whether reverse mortgage holders are more sensitive to increases in home equity values than non-participants. The authors find this to be the case, suggesting that *favorable* selection occurs because reverse mortgage borrowers are especially likely to have strong tastes for expenditures earlier in life (a higher discount rate).

2. Life Insurance

In contrast to the papers on selection in annuity markets, all of which find strong evidence of selection, those testing for the presence of selection in life insurance markets reach generally negative conclusions. This is puzzling since, as several of the authors note, both annuities and insurance involve the same risk—mortality—albeit with opposite consequences for the insurer. The apparent absence of selection in life insurance and its presence in annuities is difficult or impossible to explain if selection is based (only) on the policyholder's superior knowledge of his or her mortality risk. In that case, policyholders who know they will die sooner than their insurer believes should prefer to buy life insurance, and those who know they will live longer than their insurer believes should prefer to buy annuities. We return to this puzzle below.

Cawley and Philipson (1999) used both US aggregate and micro data to produce three main findings, none of which supports the predictions of a simple adverse selection story. First, after controlling for age, gender, and smoking status, they found that the death rate for persons who had life insurance was lower than for those who lacked it. Second, they found that quoted life insurance premiums tend to fall, rather than rise, with higher levels of coverage. Such quantity discounts appear to be inconsistent with a significant role for adverse selection in life insurance, because most adverse selection equilibria require rationing, and rationing is possible

[16]The authors also examine moral hazard in the reverse-mortgage market. Although theory yields ambiguous predictions, the data strongly support the conclusion that having access to additional cash does induce most homeowners to remain in their homes longer than they otherwise would, but not by enough on net to overcome the positive selection effects discussed earlier.

only if two $100,000 policies are more costly than one $200,000 policy. In fact, Cawley and Philipson showed that premiums per dollar of coverage fell with increased amounts purchased. The study also compared people's self-assessed risk of death (based on interview data) with their insurer's predictions, and their own insurance purchases. They found that the relatively risky are *less*, rather than more, likely to hold life insurance. They then estimated models to predict whether an individual would die during a given period, using age, gender, smoking status, and the person's self-assessed likelihood of dying. After controlling for the level of a policyholder's premium as calculated from life insurance tables—which reflects the *insurer's* assessment of the individual's risk—they found no *additional* gain from knowing the policyholder's self-assessed risk of death.

McCarthy and Mitchell (forthcoming) compare mortality rates for policyholders and others in the US, the UK, and Japan, for several ages and for different types of insurance coverage, on the basis of aggregate mortality tables. They find that purchasers of life insurance (except in Japan) have substantially *lower* mortality risk than the population as a whole—a negative, rather than a positive, correlation between risk and coverage. They conclude that "Underwriters are relatively effective at screening out poor risks [and . . .] insurance companies can better assess mortality risks than can individuals themselves."

C. Long Term Care

Long-term care insurance combines aspects of annuity and health insurance. Like the former, it provides coverage against the risk that one will outlive one's assets by paying for nursing home care, which is often a large and burdensome expense for the elderly. (The average rate for a semi-private room in a nursing home in the US was more than $50,000 per year in 2002.) Like the latter, it covers some medical expenses and also insures against long-term increases in medical costs, as ordinary health insurance does not. In the US, long term care outlays verged on $150 billion in 2004, representing a significant financial risk for the elderly. However, only about 10 percent of the elderly have a private long term care insurance plan (Brown and Finkelstein, 2007). If adverse selection has prevented the development of this market, as some have suggested, it may constitute a potentially serious welfare loss. The consensus, however, is that this seems not to be the case.

Finkelstein and McGarry (2006) study the US market for long-term care insurance on the basis of data from the Asset & Health Dynamics (AHEAD) survey of the elderly. The survey contains questions about preventive-care actions by individuals, and the authors assume that individuals who take more such actions are more risk-averse. Applying the positive–correlation test to the population as a whole, Finkelstein and McGarry find that those who purchase insurance are not at higher risk than the general population. However, when comparing

individuals' subjective assessments of the likelihood that they will enter a nursing home with the insurance companies' assessments, they find that individuals *do* have residual private information. They explain this riddle by suggesting that two types of people buy this form of insurance, each with a different kind of private information: individuals who have private information about their risk type and individuals who have private information about their preferences. Individuals with private information indicating that they are at higher risk would show a positive correlation between risk and coverage. Those who have strong preferences for this insurance (Type 2s), however, tend to be less risky due to their psychologically cautious temperaments. In the aggregate, the behaviors of these two types may offset each other, causing the positive correlation to disappear despite the existence of asymmetric information about risk type.[17]

D. Crop Insurance

Crop insurance is unique among the insurance markets considered in this survey, since farmers presumably maximize some combination of utility and farm profits. Thus, Just, Calvin, and Quiggin (1999) decompose the motives for purchasing crop insurance into three components: a risk-aversion effect, a subsidy effect (in 1988, when the data were collected, the US government underwrote part of the cost of the crop insurance program), and an adverse selection effect that occurs if farmers can use private knowledge to speculate against their insurer. The study concludes that the risk-aversion effect is small and that farmers purchase crop insurance primarily to receive the subsidy or because of adverse selection possibilities.

Makki and Somwaru (2001) use the positive–correlation test to examine adverse selection in an environment in which farmers are offered a variety of yield- and revenue-insurance products. They use Generalized Polychotomous Logit techniques to explain farmers' choices among four or five different insurance alternatives, some of which cover only yield risk while others cover revenue risk (caused by decreases in *either* yield or prices). Risk is measured retrospectively for each farm, using ten-year retrospective data on yields and prices to compute the probability that yield or revenue will fall below the amount guaranteed in the insurance contract. Insurers could presumably use the same data to compute premiums. Nevertheless, the authors find strong evidence that high-risk farmers prefer revenue insurance to yield insurance and individual insurance relative to group (county-based) insurance products, perhaps because farmers can predict their coming year's efforts better than the insurer can. Moreover, for most of

[17] A somewhat different interpretation of the study's findings is that individuals who undertake precautions tend to be more cognitively able, and cognitively-able individuals are more likely to purchase long-term care insurance. This explanation would also lead to a negative correlation between risk and tendency to purchase insurance.

the insurance products studied, high-risk farmers purchased more generous coverage than low-risk farmers did.

Sil (2005) considers the possibility of "endogenous" adverse selection in the market for crop insurance. Farmers can purchase insurance against declines in yield per acre and also have the option of selling some of their crop forward at a fixed price. Sil shows, theoretically, that having a forward contract leads a farmer to exert less effort in growing crops, regardless of whether or not insurance is also purchased. Since insurers do not observe whether policyholders also have a forward contract, heterogeneity among farmers creates a kind of endogenous adverse selection. Empirically, farmers who have forward contracts are more likely to choose more generous crop insurance than those who lack such contracts and also to experience larger insured losses (by about 6 percent), in keeping with the positive–correlation prediction.

E. Health Insurance

There is a significant body of empirical work that finds evidence of adverse selection within health insurance markets, i.e., people in poorer health choose plans that offer more generous coverage. The empirical work on health insurance markets is discussed in detail in the review by Cutler and Zeckhauser (2000). That review notes fourteen studies that examine the selection of more expensive options within a given type of insurance plan, all of which find some type of adverse selection.[18] It also reports sixteen additional studies that consider other margins (e.g., whether a policyholder re-enrolls or decides to forego insurance), virtually all of which find information asymmetry in favor of policyholders.

In one well known study, Cutler and Zeckhauser (1998) offer a compelling analysis of intra-market selection, examining data from the Massachusetts Group Insurance Commission and comparing employees who chose a more-generous fee-for-service plan with those who chose a less-generous HMO (health maintenance organization) arrangement. Those in the first group spent significantly more and were more likely to experience significant medical events (giving birth, having a heart attack) than were the HMO enrollees. In another widely noted work, Cutler and Reber (1998) study health insurance provided by one large employer, Harvard University, via several different plans. Harvard moved from subsidizing only the most generous plans to a fixed-dollar subsidy (regardless of the generosity of the plan chosen), increasing the annual cost for the most generous plan by roughly $500. The positive–correlation hypothesis was dramatically borne out: the most generous plan was abandoned by the best risks. For example, those leaving the generous plan for the HMO option were four to five years younger on average

[18] Notably, however, there is also some empirical work on health insurance that finds no evidence of a positive correlation between risk and coverage. (See Buchmueller et al. (2004), Ettner (1997), and Browne and Doerphinghaus (1993).)

than those who remained. Those who quit also had lower medical expenses than those who stayed.[19]

IV. Factors Explaining the Absence of Coverage–Risk Correlation

A risk-coverage correlation appears to be a feature of some insurance markets or pools of insurance policies but not of others. Therefore, a good adverse selection theory must explain when a positive risk-coverage correlation will be present and when it will not. In this section, we discuss explanations for the diversity of results produced by the empirical evidence reviewed in the preceding section. We consider several explanations for why a risk-coverage correlation may not be observed, as well as the empirical support for these reasons.

Section A focuses on reasons why policyholders may not have an information advantage, or may fail to use whatever advantage they have. Section B discusses factors that by themselves would lead to a negative correlation between risk and coverage and, thus, may have an offsetting effect on the positive risk-coverage correlation that adverse selection might otherwise produce. Finally, Section C discusses institutional and regulatory factors.

A. Policyholders' Lack of Informational Advantage, or Their Inability to Use It

1. Absence of (Useful) Private Information

Theoretical models of adverse selection typically represent information asymmetry in a highly stylized fashion. Policyholders are usually assumed to have superior information about their own probability and/or size of loss. This, however, may be unrealistic in some contexts, such as the automobile insurance market. Chiappori and Salanié (2000) suggest that the lack of such superior information may explain their failure to find a risk-coverage correlation in their study of this market.

To fix ideas, we can define individual *i*'s expected loss as $E(L_i) = p(X_i) \cdot L(Y_i)$, where *X* and *Y* are vectors of all of the explanatory variables (primitives) that define the probability and size of the loss, including variables whose values are unknown to both the insurer and the

[19] The generalizability of these findings is somewhat limited, however, by the fact that the selection was largely based on observable variables such as age. Even though age was observable, Cutler and Reber point out that the contractual relationship between insurer and employer forbade the former from taking age into account when setting premiums. This, by itself, "explains a large part of adverse selection." Such a constraint, however, would not apply in many other contexts outside of employer-provided health care. Moreover, the ability to select among health insurance plans offered by the same employer is probably substantially larger than the ability to select *across* insurers in most other contexts.

policyholder. For example, if the loss involves an automobile accident, X may include driver i's temperament and aggressiveness, total miles driven, average speed, road congestion, and so on, and Y may include the make and model of the policyholder's car.

There are several ways in which a policyholder may possess imperfect information. First, she or he may not know some of the elements in the X or Y vectors, or may know them only with significant random errors and/or biases. Second, the policyholder may lack information about the functional forms that translate information about behavior or environment into estimates of the probability or size of loss. Finally, a factor not known to anyone—luck—may explain much of the variance in p or L across policyholders; thus, even perfect knowledge of some elements of X and Y and the functional form will not give policyholders a significant advantage over insurers in predicting expected loss.[20]

It seems plausible that policyholders may be ignorant of or mistaken about the underlying "primitives" that determine the probability or size of loss and how those compare with the value of these parameters for other policyholders. Consider automobile insurance. One might think that a policyholder could observe his or her own driving "style." However, for the policyholder to have useful private information, she must also know how his or her driving compares with that of other drivers in the insured pool; it is not clear that drivers generally have this information. For example, Svenson (1981) found that 80 percent of all drivers believed that they were in the top 20 percent of safest drivers. Guppy (1993) found that British drivers tended to guess that the probability of their having an auto accident was lower than the probability of such an event befalling an average person of their age and gender. Moreover, those with prior convictions for speeding or drunk driving generally perceived themselves as *less* likely to have an accident than members of the non-offender group did.

Even when these primitives can be assessed objectively, policyholders may not be able to translate them into a probability or dollar amount in order to accurately estimate their expected loss and compare it with their premiums in deciding how much coverage to buy. For example, it is clear that a major determinant in auto-accident risk is the total miles driven in a given year

[20] Imagine the insurer running a regression in which the dependent variable is whether or not the policyholder experiences a loss and the right hand side variables are those in the X vector that are known to the insurer. Suppose the policyholder also runs this regression, using *all* elements of X that s/he or he knows (both public and private information). If the R^2 on the second regression is not much higher than the first, the insured's extra information will not be very valuable because random factors (those not known to either party) will be responsible for much of the risk of loss.

(Butler, 1996).[21] This would seem to give policyholders an information advantage, since they know their own mileage better than their insurer does. Even so, however, the policyholder may not realize the importance of this variable as a determinant of his or her accident risk because he or she does not know the functional form of the relationship between miles driven and the probability of an accident. (Evidence suggests that drivers overemphasize their own skill and underemphasize miles driven as contributors to accident risk.)

More directly, we can observe people's proficiency in predicting outcomes in their own lives. The evidence is limited and somewhat mixed. For example, several direct studies of mortality risk suggest that people can do a reasonably good job of predicting how long they will live (Hamermesh (1985), Hurd, Smith, and Zissimopoulos (2002)), while other recent work suggests that people do not forecast their own demise very accurately (Cawley and Philpson (1999), Bhattacharya, Goldman, and Sood (2003)). Extreme cases in which policyholders have accurate private knowledge of their impending mortality do seem to lead to selective insurance purchases,[22] but there is no reason to think these results translate readily to less extreme situations.

The foregoing suggests that the key assumption of adverse selection—that policyholders have better information about their risk type than the insurer does—is not necessarily valid for all insurance markets and products. Future research should seek to identify the products and groups of policyholders where insurance customers can be expected to have useful private information.

2. Private Information in the Possession of Some but Not All Policyholders

Whether policyholders have private information that can yield a risk-coverage correlation depends, we have suggested, on the type of insurance product at issue. Such private information may exist with respect to some types of insurance products but not for others. For a given

[21] Insurers do not typically ask for (or verify) detailed information on total miles driven (although it would arguably not be difficult to verify, and its use in setting premiums could lead to significant efficiency gains (Vickrey (1968), Edlin (2003)).

[22] The test for the gene that causes Huntington's disease is extremely accurate and the disease is fatal and untreatable. Someone who tests positive for the genetic defect may well be tempted to buy substantial life-insurance coverage if he or s/he can do so without paying the much higher premium that h/is or her test outcome would warrant. Chiappori (2006) and Hoy and Witt (2007) analyze the selective effects of genetic testing for the BRCA1/2 genes that are implicated in breast cancer. Alternatively, consider *USLife Credit Life Ins. Co.* v. *McAfee* (29 Wn. App. 574 (1981)), in which the policyholder—knowing that his wife had terminal cancer—obtained credit life insurance on seventeen loans taken out in her name, exploiting the fact that the credit life applications did not require any health declaration.

insurance product, there may be a subset of policyholders who have private information and another subset who do not. In such a case, a risk–coverage correlation may be found if the researcher focuses on the first group alone. However, we would not expect to find a risk–coverage correlation in the second set of policies and, if this group is sufficiently large, such a correlation may not be uncovered by a researcher who observes the set of all policies rather than the first set of policyholders with private information.

That the existence of risk-coverage correlation may vary among subsets of policy pools within one insurance market has been shown by Cohen (2005). Examining the set of policies sold by a single insurer, Cohen finds that a risk–coverage correlation exists among those policies sold to drivers who have three or more years of driving experience, but not for those sold to less experienced drivers. These results are consistent with the hypothesis that private information sufficient to yield a risk-coverage correlation exists only when a new customer has enough driving experience before joining the insurer. Interestingly, they also suggest that customers cannot glean significant private information about automobile-accident risks from introspection or by observing their performance in other dimensions of life. Rather, it appears that only direct experience with one's own driving can provide significant private information about one's risk type, at least in the context of automobile accidents.

Given the foregoing evidence, future research would do well not to limit itself to testing for a risk–coverage correlation within the entire set of policies available to the researcher but rather to conduct separate tests for subsets of policies. The absence of a risk–coverage correlation in a universe of policies does not rule out the possibility of such a correlation in identifiable subsets of this universe. The results of Cohen (2005) point to a partitioning into subsets on the basis of policyholders' experience, and future work may do well to explore other methods of partitioning as well.

3. Failure by Policyholders to Use Private Information They Have

For a risk-coverage correlation to come about, it does not suffice for policyholders to have private information about their risk type; they must also adjust their purchasing decisions on the basis of this information. Whereas policyholders who perfectly optimize their decisions would make such adjustments, behavioral economics suggests that individuals often fail to engage in perfect optimization, especially when the stakes are not high—as is the case, for example, with the choice of deductible levels in automobile insurance.

Thus, a risk-coverage correlation may not come into being if policyholders routinely fail to act on private information that they have by altering their insurance purchases (at least for risks within the normal range). Hurd et al. (2002) conclude that elderly respondents *could* predict actual mortality fairly well, for example, but also found that despite this (except for the most

extreme cases), these predictions did not systematically translate into actual (retirement) behavior in the way that economic theory would predict. That is, even when policyholders can make more accurate predictions than their insurers, they may not use these predictions in their insurance-purchase decisions.

As Pauly et al. (2003) note, "[r]eal consumers . . . have more on their minds than paying attention to small bargains in insurance markets." In fact, Pauly et al. conclude that demand for life insurance is insensitive to changes in either price or risk, especially the latter. Their estimate of coverage elasticity with respect to mortality risk was only about 0.2 to 0.5. Therefore, even when individuals can predict their own risk better than their insurers can, they may simply fail to act on their advantage by increasing the amount of insurance they buy.

4. *Superior Information or Predictive Ability by Insurers?*

Another possible reason for the absence of a positive correlation between risk and quantity of insurance purchased is that insurers may actually possess better information about a policyholder's risk than the policyholder does. (Chiappori and Salanié (2000)). Insurers' risk classification or underwriting is done using a combination of subjective and objective/actuarial techniques. Even if the policyholder knows things about him or herself that the insurer does not, the insurer may nevertheless be in a better position to forecast risk (using the variables it *does* know) than the policyholder, if it uses a superior forecasting method.

Although not directly concerned with insurance, Grove and Meehl's (1996) survey of "clinical" vs. "statistical" prediction lends support to this view. Noting that "[h]umans simply cannot assign optimal weights to variables, and they are not consistent in applying their own weights," the authors surveyed 136 studies that compared experts' predictive judgments with those made by a simple actuarial or statistical model. In more than half of the cases, in a wide variety of settings—criminal recidivism, college grade point averages, firm bankruptcies—the simple mechanical model made better predictions than the experts' subjective ("clinical") judgments. In the remaining cases, the two methods performed equally well, and the experts almost never outperformed the actuarial prediction.

Consider information about an individual's credit score (a single number that summarizes his or her credit history). Whether or not there is a bio-psychological basis for the correlation between credit scores and risk (Brocket and Golden, 2007), the relationship seems remarkably robust. An insurer who knows a policyholders' credit score (and the robust relationship between it and, for example, accident risk) may actually be able to predict a policyholder's riskiness

better than she or he could, even if the policyholder has superior information about some aspects of his or her own behavior.[23]

B. *Offsetting Factors: Interaction between Risk and Characteristics Associated with Increased Purchase of Insurance*

Adverse selection theory focuses on individuals' incentive to purchase more insurance when they know that they face higher risks. Thus far, we have focused on reasons why policyholders may not have, or may fail to use, superior information vis-à-vis the insurer. We now ask why the data may not elicit a risk–coverage correlation *even* when policyholders have private information and are able to use it. In particular, we discuss the possibility of factors that would by themselves yield a negative correlation between risk and coverage and, thus, may "offset" the positive risk–coverage correlation that adverse selection would produce if it were on its own. We begin by discussing the possibility that risk is correlated with risk aversion, which, other things being equal, leads to increased purchase of insurance; then we discuss the possibility that risk is correlated with other variables associated with increased insurance demand.

1. Interaction between Risk and Risk Aversion: Propitious Selection

One reason why high risk may be correlated with increased insurance demand is that it may be correlated with low risk aversion. Hemenway (1990) was apparently the first to propose the term "propitious selection" to describe an observed *negative* relationship between insurance demand and riskiness. Based largely on anecdotal evidence, he concluded that high-risk individuals are less likely to purchase insurance because they are also less risk-averse. This mechanism leads to selection that is advantageous to insurers, since the insured population is less risky than the population as a whole.

DeMeza and Webb (2001) provide an elegant theoretical model based on the same idea: cautious people put more effort into preventing accidents and are also more likely to buy insurance. Using a model that combines moral hazard and selection, De Donder and Hindricks (2006) suggest that propitious selection cannot account for a negative correlation between insurance purchase and riskiness, at least when the costs of precaution are exclusively monetary. To obtain the negative correlation, they say, one must find not only that the more risk-averse take more precautions than the less risk-averse, but also that the less risk-averse exhibit decreasing

[23] We assume throughout this survey that the positive correlation between risk and coverage is conditional on the menu of coverage offered by the insurer, which is often designed to induce policyholders to reveal h/is or her type. The question is whether there is additional private information that can not be revealed by the choice of coverage.

willingness to pay for insurance, which they argue is impossible in equilibrium.[24] Karagyozova and Siegelman (2007) use simulation techniques to conclude that with a continuum of types, even a very large negative correlation between risk aversion and riskiness cannot prevent the unraveling of the market in the presence of asymmetric information (for reasonable values of risk-aversion).

Although the ability of propitious selection models to explain the absence of adverse selection remains an open question, there is some empirical support for a positive correlation between financial and non-financial risk aversion across individuals, which is what drives these models. Guiso and Paiella (2003, 2005) survey individuals and construct a direct measure of financial risk aversion from the amount that respondents say they would be willing to pay to enter a hypothetical lottery. They find that this measure is correlated with choice of job, remaining in one's region of birth, and suffering from a chronic disease. "Overall," they conclude, "the evidence . . . implies that attitudes towards [financial] risk have considerable explanatory power for several important . . . [nonfinancial] decisions." Barsky et al. (1997) use a broadly similar survey methodology and find a similar pattern, although with weaker explanatory power for financial risk aversion. Dohmen et al. (2005, 2007) also find a relationship between financial and non-financial risk aversion on the basis of survey data. Loewenstein et al. (2001, p. 275) suggest that risk aversion is largely affective rather than cognitive and tends to be inconsistent across different contexts.

Finkelstein and McGarry (2006) propose that the absence of a positive correlation between riskiness and demand for long-term care insurance stems from multiple sources of heterogeneity among policyholders. That is, some have private information about riskiness while others have private information about their own risk aversion. If the first type are high-risk, they will use this information to increase the amount of insurance they purchase; if, however, the second type are predominantly low-risk (as empirically appears to be the case in the Finkelstein and McGarry data), they will be drawn into the insurance pool by their high risk aversion and may offset the negative effect of the high-risk policyholders' participation.

Cutler, Finkelstein, and McGarry (2008) seek to explain the "puzzling" absence of selection by investigating five different insurance markets: life, health (private acute care),

[24] De Donder and Hindricks (2006) model the relationship between riskiness and investment in precautions as a moral hazard, such that insurance prompts the more risk-averse to purchase more insurance and take fewer precautions as a result. An alternative reading of Hemenway's argument is that financial risk aversion is correlated with "non-financial risk aversion," a psychological tendency to worry about physical or emotional risks and to take steps to avoid them. There is no reason that an enthusiastic skydiver or free climber could not simultaneously be risk-averse in the economic sense, although it seems implausible that physical daring would typically accompany the avoidance of financial risk.

annuities, long-term care, and Medigap. Their results are consistent with the explanation that individuals may vary in their tolerance for risk in addition to their riskiness. Gathering data from different sources on insurance coverage, several measures of occurrence of risk, and risk tolerance,[25] they test how these three factors are related. They run unconditional and conditional bivariate relationships between two regressions: one that investigates the effect of different levels of risk tolerance on whether individuals purchase a specific type of insurance and one that examines the effect of different levels of risk tolerance on their measures of the occurrence of risk. Their analysis yields two main findings: first, individuals who engage in "risky behavior" or who do not take precautions to reduce risk are less likely to purchase any of the kinds of insurance studied. Second, these individuals tend to have higher expected claims for life insurance and long-term care insurance but have lower expected claims for annuities. For Medigap and acute care health insurance, no systematic relationship was found.

2. Interaction between Risk and Other Characteristics Associated with the Tendency to Purchase Insurance

In addition to a possible correlation between low risk and high risk-aversion, low risk may also be correlated with other variables associated with a greater tendency to purchase insurance (Fang, Keane, and Silverman (2008)). Like the correlation between low risk and high risk-aversion, any such additional correlation may also contribute to advantageous selection, which may in turn counter the effects of adverse selection and lead to the absence of an observable risk–coverage correlation.

Consider, for example, "cognitive ability." Fang et al. (2008) show that US senior citizens who purchase Medigap insurance (a supplement to Medicare) spend roughly $4,000 per year *less* on health care than those who do not purchase such policies. On the other hand, conditional on health, those covered by Medigap spend roughly $2,000 more than those not covered. Taken together, these results are starkly at odds with the positive–correlation test. The authors' preferred explanation for the negative relationship between coverage and risk is that higher cognitive ability (as measured by several survey questions) is associated with both increased demand for insurance (because the more sophisticated are better able to understand the

[25] For each type of insurance, they use a different measure of risk. For life insurance, the indicator is whether the individual died between 1992 and 2002. For acute private health insurance, they ask whether the individual reported any use of hospital services in the previous two years. For annuities, they look at whether the individual survived from 1995 to 2002. For Medigap, they study expenses reported in 1995, and for long-term care they ask whether the individual entered a nursing home. To assess risk tolerance, they use five behaviors: smoking, drinking, occupational mortality risk, receipt of preventive health care, and use of seat belts.

need for insurance or better able to understand the complex rules governing the Medicare program) and better health (because the more sophisticated take better care of themselves).

Similarly, an important point raised in many annuity studies and several long-term care insurance studies is the possibility of selection on grounds other than mortality risk. The apparent absence of selection in life insurance and its presence in annuities seems hard to explain if selection is based (only) on the policyholder's superior knowledge of his or her mortality risk; in this case, policyholders who know they will die sooner than their insurer believes should prefer to buy life insurance and those who know they will live longer than their insurer believes should prefer annuities.

Several papers find evidence of selection in insurance purchases that appears to be based on wealth or income, education, or socio-economic status (e.g., Finkelstein and Poterba, 2006). A positive relationship between income and the decision to purchase insurance may explain the different selection results between annuities and life insurance. Higher income is negatively associated with mortality risk and positively associated with insurance purchases of all kinds. This means that demand for life insurance is driven by conflicting factors—higher income on the one hand, higher mortality risk on the other—that may partly cancel each other out. (This conclusion is strengthened by Pauly et al. (2003)'s finding of low risk-elasticity of insurance demand.) The net result would be that life-insurance purchasers strongly resemble the population at large. In contrast, higher incomes and *lower* mortality risks drive selection toward annuity purchases, and these two factors work in the same direction, with the result being that annuitants outlive the population at large.

C. Institutional and Regulatory Factors

Policyholders' decisions about whether to buy insurance (and if so, how much) may also depend on "institutions," an admittedly imprecise term (Finkelstein and Poterba (2006)). For example, in many contexts, insurance is not sold directly to consumers but is intermediated by "producers," brokers or agents whose job it is to match customers with insurance providers and with particular policies and coverage. Cummins and Doherty (2006) argue that one justification for having such intermediaries—and, in particular, for paying them in a way that hinges on the profitability of the business that they place with an insurer—is that they can mitigate information asymmetries and adverse selection. (For a contrary view, see Schwarcz (2007).) If Cummins and Doherty are correct, then cross-industry or cross-national differences in the structure of insurance intermediaries may account for some of the previously described heterogeneity in adverse selection. Since few if any of the studies surveyed above provide significant details on the presence or absence of institutions of this kind, it is difficult to know whether such institutions actually play a significant role in explaining observed differences in selection.

The rubric of "institutions" should also include the way that insurance is advertised or marketed to the public, which is often tightly controlled by national or subnational regulation. For example, many countries allow insurers to market health and life insurance by offering lottery tickets as an enticement (Baker and Siegelman, forthcoming). Such marketing techniques appear especially effective in recruiting certain groups who tend to under–purchase insurance—especially the young—and hence may offset selective pressures that might otherwise materialize. The cliché that "life insurance is sold and not bought" suggests that regulations governing permissible sales practices may have a significant effect on selection pressures: Since insurers are reluctant to sell to customers who are eager to buy (presumptively the worst risks), the extent of sales to good risks is likely to be constrained by restrictions on marketing techniques.

The presence or availability of alternative forms of insurance may also play a role in selection for a given type of insurance. Universal health insurance, for example, may limit selection in demand for annuities by eliminating a major source of expense for older people while leaving demand for life insurance relatively unchanged.[26]

It is also important to realize that risk classification by insurers is heavily regulated in almost all jurisdictions. Life insurers in the US are forbidden both by civil-rights laws and insurance regulations from using race in setting premiums, even though race is predictive of longevity and had been used in the past (Paltrow (2001)). Health insurers may not use "pre-existing conditions," auto insurers are barred from relying on "credit scores" in some jurisdictions, and so on. Under such circumstances, individuals who know that they are high-risk will have an added incentive to purchase high coverage. This kind of selection is not brought about by asymmetric information. Instead, it is attributable to the insurer's not using its full information in a way that leads to the cross-subsidization of some insurance policies at the expense of others. This is the mechanism that Finkelstein and Poterba (2006) describe in the UK annuities market, where insurers fail to use policyholders' residential addresses in setting premiums even though address does predict mortality risk.

[26] Or, as De Donder and Hindricks (2006) point out, social insurance available at a pooled rate may exacerbate selection in supplemental private-insurance markets through differential effects of the "free" social insurance on the moral hazard of high- and low-risk insureds. A similar argument is made by Sil (2005), who found that futures markets, which serve as insurance for growers who participate in them, exacerbate adverse selection in the market for crop insurance.

V. Disentangling Adverse Selection and Moral Hazard

A. The Problem

In the preceding section we considered why the data may not reveal a coverage–risk correlation despite the presence of adverse selection. We now discuss the possibility that adverse selection may not exist despite the presence of a coverage–risk correlation. In particular, such a correlation may arise even in the absence of adverse selection due to moral hazard.

Unlike adverse selection, which has to do with "hidden information," moral hazard has to do with "hidden action." The risk of an accident and the losses it produces may be a product not only of the policyholder's (unalterable) "type" but also of his or her behavior. A policyholder may invest in precautions that reduce the probability or the size of the loss. The purchase of insurance diminishes the policyholder's incentives to invest in such precautions because the insured policyholder no longer captures their full benefits, at least part of which accrue to the insurer instead.

The presence of moral hazard provides an important reason for the use of deductibles in insurance contracts. (See Harris and Raviv (1978), Hölmstrom (1979), Shavell (1979), and survey by Winter (2000); Baker (1996) provides an intellectual history of moral hazard in the insurance industry.)[27] If an insurance policy covers the full loss, the policyholder has no incentive to take precautions; a deductible restricts the payout to only part of the loss, thereby giving the policyholder an incentive to take precautions. The higher the share of loss covered (the lower the deductible), the lower the level of precaution and, in turn, the higher the expected value of a policyholder's loss.[28]

Thus, the presence of moral hazard can be expected to produce a correlation between coverage and losses, with all known observables controlled for. This coverage–risk correlation is, of course, predicted by adverse selection as well. In the adverse selection story, the correlation originates from the choice of lower deductibles by riskier insureds. In the moral-hazard story, the correlation results from lower deductibles that lead to lower levels of caution and, in turn, higher risks.

---

[27] When more than one loss may occur during the life of a policy, moral-hazard considerations may call for having a deductible for each loss and not just a deductible applying to the aggregate losses during the life of the policy. See Cohen (2006).

[28] Research on moral hazard has long distinguished between ex ante moral hazard, which refers to the tendency of policyholders to take reduced precautions, and ex post moral hazard, which refers to policyholders' actions after a loss occurs. (See, for example, Dionne and St-Michel (1991).) The discussion in this section focuses largely on ex ante moral hazard.

The finding (in cross-sectional data, and conditional on observables) that coverage is correlated with risk is consistent with either adverse selection or moral hazard. Thus, by itself, the observation of such a correlation does not suffice to tell us whether it is caused by adverse selection alone, moral hazard alone, or both. The disentanglement of adverse selection and moral hazard is probably the most significant and difficult challenge that empirical work on adverse selection in insurance markets faces. Below we discuss three approaches to this challenge, based respectively on randomized and natural experiments, the dynamics of insurance contracts, and the interaction of the coverage–risk correlation and policyholder characteristics.

We abstract below from the significant literature testing for the presence of moral hazard in contexts that do not involve individual choice among insurance contracts. This literature identifies a moral hazard effect in some contexts but not in others. Kaestner and Carroll (1997) and Fortin and Lanoie (2000), for example, find that more generous coverage of workers' compensation is associated with more workplace injuries. Cohen and Dehejia (2004) find that increases in the incidence of automobile insurance and changeovers to no-fault liability systems have significant negative effects on traffic fatalities. Klick and Stratmann (2007) find that diabetics in the US exhibit higher BMIs (body/mass indices) after their states pass laws mandating health insurance coverage of medical treatment for diabetics. Cohen and Einav (2003), however, find no evidence of a connection between increased use of seatbelts and riskier driving.

B. Randomized Experiments and Natural Experiments

When insurance coverage of a set of policyholders changes for exogenous reasons—a randomized experiment or a natural experiment—it is reasonable to assume that the change affects only the policyholders' behavior and not their underlying risk. Therefore, insofar as the coverage changes lead to changes in policyholders' losses (and as long as the changes are solely due to exogenous factors and not to policyholders' choices), such a pattern may be explained by moral hazard but not by adverse selection.

Manning et al. (1987) used the RAND Health Insurance Experiment to test whether individuals who were randomly assigned more coverage chose higher levels of spending. The study found evidence consistent with *ex post* moral hazard: individuals in plans with more coverage spent more on health care and, therefore, were more costly to their insurer.

Since randomized experiments are often impossible or, at least, quite expensive to conduct, researches often take advantage of natural or quasi-natural experiments such as changes brought about by new regulations and policies. Chiappori, Durand, and Geoffard (1998) exploited such an exogenous change in French health insurance: the replacement of full coverage with a 10% copayment in 1994. This change could reasonably have been expected to affect the

incentives of policyholders, without changing the composition of the insured pool. The study found moral hazard in some dimensions of health insurance use but not in others. In particular, the introduction of the copayment had no effect on doctor office visits but did have an effect on doctor home visits.

Cardon and Hendel (2001) use data from the 1987 Natural Medical Expenditure Survey (NMES), including all health policies offered by employers to each of their employees, and the policy the worker actually chose. In the data, employees working for the same firm are offered the same menu, although employees from different firms face widely different choice sets. Insofar as employees do not select employers on the basis of the health-insurance menus that the employer offers, the situation examined by the authors is one in which employees with similar characteristics choose different contracts for exogenous reasons. To test for adverse selection, Cardon and Hendel examine similar employees who face the same set of choices. To test for moral hazard, they examine similar employees who face different coinsurance rates, with price sensitivity identified using the coinsurance variability across individuals. While they find no evidence of adverse selection, they do find price elasticities that are negative and close to those obtained in the RAND survey. The authors conclude that moral hazard, rather than adverse selection, is likely to be the main consequence of asymmetric information in their data.

Finally, Dionne et al. (2005) focus on a 1992 change by Quebec's monopoly public insurer in the sale of drivers' insurance against bodily injuries caused to others. The change was such that the insurer began increasing the premiums that it charged to drivers who accumulated demerit points through traffic violations. Consistent with moral hazard, the change was associated with a reduction in accidents.

C. Dynamic Properties

Another approach to distinguishing between moral hazard and adverse selection is based on dynamic properties of the two conditions, since moral hazard and adverse selection generate a different relationship between past and future risks. Abbring et al. (2003a, 2003b) provide a model that explains how dynamic data can help in distinguishing moral hazard from adverse selection. Insurance contracts often make premiums dependent on the policyholder's prior claim history. Under commonly used forms of the "bonus-malus" system, the cost of an accident in terms of future premiums depends on the number of previous accidents. Given this feature of insurance contracts, moral hazard should lead to a negative correlation between prior claims and accidents in a subsequent year. Under adverse selection, in contrast, prior claims reflect a policyholder's risk type; therefore, we should expect a *positive* correlation between past and future claims. Abbring et al. (2003b) apply this approach to data from France, where all insurers

are required by regulation to use a bonus-malus system, and they find no evidence of moral hazard. A critical aspect of their approach is control for unobserved heterogeneity.

Israel (2007) follows the approach of Abbring et al. (2003a, b), using ten-year tracking data from a private automobile insurer in Illinois on 30,000 policyholders. In the market that he studies, a driver's experience rating is based only on his or her claims history in the previous three years. As a result, after three years, each claim drops off the policyholder's record and changes the policyholder's position in the experience rating scheme, irrespective of whether any claim is submitted at that time. This feature of the data allows the researcher to isolate moral hazard without using restrictive assumptions about state dependence, as in Abbring et al. (2003a). The study finds a small but statistically significant evidence of moral hazard. But the effect disappears when state dependence is not controlled for, which highlights the importance of controlling for state dependence.

Another study that focuses on dynamics is Dionne et al. (2006), which uses a longitudinal sample of French automobile insurance policies during the period from 1995-1997. The dataset has the advantage of including both the number of accidents and the number of claims made by the policyholder. This allows the authors to base their test on actual accidents, rather than claims, with two advantages. First, using actual accident data provides for superior estimation of policyholders' risks. Moreover, the panel structure also allows the researchers to follow policyholders who switch to another insurer. The paper proposes a dynamic causality test for the separation of moral hazard from learning and adverse selection. The study finds asymmetric learning among policyholders with five or fewer years of driving experience, moral hazard in the group of policyholders with less than fifteen years of experience, and no residual information problem for policyholders with more than 15 years of experience. The authors attribute this result to the fact that, under the French system of bonus-malus, older policyholders have the strongest incentives for safe driving because a large portion of these policyholders have the maximum bonus-malus socio and they are motivated to stay there.

D. Interaction of the Coverage–Risk Correlation with Policyholder Characteristics and Behavior

Finally, a third approach to untangling moral hazard and adverse selection utilizes a static, single-period analysis but focuses on disaggregating the coverage–risk correlation in ways that can help distinguish between the two explanations for the correlation. This approach is pursued by Cohen (2005), who, after finding a coverage–risk correlation in the set of all policies sold to an insurer's new customers, goes on to show that the identified correlation interacts with policyholder characteristics in a way that is easier to explain under adverse selection than under moral hazard. In particular, the study shows that the correlation exists only for policyholders who have three or more years of driving experience. Consistent with adverse selection, this pattern

may be explained by customers' obtaining private information about their risk type only after amassing some driving experience. In contrast, moral hazard can not readily explain this pattern unless one argues that expected losses from accidents are sensitive to precautions taken by experienced drivers, but not those taken by new drivers.

Furthermore, the study shows that there is some underreporting of prior claims by new customers with previous driving experience. Because the insurer cannot tell which new customers under–report, this pattern provides direct evidence that at least some new customers have private information about their risk type. In contrast, such underreporting is not part of a moral hazard story in which customers with different coverage levels present different risks due to different behavior (as opposed to different risk types that they had before they bought the policies).

VI. Learning over Time

Adverse selection involves asymmetric information between policyholders and insurers with respect to policyholders' risk types. Such information is not static. Over time, both policyholders and insurers may learn information about policyholders' risk types. In this section, we review the evidence about the existence and consequences of such learning.[29]

Studying the life-insurance market, Hendel and Lizzeri (2003) provide strong evidence of the existence and significance of learning over time. Their focus is on symmetric learning, in which both policyholders and insurers gradually gain new information about the policyholder's risk of mortality. In this setting, purchasing a new contract each year, with no renewal commitment on the part of either the insurer or the policyholder, would leave policyholders with an uninsured risk of being reclassified into a high-risk category in the future and having to pay the resulting high premium. Insuring against this risk is addressed by long-term contracts with a commitment to renew on the part of the insurer but not the policyholder; the problem of future defection by policyholders who learn that they have higher-than-average mortality risk is checked by the front-loading of premiums. The Hendel-Lizzeri study finds evidence that is consistent with such front-loading and the resulting partial lock-in of consumers, the predicted response to learning over time in this context.

Finkelstein, McGarry, and Sufi (2005) find evidence in support of the Hendel-Lizzeri findings using US data on long-term care insurance, a market that also involves learning about mortality risk over time. The study shows that long-term care insurance contracts also involve

[29] We abstract from studies (Crawford and Shum (2005) and Israel (2005a, 2005b)) of learning by policyholders about the quality of their insurer's service, learning that does not involve information about policyholders' risk types.

substantial front-loading, with policyholders paying initial premiums that exceed their actuarial costs to the insurer. The study also provides evidence that policyholders who drop their coverage are subsequently less likely to use a nursing home, and that the dropping of coverage is at least partly a response to positive information about the policyholder's health situation.

These studies show the existence of learning about risk types, but they focus on symmetric learning that has no effect on any information asymmetry that may exist between policyholders and insurers. Learning need not be symmetric, however. Some information may reach only some but not all agents in the market and, therefore, may affect the presence of adverse selection. Furthermore, there is empirical evidence that such learning takes place.

In particular, experience may allow policyholders to amass information that allows them to gain an advantage over their insurer. Studying a market in which there is no information sharing among insurers, Cohen (2005) shows the existence of learning by policyholders. In particular, by finding a coverage–risk correlation among new customers with three or more years of driving experience but uncovering no such correlation among new customers with less driving experience, and by also finding that new customers underreport their prior claims histories, the study shows that prior learning by customers can be relevant to the existence of adverse selection and coverage–risk correlation.[30]

Given that policyholders with driving experience can underreport prior claims when switching to a new insurer and insofar as insurance policies are for one term only, each insurer has an informational advantage over its rivals with respect to its repeat customers. Such an information asymmetry *among* insurers provides each of them with market power vis-à-vis its current customer base. This market power, and the higher profits it may facilitate, offer another dimension for empirical work that seeks to identify learning.

In the auto-insurance market, D'Arcy and Doherty (1990) show that, consistent with learning by insurers about their repeat customers, insurers contractually forbid their agents from selling private information about customer records to rival insurers, and that entry into the insurance market is costly. The study also demonstrates that the longer various policyholder cohorts remain with the company, the less profit the company makes on them. However, the authors lacked data that would allow them to control for differences in policyholder characteristics other than experience with the company and, therefore, could not rule out the

[30] Israel (2006) obtains results similar to those of Cohen (2005) using data from the US. Studying a pool of an insurer's new customers who bought more than the legally required minimum insurance, the study finds coverage–risk correlation only among "informed consumers," defined as consumers who are more than forty years old when joining the insurer.

possibility that the compared cohorts of policyholders differed in aspects other than their experience with the company.

Cohen (2008) uses a panel dataset of repeated contracting that includes all information that the insurer has about each policyholder, including information relevant to assessing the policyholder's risk and calculating the insurer's profits on each customer. The study yields three findings that are consistent with the insurer's having private information about repeat customers that rival insurers do not share. First, the insurer makes higher profits on repeat customers. And, more importantly, these higher profits are driven by profits made on customers who have good records with the insurer. If a repeat customer with a good claims record with the insurer were to switch to a rival, the rival might be uncertain about his or her quality; this state of affairs provides the insurer with some information-based market power over those customers who have a good record. Furthermore, the longer the period over which the customer has a good record, the higher the profits made by the insurer.

Second, the study finds that the insurer's higher profits on repeat customers with good claims records are due to its ability to charge such customers a higher premium than their low risk justifies. While the insurer charges low-risk repeat customers lower premiums, the reduction in premium is smaller than the reduction in costs (insurance payouts) associated with such repeat customers. Finally, consistent with the insurer's possession of private information about repeat customers' risk type, the study finds that customers with poor claim records tend to leave the insurer while those with good records tend to stay for another term. Customers with a bad claims history have an incentive to flee their record and pool with customers who leave their insurers because of exogenous shocks.

It is worth noting that the analysis of learning about repeat customers applies only to the (many) insurance markets in which policies are issued for one term only. Building on earlier theoretical work (Dionne and Lasserre (1985), Cooper and Hayes (1987)), Dionne and Doherty (1994) show that it may be advantageous for insurers facing adverse selection to commit to offer policyholders an option to renew their policies at a specified price, and they provide evidence consistent with such behavior on the basis of automobile-insurance data from California.

It should also be stressed that the ability of insurers to obtain an informational advantage over their rivals with respect to repeat customers depends on the extent to which there are systems—resulting from regulation, contract, or otherwise—for information-sharing among insurers. Such requirements may eliminate the underreporting of prior claims by experienced drivers who join a new insurer. The desirability of information-pooling systems has been much debated in Europe, where the European Commission ruled that some systems of information sharing are anticompetitive (de Garidel-Thoron (2005)).

In the U.S., some information about accident history is maintained in state records, but it has long been observed (e.g., by D'Arcy and Doherty (1990)) that the publicly available information is highly incomplete and that a new insurer would not be able to cull from it a complete claims history of other insurers' customers. A lawsuit that reached the US Supreme Court focused on an attempt by several large insurers and reinsurers to force the Commercial General Liability insurance industry to use standard forms, and it was suggested by some that this attempt was partly motivated by a desire to make experience-sharing easier.[31] Future empirical work should try to identify the consequences of information-sharing systems.

It deserves emphasis that, even though information-sharing systems can eliminate the underreporting of prior claims to a new insurer, they do not necessarily eliminate all types of asymmetric learning. In particular, driving experience may provide policyholders with information about their risk type that is not fully reflected in claims records—for example, information about accidents that were narrowly avoided by the policyholder or the policyholder's driving ability. Whether such learning is significant is another interesting subject for future empirical research. Cohen and Einav (2007) provide some evidence suggesting that the incidence of accidents that are unreported because the level of damages is just below the deductible is not large, but there may be many accidents that were narrowly avoided or that ended with very minor damage.

VII. Conclusion

In the first decade following the appearance of Rothschild and Stiglitz's seminal paper, work on adverse selection was largely theoretical. Since then, empirical work has done an impressive job of catching up, so that most research on adverse selection in insurance markets in the past decade has been empirical.

We have examined a large number of studies that test the basic coverage–risk correlation predicted by adverse selection across a wide range of insurance markets. Although the studies yield different findings about the existence of the coverage–risk correlation, we argued that the fact that the correlation is found by some studies and not by others does not indicate that work in this area is still at an inconclusive stage, or that more consistent results should be expected to appear over time. Rather, our assessment of the work in this area is that there is a good basis for expecting the existence of adverse selection and a coverage–risk correlation to vary across markets and, indeed, even across segments of the same market. What we should expect empirical work to provide in this area, then, is not a once-and-for-all answer to the question about whether adverse selection exists but rather an ever-improving understanding of the circumstances under

[31] See *Hartford Fire Insurance Co. v. California*, 509 U.S. 764 (1993), and, for commentary, Ayres and Siegelman (1989).

which a coverage–risk correlation should and should not be expected to arise. We have sought in this paper to advance this objective by providing a classification of reasons why a coverage–risk correlation may not exist and a framework for thinking about them.

Our conclusions have important implications for policy discussions in this area. Such discussions should not be based on a general assumption that adverse selection and the coverage–risk correlation may be expected to exist generally in insurance markets, nor should they rely on studies that fail to find such a correlation to suggest the opposite general assumption. Policy discussions should try to tailor themselves to the specific insurance market under consideration, recognizing that adverse selection and coverage–risk correlations vary across insurance markets (and even among pools of risks within a market), and that they do so in ways that are at least somewhat predictable on the basis of existing research.

One important direction for subsequent work is to further study and understand the factors that cause the coverage–risk correlation to arise in some circumstances but not in others. More work is warranted with respect to disentangling moral hazard and adverse selection and with respect to learning over time. Furthermore, transcending the question of when and to what extent adverse selection occurs, future research should examine the consequences of adverse selection, when it occurs, for the supply and purchase of insurance and, in turn, for efficiency and welfare.[32] As researchers gain better and better access to comprehensive datasets that allow them to perform tests that they could not have performed earlier, we hope that significant progress will be made in all of these directions in the years to come.

[32] A recent study that moves in this direction is Einav, Finkelstein, and Schrimpf (2007). They estimate that adverse selection causes a welfare loss of about 2% of annual premiums in the UK annuities market.

References

Abbring, Jaap, Pierre-André Chiappori, James Heckman, and Jean Pinquet (2003a), "Adverse Selection and Moral Hazard in Insurance: Can Dynamic Data Help to Distinguish?" *Journal of The European Economic Association*, Vol. 1 (2–3), pp. 512–521.

Abbring, Jaap, Pierre-André Chiappori, and Jean Pinquet (2003b), "Moral Hazard and Dynamic Insurance Data," *Journal of the European Economic Association*, Vol.1(4), pp. 767–820.

Arrow, Kenneth J. (1963), "Uncertainty and the Welfare Economics of Medical Care**,"** *The American Economic Review*, Vol. 53 (5), pp. 941–973.

Ayres, Ian, and Peter Siegelman (1989), "The Economics of the Insurance Anti-Trust Suits: Toward an Exclusionary Theory," *Tulane Law Review*, Vol. 63, pp. 971–996.

Baker, Tom (1996), "On the Genealogy of Moral Hazard," *Texas Law Review*, Vol. 75, pp. 237–292.

Baker, Tom, and Siegelman, Peter (2010), "Tontines for the Invincibles," Wisconsin Law Review.

Barsky, Robert B., Frank T. Juster, Miles Kimball, and Matthew Shapiro (1997), "Preference Parameters and Behavioral Heterogeneity: An Experimental Approach in the Health and Retirement Study," *Quarterly Journal of Economics*, Vol. 112, pp. 537–579.

Bhattacharya, Jayanta, Dana Goldman, and Neeraj Sood (2003), "Market Evidence of Misperceived Prices and Mistaken Mortality Risks," NBER Working Paper No. 9863.

Boyer, Marcel, and Georges Dionne (1989), "An Empirical Analysis of Moral Hazard and Experience Rating," *Review of Economics and Statistics*, Vol. 71, pp. 128–134.

Brockett, Patrick L., and Linda L. Golden (2007), "Biological and Psycho-behavioral Correlates of Credit Scores and Automobile Insurance Losses: Towards an Explanation of Why Credit Scoring Works," *Journal of Risk & Insurance*, Vol. 74, pp. 22–63.

Brown, Jeffrey and Amy Finkelstein (2007), "Why is the Market for Long Term Care Insurance so Small," *Journal of Public Economics*, Vol.91, No. 10, pp. 1967–1991.

Browne, Mark J., and Helen I. Doerpinghaus (1993), "Information Asymmetries and Adverse Selection in the Market for Individual Medical Expense Insurance," *Journal of Risk and Insurance*, Vol. 60, pp. 300–312.

Buchmueller, Thomas, Agnes Couffinahl, Michel Grignon and Marc Perronnin (2004), "Access to Physician Services: Does Supplemental Insurance Matter? Evidence from France," *Health Economics*, Vol. 13, pp. 669–687.

Butler, Patrick (1996), "Automobile Insurance Pricing: Operating versus Ownership Costs," Ch. 39 in *Women's Travel Issues: Proceedings from the Second National Conference*, US Department of Transportation, Federal Highway Administration Publ. FHWA-PL-97-024.

Cardon, James H., and Igal Hendel (2001), "Asymmetric Information in Health Insurance: Evidence from the National Health Expenditure Survey," *RAND Journal of Economics*, Vol. 32, pp. 408–427.

Cawley, John, and Tomas Philipson (1999), "An Empirical Examination of Information Barriers to Trade in Insurance," *American Economic Review*, Vol. 89, pp. 827–846.

Chiappori, Pierre-André (1999), "Asymmetric Information in Automobile Insurance: An Overview," in *Automobile Insurance: Road Safety, New Drivers, Risks, Insurance Fraud and Regulation*, ed. Georges Dionne and Claire Laberge-Nadeau, pp. 1–12 (Norwell, MA: Kluwer).

Chiappori, Pierre-André (2000), "Econometric Models of Insurance under Asymmetric Information," in *Handbook of Insurance*, ed. Georges Dionne, pp. 365–394 (Norwell, MA: Kluwer).

Chiappori, Pierre-André (2006), "The Welfare Effects of Predictive Medicine," in *Competitive Failures in Insurance Markets: Theory and Policy Implications*, ed. Pierre-André Chiappori and Christian Gollier (Cambridge, MA: MIT Press).

Chiappori, Pierre-André, Franck Durand, and Pierre Y. Geoffard (1998), "Moral Hazard and the Demand for Physician Services: First Lessons from a French Natural Experiment," *European Economic Review*, Vol. 42, pp. 499–511.

Chiappori, Pierre-Andre, and Bernard Salanié (1997), "Empirical Contract Theory: The Case of Insurance Data," *European Economic Review*, Vol. 41, pp. 943–950.

Chiappori, Pierre-André, and Bernard Salanié (2000), "Testing for Asymmetric Information in Insurance Markets," *Journal of Political Economy*, Vol. 108, pp. 56–78.

Chiappori, Pierre-André, and Bernard Salanié (2003), "Testing Contract Theory: A Survey of Some Recent Work," in *Advances in Economics and Econometrics: Theory and Applications, Eighth World Congress* Vol 1, 115-149, ed. Mathias Dewatripont, Lars P. Hansen and Stephen Turnovsky (Cambridge: Cambridge University Press).

Chiappori, Pierre-André, Bruno Jullien, Bernard Salanié and François Salanié (2006) "Asymmetric Information in Insurance: General Testable Implications," *Rand Journal of Economics*, Vol. 37, pp. 783–798.

Cohen, Alma (2005), "Asymmetric Information and Learning in the Automobile Insurance Market," *The Review of Economics and Statistics*, Vol. 87, pp. 197–207.

Cohen, Alma (2006), "The Disadvantages of Aggregate Deductibles," *B.E. Journals in Economic Analysis & Policy*, Vol. 6, No. 1, article 7.

Cohen, Alma (2008), "Asymmetric Learning in Repeated Contracting: An Empirical Study," NBER Working Paper No. 13752.

Cohen, Alma, and Rajeev Dehejia (2004), "The Effects of Automobile Insurance and Accident Liability Laws on Traffic Fatalities," *Journal of Law and Economics*, Vol. 47 No. 2, pp. 357–393.

Cohen, Alma, and Liran Einav (2003), "The Effect of Mandatory Seat Belt Laws on Driving Behavior and Traffic Fatalities," *Review of Economics and Statistics*, Vol. 85, pp. 828–843.

Cohen, Alma, and Liran Einav (2007), "Estimating Risk Preferences from Deductible Choice," *American Economic Review*, Vol. 97, No. 3, pp. 745–788.

Cooper, Russell, and Beth Hayes (1987), "Multi-period Insurance Policies," *International Journal of Industrial Organization*, Vol. 5, pp. 211–231.

Crawford, Gregory, and Matthew Shum (2005), "Uncertainty and Learning in Pharmaceutical Demand," *Econometrica*, 73, pp. 1137–1173.

Cummins, J. David, and Neil A. Doherty (2006), "The Economics of Insurance Intermediaries," *Journal of Risk and Insurance*, Vol. 73, pp. 359–396.

Cutler, David M., and Sarah J. Reber (1998), "Paying For Health Insurance: The Trade-Off Between Competition and Adverse Selection," *The Quarterly Journal of Economics*, Vol. 113, pp. 433–466.

Cutler, David M., Amy Finkelstein, and Kathleen McGarry (2008), "Preference Heterogeneity in Insurance Markets: Explaining a Puzzle," *American Economic Review Papers and Proceedings*, Vol. 98(2), pp. 157–162.

Cutler, David M., and Richard J. Zeckhauser (1998) "Adverse Selection in Health Insurance," *Forum for Health Economics & Policy*: Vol. 1: (Frontiers in Health Policy Research), Article 2.

Cutler, David M., and Richard J. Zeckhauser (2000), "The Anatomy of Health Insurance," in Anthony J. Culyer and Joseph P. Newhouse, eds. *Handbook of Health Economics*, Vol. 1, Ch. 11, 563–643 (Amsterdam: Elsevier Science).

Dahlby, Bev (1983), "Adverse Selection and Statistical Discrimination: An Analysis of Canadian Automobile Insurance," *Journal of Public Economics*, Vol. 20, pp. 121–130.

Dahlby, Bev (1992), "Testing for Asymmetric Information in Canadian Automobile Insurance," in Georges Dionne, ed., *Contributions to Insurance Economics*, pp. 423–443, (Boston:

Kluwer Academic Publishers).

Davidoff, Thomas, and Gerd Welke (2005), "Selection and Moral Hazard in the Reverse Mortgage Market," working paper, Haas School of Business, University of California, Berkeley, available at http://faculty.haas.berkeley.edu/davidoff/RMsubmit1.pdf.

D'Arcy, Stephen P., and Neil A. Doherty (1990), "Adverse Selection, Private Information, and Lowballing in Insurance Markets," *Journal of Business*, Vol. 63, pp. 145–164.

De Garidel-Thoron, Thomas (2005), "Welfare-Improving Asymmetric Information in Dynamic Insurance Markets," *Journal of Political Economy*, Vol. 113, pp. 121–150.

De Meza, David, and David C. Webb (2001), "Advantageous Selection in Insurance Markets," *Rand Journal of Economics*, Vol. 32, No. 2, pp. 249–262.

DeDonder, Philippe, and Jean Hindriks (2006), "Does Propitious Selection Explain Why Riskier People by Less Insurance?" CEPR Working Paper DP5640.

Dionne, Georges, and Neil A. Doherty (1992), "Adverse Selection in Insurance Markets: A Selective Survey," in Georges Dionne, ed., *Contributions to Insurance Economics* (Boston: Kluwer Academic Publishers), pp. 97–140.

Dionne, Georges, and Neil A. Doherty (1994), "Adverse Selection, Commitment and Renegotiation: Extension to and Evidence from Insurance Markets," *Journal of Political Economy*, Vol. 102 No. 2, pp. 210–235.

Dionne, Georges, Neil A. Doherty and Nathalie Fombaron (2001), "Adverse Selection in Insurance Markets," in Georges Dionne, ed., *Handbook of Insurance* (Boston: Kluwer Academic Publishers), pp. 185–243.

Dionne, Georges and Robert Gagné (2001) "Deductible Contracts Against Fraudulent Claims: Evidence From Automobile Insurance," *The Review of Economics and Statistics*, Vol. 83(2), pages 290-301.

Dionne, Georges, Christian Gouriéroux and Charles Vanasse (2001), "Testing for Evidence of Adverse Selection in the Automobile Insurance Market: A Comment," *Journal of Political Economy*, Vol. 109, pp. 444–473.

Dionne, Georges, and Pierre Lasserre (1985), "Adverse Selection, Repeated Insurance Contracts and Announcement Strategy," *Review of Economic Studies*, Vol. 50, pp. 719–723.

Dionne, Georges, Mathieu Maurice, Jean Pinquet, and Charles Vanasse (2005), "The Role of Memory in Long-Term Contracting with Moral Hazard: Empirical Evidence in Automobile Insurance," Available at SSRN: http://ssrn.com/abstract=764705. (Accessed June 30, 2008.)

Dionne, Georges, Pierre-Carl Michaud, and Maki Dahchour (2007), "Separating Moral Hazard

from Adverse Selection and Learning in Automobile Insurance: Longitudinal Evidence from France," Available at SSRN: http://ssrn.com/abstract=583063. (Accessed June 30, 2008.)

Dionne, Georges, Pascal St-Amour and Désiré Vencatachellum (2008), "Asymmetric Information and Adverse Selection in Mauritian Slave Auctions" (mimeo).

Dionne, Georges, and Pierre St-Michel (1991), "Worker's Compensation and Moral Hazard," *Review of Economics and Statistics*, Vol. 73, pp. 236–244.

Dohmen, Thomas, Armin Falk, David Huffman, Uwe Sunde, Jürgen Schupp and Gert G. Wagner (2005), "Individual Risk Attitudes: New Evidence from a Large, Representative, Experimentally-Validated Survey," IZA Discussion Paper No. 1730.

Dohmen, Thomas, Armin Falk, David Huffman and Uwe Sunde (2007), "Are Risk Aversion and Impatience Related to Cognitive Ability?" IZA Discussion Paper No. 2735.

Edlin, Aaron S. (2003), "Per-Mile Premiums for Auto Insurance," in Richard Arnott et al., eds., *Economics for an Imperfect World: Essays in Honor of Joseph Stiglitz,* (Cambridge, MA: MIT Press), pp. 53–82.

Einav, Liran, Amy Finkelstein, and Paul Schrimpf (2007), "The Welfare Costs of Asymmetric Information: Evidence from the U.K. Annuity Market," NBER Working Paper 13228.

Ettner, Susan L. (1997), "Adverse Selection and the Purchase of Medigap Insurance by the Elderly," *Journal of Health Economics*, Vol. 16, pp. 543–562.

Fang, Hanming, Michael P. Keane, and Dan Silverman (2008), "Sources of Advantageous Selection: Evidence from the Medigap Insurance Market," *Journal of Political Economy*, Vol. 116, pp. 303–350.

Finkelstein, Amy, and Kathleen McGarry (2006), "Multiple Dimensions of Private Information: Evidence from the Long-Term Care Insurance Market," *American Economic Review*, Vol. 96, pp. 938–958.

Finkelstein, Amy, Kathleen McGarry, and Amir Sufi (2005), "Dynamic Inefficiencies in Insurance Markets: Evidence from Long-Term Care Insurance," *American Economic Review Papers and Proceedings*, Vol. 95, pp. 224–228.

Finkelstein, Amy, and James Poterba (2004), "Adverse Selection in Insurance Markets: Policyholder Evidence from the U.K. Annuity Market," *Journal of Political Economy*, Vol. 112, pp. 183–208.

Finkelstein, Amy, and James Poterba (2002), "Selection Effects in the United Kingdom Individual Annuities Market," *The Economic Journal*, Vol. 112, pp. 28–50.

Finkelstein, Amy, and James Poterba (2006) "Testing for Adverse Selection with 'Unused

Observables','' NBER Working Paper No. 12112.

Fortin, Bernard, and Paul Lanoie (2000), "Incentive Effects of Workers' Compensation: A Survey," in Georges Dionne, ed., *Handbook of Insurance*, Ch. 13 (Boston: Kluwer Academic Publishers), pp. 421–458.

Friedman, Benjamin M., and Mark J. Warshawsky (1990), "The Cost of Annuities: Implications for Savings Behavior and Bequests," *Quarterly Journal of Economics,* Vol. 420, pp. 135–154.

Grove, William M., and Paul E. Meehl (1996), "Comparative Efficiency of Informal (Subjective, Impressionistic) and Formal (Mechanical, Algorithmic) Prediction Procedures: the Clinical-Statistical Controversy," *Psychology, Public Policy and Law*, Vol. 2, pp. 1–31.

Guiso, Luigi, and Monica Paiella (2003), "Risk Aversion, Wealth and Background Risk," Banca D'Italia Working Paper 483.

Guiso, Luigi, and Monica Paiella (2005), "The Role of Risk Aversion in Predicting Individual Behavior," Banca D'Italia Working Paper 546.

Guppy, Andrew (1993), "Subjective Probability of Accident and Apprehension in Relation to Self-Other Bias, Age, and Reported Behavior," *Accident Analysis and Prevention*, Vol. 25, No. 4, pp. 375–382.

Hall, D. Jonathan (2006), "Adverse Selection in the Market for Health Insurance," Mimeo, Brigham Young University Department of Economics.

Hamermesh, Daniel S. (1985), "Expectations, Life Expectancy, and Economic Behavior," *Quarterly. Journal of Economics*, Vol. 10, pp. 389–408.

Harris, Milton, and Artur Raviv (1978), "Some Results on Incentive Contracts with Applications to Education and Employment, Health Insurance, and Law Enforcement," *American Economic Review*, Vol. 68, No. 1, pp. 20–30.

Hemenway, David (1990), "Propitious Selection," *Quarterly Journal of Economics*, Vol. 105, pp. 1073–1069.

Hendel, Igal, and Alessandro Lizzeri (2003), "The Role of Commitment in Dynamic Contracts: Evidence from Life Insurance," *Quarterly Journal of Economics*, Vol. 118, pp. 299–327.

Hölmstrom, Bengt (1979), "Moral Hazard and Observability," *The Bell Journal of Economics,* Vol. 10, No. 1, pp. 74–91.

Hosios, J. Arthur, and Michael Peters (1989), "Repeat Insurance Contracts with Adverse Selection and Limited Commitment," *Quarterly Journal of Economics*, Vol. 104, No. 2. pp. 229–253.

Hoy, Michael, and Julia Witt (2007), "Welfare Effects of Banning Genetic Information in the Life Insurance Market: The Case of BRCA1/2 Genes," *Journal of Risk and Insurance* Vol. 74, pp. 523–546.

Hurd, Michael D., James P. Smith and Julie M. Zissimopoulos (2002), "The Effects of Subjective Survival on Retirement and Social Security Claiming," NBER working paper No. 9140.

Israel, Mark (2005a), "Tenure Dependence in Consumer-Firm Relationships: An Empirical Analysis of Consumer Departures from Automobile Insurance Firms," *RAND Journal of Economics*, Vol. 36, pp. 165–192.

Israel, Mark (2005b), "Services as Experience Goods: An Empirical Examination of Consumer Learning in Automobile Insurance," *American Economic Review*, Vol. 95, No. 5, pp. 1444–1463.

Israel, Mark (2006), "Where Is All the Hidden Information Hiding? Evidence from Automobile Insurance Panel Data," *mimeo*, Northwestern University.

Israel, Mark (2007), "Do We Drive More Safely When Accidents Are More Expensive? Identifying Moral Hazard from Experience Rating Schemes," *mimeo*, Wharton School, University of Pennsylvania.

Just, Richard E., Linda Calvin, and John Quiggin (1999), "Adverse Selection in Crop Insurance: Actuarial and Asymmetric Information Incentives," *American Journal of Agriculture Economics*, Vol. 81, pp. 834–849.

Karagyozova, Tsvetanka, and Peter Siegelman (2007), "Propitious Selection in Insurance Markets: A Simulation Study Using Continuous Distributions," Working Paper, University of Connecticut Law School.

Kaestner, Robert, and Anne Carroll (1997), "New Estimates of the Labor Market Effects of Workers' Compensation Insurance," *Southern Economic Journal*, Vol. 63, No. 3, pp. 635–651.

Klick, Jonathan, and Thomas Stratmann (2007), "Diabetes Treatments and Moral Hazard," *The Journal of Law and Economics,* Vol. 50, No. 3, pp. 519–538.

Koufopolous, Kostas (2007), "On the Positive Correlation Property in Competitive Insurance Markets," *Journal of Mathematical Economics*, Vol. 43, pp. 597–605.

Löfgren, Karl-Gustav, Torsten Persson, and W. Jörgen Weibull (2002), "Markets with Asymmetric Information: The Contributions of George Akerlof, Michael Spence and Joseph Stiglitz," *Scandinavian Journal of Economics*, Vol. 104, No. 2, pp. 195–211.

Loewenstein, George F., Christopher K. Hsee, Elke U. Weber and Ned Welch (2001), "Risk as

Feelings," *Psychological Bulletin*, Vol. 127, pp. 267–286.

McCarthy, David, and Olivia S. Mitchell (forthcoming), "International Adverse Selection in Life Insurance and Annuities," in Shripad Tuljapurkar, Naohiro Ogawa, and Anne Gauthier, eds. *Riding the Age Waves: Responses to Aging in Advanced Industrial States* (Amsterdam: Elsevier).

Makki, Shiva S., and Agapi Somwaru (2001), "Evidence of Adverse Selection in Crop Insurance Markets," *Journal of Risk and Insurance*, Vol. 68, No. 4, pp. 685–708.

Manning, Willard G., Joseph P. Newhouse, Naihua Duan, Emmett B. Keeler and Arleen Leibowitz (1987), "Health Insurance and the Demand For Medical Care: Evidence From a Randomized Experiment," *American Economic Review*, Vol. 77(3), pp. 251–277.

Paltrow, Scott (2001), "Old Memos Show in Black and White MetLife's Use of Race to Screen Clients," *Wall Street Journal*, July 24, p. 1.

Pauly, Mark V. (1974), "Overinsurance and Public Provision of Insurance: The Roles of Moral Hazard and Adverse Selection," *The Quarterly Journal of Economics*, Vol. 88 (1), pp. 44–62.

Pauly, Mark V., Kate H. Withers, Krupa S. Viswanathan, Jean Lemaire, John C. Hershey, Katrina Armstrong and David A. Asch (2003), "Price Elasticity of Demand for Term Life Insurance and Adverse Selection," NBER Working Paper No. 9925.

Priest, George L. (1987), "The Current Insurance Crisis and Modern Tort Law," *The Yale Law Journal*, Vol. 96, No. 7, pp. 1521–1541.

Puelz, Robert, and Arthur Snow (1994), "Evidence on Adverse Selection: Equilibrium Signaling and Cross-Subsidization in the Insurance Market," *Journal of Political Economy*, Vol. 102, pp. 236–257.

Richaudeau, Didier (1999), "Automobile Insurance Contracts and Risk of Accident: An Empirical Test Using French Individual Data," *The Geneva Papers on Risk and Insurance Theory*, Vol. 24, pp. 97–114.

Rothschild, Michael, and Joseph Stiglitz (1976), "Equilibrium in Competitive Insurance Markets," *Quarterly Journal of Economics*, Vol. 90, No. 4, pp. 629–649.

Romano, Roberta (1989), "What Went Wrong with Directors' and Officers' Liability Insurance?" *Delaware Journal of Corporate Law*, Vol. 14, pp. 1–27.

Saito, Kuniyoshi (2006) "Testing for Asymmetric Information in the Automobile Insurance Market Under Rate Regulation," *Journal of Risk & Insurance*, Vol. 73, No. 2, pp. 335–356.

Schwarcz, Daniel B. (2007), "Beyond Disclosure: The Case for Banning Contingent

Commissions," *Yale Law and Policy Review*, Vol. 25, pp. 289–331.

Shavell, Steven (1979), "On Moral Hazard and Insurance," *The Quarterly Journal of Economics*, Vol. 93, No. 4, pp. 541–562.

Siegelman, Peter (2004), "Adverse Selection in Insurance Markets: An Exaggerated Threat," *The Yale Law Journal*, Vol. 113 (6), pp. 1123–1281.

Sil, Jayashree (2005), "Endogenous Adverse Selection: Evidence from U.S. Crop Insurance," Mimeo, University of California, Berkeley, Department of Agricultural and Resource Economics, available at http://are.berkeley.edu/~jsil/jobmarket.html.

Svenson, Ola (1981), "Are We All Less Risky and More Skillful Than Our Fellow Drivers?" *Acta Psychologica*, Vol. 47, No. 2, pp. 143–148.

Vickrey, William (1968), "Automobile Accidents, Tort Law, Externalities and Insurance: An Economist's Critique," *Law & Contemporary Problems*, Vol. 33, No. 3, pp. 464–487.

Winter, Ralph (2000), "Optimal Insurance under Moral Hazard," in Georges Dionne, ed., *Handbook of Insurance* (Boston: Kluwer Academic Publishers).

Table 1: Evidence on Adverse Selection in Insurance Markets

Author	Data Description	Empirical Method	Complete Info?	Asymmetry Found?	Comments
			Automobiles		
Puelz & Snow (1994)	3280 individuals from single insurer. US: 1986	Ordered logit for the deductible, using observed number of accidents as RHS variable	No	Yes	Those with higher accident risk choose higher deductibles. Heterogeneous group by age/experience, and effect of accidents constrained to be linear.
Richaudeau (1999)	5703 individuals from national survey. France: 1995	Probit estimate of choice of deductible; this est. prob. used in negative binomial model of number of accidents.	Unclear: survey contains some info insurer did not have, but lacks premium data.	No/Yes	Accident risk not associated with increased insurance purchased (total miles driven controlled for); but those who drove more did buy more insurance.
Chiappori & Salanié (2000)	20,716 drivers who got licenses in 1986–88, insured by any of 21 companies. France: 1989	Several parametric and non-parametric tests	Unclear	No	Those with higher accident risk do not purchase more coverage. Relatively homogeneous group by age/experience.
Dionne, Doherty and Fombaron (2001)	4772 drivers covered by a private insurer. Quebec: date uncl.	As in Puelz and Snow, plus additional specifications	Unclear	No	No coverage/risk correlation observed. Replication of Puelz and Snow methods with different data, and with nonlinear effects fully accounted for.
Cohen (2005)	Complete data used by insurer. Israel: 1994–99.	Bivariate probit for choice of deductible and occurrence of at least one claim, plus other tests	Yes	Yes/No	No coverage/risk correlation found for beginning drivers, but those with more experience did purchase more coverage.
Saito (2006)	21,997 policies purchased from single insurer in April 1999–April 2000. Japan.	Bivariate probit for choice of deductible (or extended coverage) and occurrence of claim, plus other tests	Yes, apparently	No	Those with higher accident risks do not purchase more coverage. Results apply to both beginning and experienced drivers.
			Annuities		
Finkelstein & Poterba (2002)	Standardized life tables. UK: 1998	Comparison of mortality rates of annuitants and non-annuitants.	Essentially yes, since annuities are not individually priced	Yes	Annuitants live longer than non-annuitants with similar characteristics.
Finkelstein & Poterba (2004)	42,000 annuities sold by large UK insurer from 1981–98	Comparisons of risk (estimated via hazard model) with choice of annuity along several margins.	Yes	Yes	Annuitants with longer expected lives are more likely to purchase policies with greater "back-loading" and less likely to purchase survivor benefits. Less evidence of selection on initial payout.

Author	Data Description	Empirical Method	Complete Info?	Asymmetry Found?	Comments
McCarthy & Mitchell (forthcoming)	Standardized life tables, US, UK, Japan	Comparison of mortality rates of annuitants and non-annuitants	Essentially yes, since annuities not individually priced	Yes	Annuitants live longer than non-annuitants with similar characteristics.
			Life Insurance		
Cawley & Philipson (1999)	Several micro and aggregate datasets on pricing and insurance takeup. US: date unspecified	Various	No	No	Mortality is lower for those with life insurance; prices fall with quantity of insurance purchased; policyholders are not good at predicting their own demise.
McCarthy & Mitchell (forthcoming)	Aggregate life tables, US, UK, Japan	Comparison of mortality rates of insured/uninsured	No	No	Policyholders have lower mortality than non-insureds.
Hendel & Lizzeri (2003)	Pricing and lapsation rates for several large insurers	Comparison of NPV of contracts across different amounts of "front-loading"	Yes	No	Contracts with more prepayment are associated with lower future voluntary termination rates and lower total cost, which contradicts prediction that high-risk policyholders would pay up-front to avoid classification risk.
			Long-Term Care		
Finkelstein & McGarry (2006)	AHEAD survey of the elderly. US: 1995–2000	Comparison of individual's beliefs that they will enter nursing home with subsequent use, and with purchase of long-term care insurance	Yes	Yes, but no adverse selection	Individual's subjective estimates of likelihood of nursing-home use do predict subsequent use, after controlling for observables, and are correlated with insurance coverage. However, insurance coverage is *not* correlated with risk.
			Reverse Mortgages		
Davidoff & Welke (2005)	45,000 reverse mortgage holders and 3,000 others	Various, incl. calibrated simulation that also accounts for possible moral hazard	Unclear	Yes, but no adverse selection	Reverse mortgage holders have higher death rates than observationally equivalent non-holders and are more sensitive to changes in housing prices; the opposite of what adverse selection would predict
Cutler & Zeckhauser (1998)	Claims data from enrollees in different types of plans. US: 1990s	Comparison of characteristics of policyholders by choice of plan	Probably	Yes	Employees who chose more generous plans were older, spent more, and had more claims for high-cost procedures such as heart attacks and births. Changes in employer subsidy led to selection death spiral as best risks left most expensive plan in favor of cheaper alternatives.
Cutler & Reber (1998)	Claims and enrollment data from Harvard University. US: 1990s	Ditto	Probably	Yes	Ditto, largely same data as above. Welfare losses from adverse selection are about 2%–4% of total spending.

Author	Data Description	Empirical Method	Complete Info?	Asymmetry Found?	Comments
Cardon & Hendel (2001)	826 individuals from National Medical Expenditure Survey. US:	Two-stage structural model in which policyholders choose from available insurance policies and then make claims.	No	No	Most of variance in claims is explained by observables; private signal of health has low variance.
Fang, Keane & Silverman (2008)	Medicare Current Beneficiary Survey and Health and Retirement Survey. US: 2000–2002	OLS regressions of medical spending on dummy for presence of Medigap insurance plus extensive controls.	Yes	Yes, but no adverse selection	Those with Medigap insurance spend $4,000 less than those without; but conditional on observable health status, those with insurance spend $1,900 *more*. Medigap purchasers have *better* health than non-purchasers. Cognitive differences between purchasers and non-purchasers account for selection.
		Crop Insurance			
Just, et al (1999)	Several hundred corn and soybean farmers from national sample. US: 1988	Decomposition of insurance purchase incentives into subsidy, risk aversion and informational advantage components based on parametric model of crop yields.	Unclear	Yes, but selection effects importantly due to subsidy, not information al advantage	For both crops, risk aversion is a relatively unimportant motive for insurance purchases. Both insured and uninsured farmers would lose money on insurance purchases were it not for the federal subsidy.
Makki & Somwaru (2001)	60,000 corn farmers. Iowa (US): 1987–97.	Multinomial logit prediction of choice among various insurance coverage controlling for risk	Unclear	Yes	High-risk farmers are more likely to select broader revenue-based rather than yield-based insurance and to purchase more coverage.
Sil (2005)	Agric. Resource Management Survey data for corn and soybean growers. Indiana (US)	Logistic predictions of effect of having a forward contract on probability of loss, and on the choice of insurance coverage.	Probably yes	Yes	Policyholders with forward contracts have a 6% higher risk and choose more generous forms of insurance coverage.