Asymmetric Information in Insurance Markets: Empirical Assessments[*]

Pierre-André Chiappori[†] Bernard Salanié[‡]

January 24, 2012

Abstract

The paper surveys a number of approaches to testing for or evaluating the importance of asymmetric information in insurance relationships. Our focus throughout is on the methodology rather than on the empirical results.

Keywords: Insurance, adverse selection, moral hazard, contract theory, tests.

---

[*] To appear in the *Handbook of Insurance*, 2nd edition (G. Dionne, ed.).

[†] Department of Economics, Columbia University, 420 West 118th St., New York, NY 10027, USA. E-mail: pc2167@columbia.edu

[‡] Department of Economics, Columbia University, 420 West 118th St., New York, NY 10027, USA. E-mail: bsalanie@columbia.edu

Introduction

Modern insurance economics has been deeply influenced by the developments of contract theory. Our understanding of several crucial aspects, such as the design of optimal insurance contracts, the form of competition on insurance markets or the role of public regulation, just to name a few, systematically refers to the basic concepts of contract theory—moral hazard, adverse selection, commitment, renegotiation and others. Conversely, it is fair to say that insurance has been, and to a large extent still remains, one of the most important and promising fields of empirical application for contract theory.

It can even be argued that, by their very nature, insurance data provide nearly ideal material for testing the predictions of contract theory. Chiappori (1994) and Chiappori and Salanié (1997) remark that most predictions of contract theory are expressed in terms of a relationship between the form of the contract, a "performance" that characterizes the outcome of the relationship under consideration, and the resulting transfers between the parties. Under moral hazard, for instance, transfers will be positively correlated to but less volatile than outcomes, in order to conjugate incentives and risk sharing; under adverse selection, the informed party will typically be asked to choose a particular relationship between transfer and performance within a menu. The exact translation of the notions of "performance" and "transfer" varies with the particular field at stake. Depending on the particular context, the "performance" may be a production or profit level, the performance of a given task or the occurrence of an accident; whereas the transfer can take the form of a wage, a dividend, an insurance premium and others.

In all cases, empirical estimation of the underlying theoretical model would ideally require a precise recording of (i) the contract, (ii) the information available to both parties, (iii) the performance, and (iv) the transfers. In addition, the contracts should be to a large extent standardized, and large samples should be considered, so that the usual tools of econometric analysis can apply. As it turns out, data of this kind are quite scarce. In some contexts, the contract is essentially implicit, and its detailed features are not observed by the econometrician. More frequently, contracts do not present a standardized form because of the complexity of the information required either to characterize the various (and possibly numerous) states of the world that should be considered, or to precisely describe available information[1]. In many cases, part of the information at the parties' disposal is simply not observed by the econometrician, so that it is de facto impossible to condition on it as required by the theory. Last but not least, the "performance" is often not recorded, and even not precisely defined. In the case of labor contracts, for instance, the employee's "performance" often is the product of a supervisor's subjective evaluation, which is very rarely recorded in the data that the firm makes available to the econometrician.

In contrast, most insurance contracts fulfill all of the previous requirements. Individual insurance contracts (automobile, housing, health, life, etc.) are largely standardized. The insurer's information is accessible, and can generally be summarized through a reasonably small number of quantitative or qualitative indicators. The "performance"—whether it represents the occurrence of an accident, its cost, or some level of expenditure—is very precisely recorded in the firms' files. Finally, insurance companies frequently use data bases containing several millions of contracts, which is as close to asymptotic properties as one can probably be. It should thus be no surprise that empirical tests of adverse selection, moral hazard or repeated contract theory on insurance data have attracted renewed attention.

In what follows, we shall concentrate on empirical models that explicitly aim at testing for or evaluating the importance of asymmetric information in insurance relationships. This obviously excludes huge parts of the empirical literature on insurance, that are covered by other chapters of this volume. Some recent research has focused on evaluating the welfare consequences of asymmetric information, "beyond testing" to use the title of the survey by Einav, Finkelstein and Levin (2010). For lack of space we will not cover it here. Also, we will leave aside the important literature on fraud—a topic that is explicitly addressed by Picard in this volume. Similarly, since the major field of health insurance is comprehensively surveyed by Morrisey, we shall only allude to a few studies relating to information asymmetries in this context.

Finally, we chose to focus on the methodological aspects of the topic. In the past fifteen years, a large volume of empirical work has evaluated the importance of asymmetric information in various insurance markets. There are excellent surveys that present their results, such as Cohen-Siegelman

---

[1] This problem, for instance, is frequently encountered with data related to firms' behavior.

(2010); and we will limit ourselves to the broad conclusions we draw from these many studies.

The structure of this contribution is as follows. Section 1 discusses the main conclusions reached by the economic theory of insurance. We place particular emphasis on the testable consequences that can be derived from existing models. Section 2 reviews a few studies that exploit these theoretical insights in a static context. Section 3 briefly considers the dynamic aspects of the issue. We conclude with some ideas for future research.

1 Empirical tests of information asymmetries: the theoretical background

It is by now customary to outline two polar cases of asymmetric information, namely adverse selection and moral hazard. Each case exhibits specific features that must be understood before any attempt at quantifying their empirical importance[2].

1.1 Asymmetric information in insurance: what does theory predict?

1.1.1 Adverse selection

The basic story and its interpretations At a very general level, adverse selection arises when one party has a better information than other parties about some parameters that are relevant for the relationship. In most theoretical models of insurance under adverse selection, the subscriber is taken to have superior information. The presumption is usually that the insuree has better information than her insurer on her accident probability and/or on the (conditional) distribution of losses incurred in case of accident. A key feature is that, in such cases, the agent's informational advantage bears on a variable (risk) that directly impacts the insurer's expected costs. Agents who know that they face a higher level of risk will buy more coverage, introducing a correlation between the agents' contract choice and the unobservable component of their risk. The insurer's profit will suffer since the cost of

---

[2] We refer the reader to Salanié (2005) for a comprehensive presentation of the various theoretical models.

providing coverage is higher for higher-risk agents. In the terminology of contract theory, this is a model of *common values*; and this feature is what creates problems with competitive equilibrium.

This general definition may however be qualified in several ways. First, a finding that agents who buy more insurance have riskier outcomes is consistent with the standard story - they bought more insurance *because* they realized that they were more likely to have an accident - but also with alternative, observationally equivalent interpretations. To give but a simple example, assume that agents are of two types, green and red, and that these types are not observable by the insurer. Assume, furthermore, that red agents have two characteristics: their risk is larger *and* they have a higher predisposition to buy insurance (or contracts offering a more extensive coverage). These two characteristics could be linked by a causal relationship: agents want more coverage because they realize they are more accident-prone; or they could just be caused by some third factor—say, wealthier agents have a longer life expectancy and can also better afford to buy annuities. The distinction is irrelevant for most theoretical predictions, at least as long as the putative third factor is not observed by the insurance company[3]. As we shall discuss below, the important feature is the existence of an exogenous correlation between the agent's risk and her demand for insurance, not the source of this correlation.

Secondly, the focus in most theoretical models on one particular source of adverse selection—the agent's better knowledge of her risk—is very restrictive. In many real-life applications, risk is not the only possible source of informational asymmetry, and arguably not the most important one. Individuals also have a better knowledge of their own preferences, and particularly their level of risk aversion—an aspect that is often disregarded in theoretical models. A possible justification for this lack of interest is that if adverse selection only bears on preferences, it should have negligible consequences upon the form and the outcome of the relationship in competitive markets. Pure competition typically imposes that companies charge a fair premium, at least whenever the latter can be directly computed (which is precisely the case when the agent's risk is known.) The insurer's costs do not directly depend on the insuree's preferences: values are private. Then the equilibrium contract should not depend on whether the subscriber's pref-

---

[3] Indeed, the underwriting of standard annuity contracts is not contingent on the client's wealth or income.

erences are public or private information. To be a little more specific: in a model of frictionless, perfectly competitive insurance markets with symmetric information, the introduction of hidden information on preferences only will not alter the equilibrium outcome[4]

This conclusion should however be qualified, for at least two reasons. First, perfect competition does not approximate insurance markets that well. Fixed costs, product differentiation, price stickiness, switching costs and cross-subsidization are common; oligopoly is probably the rule rather than the exception. In such a context, firms are able to make positive profits; their profitability depends on the agents' demand elasticity, which depends on their risk aversion. Take the extreme case of a monopoly insurer, which corresponds to the principal-agent framework: it is well-known that adverse selection on risk aversion does matter for the form of the optimal contract, as more rent can be extracted from more risk-averse buyers.

A second caveat is that even when adverse selection on preferences *alone* does not matter, when added to asymmetric information of a more standard form it may considerably alter the properties of equilibria. In a standard Rothschild-Stiglitz context, for instance, heterogeneity in risk aversion may result in violations of the classical, "Spence-Mirrlees" single-crossing property of indifference curves, which in turn generates new types of competitive equilibria[5]. More generally, situations of bi- or multi-dimensional adverse selection are much more complex than the standard ones, and may require more sophisticated policies[6].

The previous remarks only illustrate a basic conclusion: when it comes to empirical testing, one should carefully check the robustness of the conclusions under consideration to various natural extensions of the theoretical background. Now, what are the main robust predictions emerging from the theoretical models?

---

[4] See Pouyet-Salanié-Salanié (2008) for a general proof that adverse selection does not change the set of competitive equilibria when values are private.

[5] See for instance Villeneuve (2003). The same remark applies to models with adverse selection and moral hazard, whether adverse selection is relative to risk, as in Chassagnon and Chiappori (1997), or to risk aversion, as in Julien, Salanié and Salanié (2007).

[6] Typically, they may require more instrument than in the standard models. In addition, one may have to introduce randomized contracts, and bunching may take specific forms. See Rochet and Stole (2003) for a survey.

The exclusivity issue A first and crucial distinction, at this stage, must be made between *exclusive* and *non exclusive* contracts. The issue, here, is whether the insurer can impose an exclusive relationship or individuals are free to buy an arbitrary number of contracts from different insurance companies. Both situations coexist in insurance markets; for instance, automobile insurance contracts are almost always exclusive, whereas annuities or life insurance contracts are typically sold without exclusivity.[7]

Non exclusive contracts and price competition Non exclusivity strongly restricts the set of possible contracts. For instance, no convex price schedule can be implemented: if unit prices rise with quantities (which is typically what adverse selection requires), agents can always "linearize" the schedule by buying a large number of small contracts from different insurers[8]. The same holds true for quantity constraints, which can be considered as a particular form of price convexity. To a large extent, the market, in the absence of exclusivity, entail standard (linear) pricing.

In this context, adverse selection has well-known consequences. Since all agents face the same (unit) price, high risk individuals are de facto subsidized (with respect to fair pricing), whereas low risk agents are taxed. The latter are likely to buy less insurance, or even to leave the market. A first prediction of the theory is precisely that, in the presence of adverse selection, the market typically shrinks, and the high risk agents are over-represented among buyers. In addition, purchased quantities should be positively correlated with risk; i.e., high risk agents should, everything equal, buy more insurance. Both predictions are testable using insurers' data.

The presence of adverse selection will also have an impact on prices. Because of the over-representation (in number and in quantity purchased) of high risk agents in the insurers' portfolios, unit prices will, at equilibrium, exceed the level that would obtain in the absence of adverse selection. Although the latter is not observable, it may in general be computed from the average characteristics of the general population. A typical example is provided by annuities, since the distribution of life expectancy conditional on age is well documented. It is in principle possible to compute the fair price

---

[7] A different but related issue is whether, in a non exclusive setting, each insurer is informed of the agent's relationships with other insurers.

[8] The benefits of linearization can be mitigated by the presence of fixed contracting costs. For large amounts of coverage, however, this limitation is likely to be negligible.

of a given annuity, and to compare it to actual market price. A difference that exceeds the "normal" loading can be considered as indirect evidence of adverse selection (provided, of course, that the normal level of loading can be precisely defined).

Exclusive contracts In the alternative situation of exclusivity, the set of available contracts is much larger. In particular, price schedules may be convex, and ceilings over insurance purchases can be imposed. Theoretical predictions regarding outcomes depend, among other things, on the particular definition of an equilibrium that is adopted—an issue on which it is fair to say that no general agreement has been reached. Using Rothschild and Stiglitz's concept, equilibrium may fail to exist, and cannot be pooling. However, an equilibrium à la Riley always exists. The same property holds for equilibria a la Wilson; in addition, the latter can be pooling or separating, depending on the parameters. Referring to more complex settings—for instance, game-theoretic frameworks with several stages—does not simplify the problem, because the properties of equilibria are extremely sensitive to the detailed structure of the game (for instance, the exact timing of the moves, the exact strategy spaces, ...), as emphasized by Hellwig (1987).

These remarks again suggest that empirically testing the predictions coming from the theory is a delicate exercise; it is important to select properties that can be expected to hold in very general settings. Here, the distinction between exclusive and non exclusive contracts is crucial. For instance, convex pricing—whereby the unit price of insurance increases with the purchased quantity—is a common prediction of most models involving exclusive contracts, but it cannot be expected to hold in a non exclusive framework.

A particularly important feature, emphasized in particular by Chiappori and Salanié (2000), is the so called *positive correlation property*, whereby a increasing relationship exists, conditional on all variables used for underwriting, between an agent's risk and the amount of insurance she purchases. Prior to any empirical test, however, it is crucial to clearly understand the scope and limits of this prediction; we analyze this issue in section 1.2.

1.1.2 Moral hazard

Moral hazard occurs when the probability of a claim is not exogenous, but depends on some decision made by the subscriber (e.g., effort of prevention). When the latter is observable and contractible, then the optimal decision will be an explicit part of the contractual agreement. For instance, an insurance contract covering a fire peril may impose some minimal level of firefighting capability, or alternatively adjust the rate to the existing devices. When, on the contrary, the decision is not observable, or not verifiable, then one has to examine the incentives the subscriber is facing. The curse of insurance contracts is that their mere existence tends to weaken incentives to reduce risk. Different contracts provide different incentives, hence result in different observed accident rates. This is the bottom line of most empirical tests of moral hazard.

Ex ante versus ex post An additional distinction that is specific to insurance economics is between an *accident* and a *claim*. The textbook definition of moral hazard is *ex ante*: the consequence of the agent's effort is a reduction in *accident* probability or severity, as one would expect of unobservable self-insurance or self-protection efforts. But insurance companies are interested in claims, not in accidents. Whether an accident results in a claim is at least in part the agent's decision, and as such it is influenced by the form of the insurance contract—a phenomenon usually called *ex post* moral hazard. Of course, the previous argument holds for both notions: more comprehensive coverage discourages accident prevention *and* increases incentives to file a claim for small accidents. However, the econometrician will in general be eager to separate "true" moral hazard, which results in changes in the accident rates, from ex post moral hazard. Their welfare implications are indeed very different. For instance, a deductible is more likely to be welfare increasing when it reduces accident probability than when its only effect is to discourage victims from filing a claim. The latter only results in a transfer from insurer to insuree; and this matters much less for welfare[9].

The distinction between claims and accidents has two consequences. One is that the incentives to file a claim should be (and indeed are) monitored by the insurance company, particularly when the processing of a small claim involves important fixed costs for the company. A deductible, for instance, is often seen by insurance companies as a simple and efficient way of avoiding small claims. Secondly, and from a more empirical perspective, the empirical

---

[9] A related problem is fraud, defined as any situation where a subscriber files a claim for a false accident or overstates its severity in order to obtain a more generous compensation. The optimal contract, in that case, typically requires selective auditing procedures (see the chapter by Picard in this volume).

distribution of claims will in general be a truncation of that of accidents—since "small" accidents are typically not declared. However, the truncation is endogenous; it depends on the contract (typically, on the size of the deductible or the presence of experience rating), and also on the individual characteristics of the insured (if only because the cost of higher future premia is related to the expected frequency of future accidents). This can potentially generate severe biases. If a high deductible discourages small claims, a (spurious) correlation will appear between the choice of the contract and the number of filed claims, even in the absence of adverse selection or ex ante moral hazard. The obvious conclusion is that any empirical estimation must very carefully control for potential biases due to the distinction between accidents and claims.

Moral hazard and adverse selection Quite interestingly, moral hazard and adverse selection have similar empirical implications, but with an inverted causality. Under adverse selection, people are characterized by different levels of ex ante risk, which translate into different ex post risk (accident rates); therefore different insurees choose different contracts. In a context of moral hazard, agents first choose different contracts; then, faced with different incentive schemes they adopt more or less cautious behavior, which ultimately results in heterogeneous accident probabilities. In both cases however, controlling for observables, the choice of a contract will be correlated with the accident probability: more comprehensive coverage is associated with higher risk.

This suggests that it may be difficult to distinguish between adverse selection and moral hazard in the static framework (i.e., using cross-sectional data). An econometrician may find out that, conditionally on observables, agents covered by a comprehensive automobile insurance contract are more likely to have an accident. But this could be because the comprehensive contract they chose (for some exogenous and independent reason) reduced their incentives to drive safely; or because they chose full coverage knowing that their risk was higher; or because both contract choice and risk were determined by some exogenous, third factor. Discriminating between these explanations is a much harder task, to which we return in section 2.4.

1.2 The positive correlation property: general results

The argument that in the presence of asymmetric information and with exclusive contracts, ex post risk and coverage should be positively correlated is quite intuitive; and in fact such tests were used in the health insurance literature[10] before they were formally analyzed by Chiappori and Salanié (1997, 2000) and by Chiappori et al (2006). Yet the very idea of a positive correlation begs the question of what are the appropriate measures of the variables that are to be correlated when contracts can be complex, and also of the conditioning set—the latter is crucial as we will see. Often-used pricing schemes, such as experience rating, or regulation may also complicate the picture, not to mention ex post moral hazard (when insurees may not file low-value claims so as to preserve their risk rating.)

In addition, there are a priori appealing objections to the intuitive argument. The one that comes up most frequently may be that insurers attempt to "cherry-pick" insurees: more risk-averse insurees may both buy higher coverage and behave more cautiously, generating fewer claims. Such a "propitious" or "advantageous selection"[11] suggests that the correlation of risk and coverage may in fact be negative. As it turns out, this counter-argument is much less convincing than it seems; but it does require proper analysis.

Before we proceed with the formal analysis, it is important to note that all of the arguments below assume an observably homogeneous population of insurees: more precisely, we focus on a subpopulation whose pricing-relevant characteristics are identical. What is "pricing-relevant" depends on what insurers can observe, but also on regulation (e.g., rules that forbid discrimination.) We will assume that the econometrician also observes all pricing-relevant characteristics, which is typically true if he has access to the insurer's data.

A simple counterexample We may start with a simple but basic remark—namely that the positive correlation property is, broadly speaking, typical of a *competitive* environment. While we will be more precise below, it is easy to see that in a monopoly context, the correlation between coverage and accidents may take any sign, at least when the analyst cannot fully control for risk aversion. An intuitive argument goes as follows. Start from a monopoly situation in which agents have the same risk but different risk aversions; to

---

[10] See for instance the surveys by Cutler and Zeckhauser (2000) and Glied (2000).

[11] See Hemenway (1990), and de Meza and Webb (2001) for a recent analysis.

keep things simple, assume there exist two types of agents, some very risk-averse, the others much less. The monopoly outcome is easy to characterize. Two contracts are offered; one, with full insurance and a larger unit premium, will attract more risk-averse agents, while the other entails a smaller premium but partial coverage and targets the less risk-averse ones. Now, slightly modify individual risks, in a way that is perfectly correlated with risk aversion. By continuity, the features just described will remain valid, whether more risk-averse agents become slightly riskier or slightly less risky. In the first case, we are back to the positive correlation situation; in the second case, we reach an opposite, *negative* correlation conclusion.

The logic underlying this example is clear. When agents differ in several characteristics—say, risk and risk aversion—contract choices reflect not only relative riskiness but also these alternative characteristics. The structure of the equilibrium may well be mostly driven by the latter (risk aversion in the example above), leading to arbitrary correlations with risk[12].

However, and somewhat surprisingly, this intuition does *not* hold in a competitive context. Unlike other differences, riskiness *directly* impacts the insurer's profit; under competition, this fact implies that the correlation can only be positive (or zero), but never negative, provided that it is calculated in an adequate way. To see why, let us first come back to our simple example, this time in a perfectly competitive context. Again, we start from the benchmark of agents with identical risk but different risk aversions, and we marginally modify this benchmark by slightly altering riskiness. If more risk averse agents are riskier, any Rothschild-Stiglitz (from now RS) equilibrium will take the usual form: namely, a full insurance contract attracting high risk/high risk aversion agents, and a partial coverage one targeting the remaining, low risk/low risk aversion ones; in particular, the positive correlation property is satisfied. Assume, now, that risk averse agents are *less* risky. Then *the RS equilibrium contract can only be pooling.* Indeed, assume that a separating contract exists. Under the standard, zero-profit condition, the comprehensive coverage contract, which attracts the more risk averse agents, must now have a *lower* unit price, because of their lower risk. But then the revelation constraint cannot be satisfied, since all agents - including the less risk averse ones - prefer more coverage *and* a cheaper price.

Of course, this example cannot by itself be fully conclusive. For one

---

[12] The analysis in Jullien-Salanié-Salanié (2007) also illustrates it, with risk-averse agents and moral hazard in a Principal-Agent model of adverse selection.

thing, it assumes perfect correlation between risk and risk aversion; in real life, much more complex patterns may exist. Also, we are disregarding moral hazard, which could in principle reverse our conclusions. Finally, RS is not the only equilibrium concept; whether our example would survive a change in equilibrium concept is not clear a priori. We now proceed to show that, in fact, the conclusion is surprisingly robust. For that purpose, we turn to the formal arguments in Chiappori et al (2006).

The formal model Consider a competitive context in which several contracts coexist. Suppose that each contract C_i offers a coverage $R_i(L)$: if the total size of the claims over a contract period is L, then the insuree will be reimbursed $R_i(L)$. For instance, $R_i(L) = \max(L - d_i, 0)$ for a straight deductible contract with deductible d_i. We say that contract C_2 "covers more" than contract C_1 if $R_2 - R_1$ (which is zero in $L = 0$) is an increasing function of L; this is a natural generalization of $d_2 \leq d_1$ for straight deductible contracts. To keep our framework fully general, we allow the probability distribution of losses to be chosen by the insuree in some set, which may be a singleton (then risk is fully exogenous) or not (as in a moral hazard context).

Now consider an insuree who chose a contract C_1, when a contract C_2 with more coverage was also available to him. Intuitively, it must be that contract C_1 had a more attractive premium. Let P_1 and P_2 denote the premia of C_1 and C_2. Now suppose that the insuree anticipates that under contract C_1, he will generate a distribution of claims G. Note that the insuree could always buy contract C_2 and otherwise behave as he does under contract C_1, generating the same distribution G of claims L that he anticipates under contract C_1. If the insuree is risk-neutral, his expected utility under contract C_1 is

$$\int R_1(L)dG(L) - P_1$$

and he knows that he could obtain

$$\int R_2(L)dG(L) - P_2$$

by buying contract C_2 and otherwise behaving as he does under contract C_1, generating the same distribution G of claims L that he anticipates under contract C_1. By revealed preferences, it must be that

$$\int R_1(L)dG(L) - P_1 \geq \int R_2(L)dG(L) - P_2.$$

Now let the insuree be risk-averse, in the very weak sense that he is averse to mean-preserving spreads. Then given that contract C_2 covers more than contract C_1, it is easy to see that $R_1 - \int R_1 dG$ is a mean-preserving spread of $R_2 - \int R_2 dG$, which makes the inequality even stronger. To summarize this step of the argument:

Lemma 1 *Assume that an insuree prefers a contract C_1 to a contract C_2 that covers more than C_1. Let G be the distribution of claims as anticipated by the agent under C_1. Then if the insuree has increasing preferences over wealth and dislikes mean-preserving spreads,*

$$P_2 - P_1 \geq \int (R_2(L) - R_1(L))\, dG(L).$$

The main result We now consider the properties of the equilibrium. As mentioned above, under adverse selection, the mere definition of an equilibrium is not totally clear. For instance, a Rothschild-Stiglitz equilibrium requires that each contract makes non negative profit, and no new contract could be introduced and make a positive profit. As is well known, such an equilibrium may fail to exist or to be (second best) efficient. Alternatively, equilibria a la Spence-Miyazaki allow for cross subsidies (insurance companies may lose on the full insurance contract and gain on the partial insurance ones); as a result, an equilibrium always exists and is efficient. We certainly want to avoid taking a stand on which notion should be preferred; and we do not even want to rule out imperfectly competitive equilibria.

Therefore, we shall simply make one assumption on the equilibrium—namely, that *the profit made on contracts entailing more comprehensive coverage cannot strictly exceed those made on contract involving partial insurance.* Note that this “non-increasing profits” property (Chiappori et al. 2006) is satisfied by the two concepts just described—and, as a matter of fact, by most (if not all) concepts of competitive equilibrium under adverse selection proposed so far.

Define the profit of the insurer on a contract C as the premium[13] minus the reimbursement, allowing for a proportional cost λ and a fixed cost K:

$$\pi = P - (1 + \lambda) \int R(L) dF(L) - K$$

---

[13] Premia are often taxed, but this is easy to incorporate in the analysis.

if the average buyer of contract C generates a distribution of claims F. Under the non-increasing profits assumption, we have that $\pi_2 \geq \pi_1$, therefore:

$$P_2 - P_1 \leq (1+\lambda_2)\int R_2(L)dF_2(L) - (1+\lambda_1)\int R_1(L)dF_1(L) + K_2 - K_1,$$

which gives us a bound on $P_2 - P_1$ in the other direction than in Lemma 1. Remember that the inequality in the lemma contains the distribution of claims G that the insuree expects to prevail under contract C_1. Assume that there is at least one insuree who is not optimistic[14], in the sense that his expectations G satisfy

$$\int (R_2(L) - R_1(L))\,(dG(L) - dF_1(L)) \geq 0.$$

Combining with Lemma 1 and rearranging terms to eliminate $P_2 - P_1$, we obtain

$$\int R_2(L)\,((1+\lambda_2)dF_2(L) - dF_1(L)) \geq \lambda_1 \int R_1(L)dF_1(L) + K_1 - K_2. \quad (1)$$

While it may not be obvious from this expression, this inequality is the positive correlation property. To see this, assume that the proportional costs are zero, and that $K_2 \leq K_1$. Then we have

Theorem 2 *Consider a contract C_1 and a contract C_2 that covers more than C_1. Assume that*

1. *at least one of the insurees who prefers C_1 to C_2 is not optimistic, has increasing preferences, and is averse to mean-preserving spreads*
2. *profits are non-increasing in coverage.*

Then if C_1 and C_2 have zero proportional costs and their fixed costs are ordered by $K_2 \leq K_1$,

$$\int R_2(L)\,(dF_2(L) - dF_1(L)) \geq 0.$$

---

[14] Since $R_2 - R_1$ is non-decreasing, this inequality holds for instance if G first-order stochastically dominates F_1, hence our choice of the term "not optimistic." Chiappori et al (2006) assumed that no insuree was optimistic. The much weaker condition stated here is in fact sufficient.

Take the simplest possible case, in which claims can either be 0 or $\bar{L}$, and the average buyer of contract C_i faces a claim $\bar{L}$ with probability p_i; then

$$\int R_2(L)\left(dF_2(L) - dF_1(L)\right) = \left(p_2 - p_1\right) R_2\left(\bar{L}\right),$$

and Theorem 2 implies that $p_2 \geq p_1$: contracts with more coverage have higher ex post risk, in the sense used in the earlier literature. In more complex settings, inequality (1) could be used directly if the econometrician observes the reimbursement schemes R_i and distributions of claim sizes F_i, and has reliable estimates of contract costs λ_i and K_i.

Note that while we assume some weak forms of risk aversion and rationality in Assumption 1, we have introduced no assumption at all on the correlation of risk and risk aversion: it does not matter whether more risk-averse agents are more or less risky, insofar as it does not invalidate our assumption that profits do not increase in coverage and if we apply the general version of the inequality (1). Take the advantageous selection story in de Meza and Webb (2001) for instance. They assume no proportional costs, zero profits (which of course fits our Assumption 2), and a $\{0, \bar{L}\}$ model of claim sizes; but they allow for administrative fixed costs, so that (1) becomes

$$(p_2 - p_1) R_2(\bar{L}) \geq K_1 - K_2.$$

In the equilibrium they consider, contract C_1 is no-insurance, which by definition has zero administrative costs. Thus their result that p_2 may be lower than p_1 fits within Theorem 2.

Let us stress again that in imperfectly competitive markets Assumption 2 may fail when agents have different risk aversions, and sometimes negative correlations will obtain; but it should be possible to check Assumption 2 directly on the data[15].

Inequality (1) is also useful as an organizing framework to understand when simpler versions like $p_2 \geq p_1$ may *not* hold. Assume again that claims can only be 0 or some (now contract-dependent) $\bar{L}_i$, and reintroduce costs. Then (1) is

$$(1 + \lambda_2) p_2 R_2(\bar{L}_2) - (R_2(\bar{L}_1) + \lambda_1 R_1(\bar{L}_1)) p_1 + K_2 - K_1 \geq 0.$$

Proportional costs, even if equal across contracts, may make this consistent with $p_2 < p_1$. Even if the proportional costs are zero, $p_2 \geq p_1$ may fail if

---

[15] Chiappori et al (2006) do it in their empirical application.

$\bar{L}_2 > \bar{L}_1$, so that higher coverage generates larger claims; or as we saw above, if $K_2 > K_1$—higher coverage entails larger fixed costs. We would argue that in such cases, the positive correlation property does not fail: it simplest form is just applied incorrectly.

Finally, Theorem 2 abstracts from experience rating. With experience rating the cost of a claim to the insuree is not only $(L - R_i(L))$: it also includes the expected increase in future premia, along with their consequences on future behavior. If switching to a new contract is costless (admittedly a strong assumption in view of the evidence collected by Handel (2011) and others), then the discounted cost of a claim $c(L)$ is the same for both contracts. It is easy to see that experience rating then does not overturn the inequality $p_2 \geq p_1$ in the simpler cases. Chiappori et al (2006) have a more detailed discussion.

As a final remark, note that, as argued in section 1.1.2, tests of the positive correlation property, at least in the static version, cannot distinguish between moral hazard and adverse selection: both phenomena generate a positive correlation, albeit with opposite causal interpretations. Still, such a distinction is quite important, if only because the different (and sometimes opposite) welfare consequences. For instance, a deductible—or, for that matter, any limitation in coverage—that reduces accident probabilities through its impact on incentives may well be welfare increasing; but if the same limitation is used as a separating device, the conclusion is less clear. Distinguishing empirically between moral hazard and adverse selection requires more structure or more data; we survey several approaches in section 2.4.

2 Empirical tests of asymmetric information

While the theoretical analysis of contracts under asymmetric information began in the 70s, the empirical estimation of insurance models entailing either adverse selection or moral hazard is more recent[16]. Much of this work revolves around the positive correlation property, as will our discussion.

We will focus here on the methodological aspects. We start with insurance markets involving non exclusive contracts. Next, we discuss the most

---

[16] Among early contributions, one may mention Boyer and Dionne (1987) and Dahlby (1983).

common specifications used to evaluate and test the correlation of risk and coverage under exclusivity. Then we give a brief discussion of the results; the survey by Cohen and Siegelman (2010) provides a very thorough review of empirical studies on various markets, and we refer the reader to it for more detail. Finally, we discuss the various approaches that have been used to try to disentangle moral hazard and adverse selection.

2.1 Non-exclusive insurance

2.1.1 Annuities

Annuities provide a typical example of non exclusive contracts, in hich moreover the information used by the insurance company is rather sparse. Despite the similarities between annuities and life insurances (in both cases, the underlying risk is related to mortality), it is striking to remark that while the underwriting of life insurance contracts (at least above some minimum amount) typically require detailed information upon the subscriber's health state, the price of an annuity only depends on the buyer's age and gender. One may expect that this parsimony leaves a lot of room for adverse selection; empirical research largely confirms this intuition.

A first line of research has focused on prices. In an important contribution, Friedman and Warshawski (1990) compute the difference between the implicit contingent yield on annuities and the available yield on alternative forms of wealth holding (in that case, US government bonds). Even when using longevity data compiled from company files, they find the yield of annuities to be about 3% lower than US bonds of comparable maturity, which they interpret as evidence of adverse selection in the company's portfolio. Similar calculations on UK data by Brugiavini (1990) find a 3% difference, but only when longevity is estimated on the general population.

A related but more direct approach studies the distribution of mortality rates in the subpopulation of subscribers, and compares it to available data on the total population in the country under consideration. Brugiavini (1990) documents the differences in life expectancy between the general population and the subpopulation actually purchasing annuities. For instance, the probability, at age 55, to survive till age 80 is 25% in the general population but close to 40% among subscribers. In a similar way, the yield difference computed by Friedman and Warshawski (1990) is 2% larger when computed from data relative to the general population.

The most convincing evidence of adverse selection on the annuity market is probably that provided by Finkelstein and Poterba (2004). They use a data base from a UK annuity firm; the data covers both a compulsory market (representing tax-deferred retirement saving accounts that must be transformed into annuities to preserve the tax exemption) and a voluntary market. The key element of their empirical strategy is the existence of different products, involving different degrees of back-loading. At one extreme, nominal annuities pay a constant nominal amount; the real value of annual payments therefore declines with inflation. Alternatively, agents may opt for escalating annuities, in which a initially lower annual payment rises with time at a predetermined rate (in practice, 3 to 8%), or for real annuities, which pay an annual amount indexed on inflation. Under adverse selection on mortality risks, one would expect agents with superior life expectancy to adopt more back-loaded products (escalating or real). Finkelstein and Poterba's results confirm this intuition; using a proportional hazard model they find that buyers of these products have a significantly smaller death hazard rate. The most striking conclusion is the magnitude of this effect. The differential impact, on the hazard rate, of indexed or escalating products versus nominal ones dominates that of gender, the standard indicator used in underwriting; for the voluntary market, the impact of contract choice is actually several times larger.

These results teach two lessons. First, adverse selection does exist in real life, and particularly affects markets in which insurers collect little information during the underwriting process is scarce—a salient characteristic of annuity markets. Second, the form taken by adverse selection on such markets goes beyond the standard correlation between risk and quantity purchased; the type of product demanded is also affected by the agent's private information, and that effect may in some cases be dominant.[17]

2.1.2 Life insurance

Life insurance contracts provide another typical example of non exclusive contracts, although adverse selection might in this case be less prevailing. In an early paper, Cawley and Philipson (1999) used direct evidence on the (self-perceived and actual) mortality risk of individuals, as well as the price

---

[17] While Finkelstein and Poterba do find a significant relationship between risk and quantity purchased, the sign of the correlation, quite interestingly, varies between the compulsory and voluntary markets.

and quantity of their life insurance. They found that unit prices fall with quantities, indicative of the presence of bulk discounts. More surprisingly, quantities purchased appeared to be *negatively* correlated with risk, even when controlling for wealth. They argued that this indicated that the market for life insurance may not be affected by adverse selection. This conclusion is however challenged in a recent paper by He (2009), who points to a serious sample selection problem in the Cawley-Philipson approach: agents with a higher, initial mortality risk are more likely to have died *before* the beginning of the observation window, in which case they are not included in the observed sample.

To avoid this bias, He suggests to concentrate on a sample of potential *new* buyers (as opposed to the entire cross section). Using the Health and Retirement Study dataset, he does find evidence for the presence of asymmetric information, taking the form of a significant and positive correlation between the decision to purchase life insurance and subsequent mortality (conditional on risk classification). The effect is actually quite strong; for instance, individuals who died within a 12-year time window after a base year were 19% more likely to have taken up life insurance in that base year than were those who survived the time window.

In summary, the existence of adverse selection effects is well documented in several non-exclusive markets.

2.2 Evaluating the correlation of risk and coverage in exclusive markets

We now turn to exclusive markets. To measure the correlation of risk and coverage, we of course need to measure them first. Since risk here means "ex-post risk", it can be proxied by realized risk: the occurrence of a claim (a binary variable), the number of claims (an integer), or the cumulative value of claims (a non-negative number) could all be used, depending on the specific application. Let y_i denote the chosen measure of ex-post risk for insuree i. Coverage D_i could be the value of the deductible, or another variable (say a binary, compulsory vs complementary variable) could be used.

Finally, a (hopefully complete) set of pricing-relevant variables X_i will be found in the insurer's files. As emphasized by Chiappori and Salanié (1997, 2000), it is very important to account for all publicly observed pricing-relevant covariates. Failure to do so can lead to very misleading results:

Dionne et al. (2001) provide a striking illustration on the early study by Puelz and Snow (1994). Even so, it is not always obvious which variables are "pricing-relevant"; we will return to this issue in section 2.3.

2.2.1 Basic approaches

Let us focus first the simplest (and very common) case in which both y and D are 0-1 variables. Then one straightforward measure of the relevant correlation[18] is

$$\rho_1(X) = \Pr(y = 1|D = 1, X) - \Pr(y = 1|D = 0, X); \tag{2}$$

and a second one is

$$\rho_2(X) = \text{cov}(y, D|X) = \Pr(y = D = 1|X) - \Pr(y = 1|x)\Pr(D = 1|X). \tag{3}$$

It is easy to see that

$$\rho_2(X) = \rho_1(X)V(D|X)$$

since $V(D|X) = \Pr(D = 1|X)(1 - \Pr(D = 1|X))$; it follows that the two measures have the same sign.

In Chiappori and Salanié (1997, 2000) we proposed to simultaneously estimate two binary choice models. One relates to the choice of coverage:

$$y = \mathbb{1}(f(X) + \varepsilon > 0); \tag{4}$$

and the second one regresses coverage on covariates:

$$D = \mathbb{1}(g(X) + \eta > 0). \tag{5}$$

We argued that the correlation can be estimated by running a bivariate probit for (y, D) and allowing for correlated ε and η, or by estimating two separate probits and then measuring the correlation of the generalized residuals $\hat{\varepsilon}$ and $\hat{\eta}$. By construction the probit assumes that ε and η are independent of X, so that a test that they are uncorrelated is equivalent to a test that ρ_1 and ρ_2 are identically zero.

According to standard theory, asymmetric information should result in a positive correlation, under the convention that $D = 1$ (resp. $y = 1$) corresponds to more comprehensive coverage (resp. the occurrence of an accident).

---

[18] Recall however from section 1.2 that given the results of Chiappori et al (2006), the positive correlation property may bear on a more complicated object.

One obvious advantage of this setting is that is does not require the estimation of the pricing policy followed by the firm, which is an extremely difficult task—and a potential source of severe bias.

An alternative way to proceed when D is a 0-1 variable is to run a linear regression of y on D:

$$E(y|X, D) = a(X) + b(X)D + u. \tag{6}$$

Given that D only takes the two values 0 and 1, the linear form is not restrictive; and it easy to see that the estimator of $b(X)$ in (6) converges to

$$\rho(X) = E(y|D = 1, X) - E(y|D = 0, X),$$

which equals $\rho_1(X)$ if y is also a 0-1 variable; if it is not, then $\rho(X)$ is a useful measure of correlation but may not be the appropriate one.

Given the often large set of covariates X used by insurers for pricing, it may be hard to find the correct functional forms for f and g, or alternatively for a and b. We also proposed a nonparametric test that relies upon the construction of a large number of "cells", each cell being defined by a particular profile of exogenous variables. Under the null (in the absence of asymmetric information), within each cell the choice of contract and the occurrence of an accident should be independent, which can easily be checked using a χ^2 test. Constructing the cells requires some prior knowledge of the context; and it is useful to restrict the analysis to relatively homogeneous classes of drivers.

Finally, while much of the literature has focused on a discrete outcome (the occurrence of a claim, or sometimes a coarse classification), we have shown in section 1.2 that there is no need to do so. For a more general implementation of the test, we refer the reader to Chiappori et al (2006, section 5) who use data on the size of claims to test the more general positive correlation property of Theorem 2.

2.2.2 Accidents versus claims

These methods can easily be generalized to the case when coverage D takes more than two values; see for instance Dionne, Gouriéroux and Vanasse (1997), Richaudeau (1999) and Gouriéroux (1999) for early contributions. However, the issues raised in section 1.2 then may become important. If for instance D is the choice of deductibles, then we need to take into account differences in per-contract and per-claim costs for the insurer. A regression using claims as the dependent variable may generate misleading results, because a larger deductible automatically discourages reporting small accidents, hence reduces the number of claims even when the accident rate remains constant.

As shown in the Appendix of Chiappori et al. (2006), if insurees follow simple, contract-independent strategies when deciding to report a loss as a claim, then under fairly weak assumptions Theorem 2 is still valid. However, the positive correlation test then becomes conservative: positive correlations can be found even without asymmetric information. In any case, we know very little about the reporting behavior of insurees and other approaches are still useful. Chiappori and Salanié (2000) discarded all accidents where one vehicle only is involved. Whenever two automobiles are involved, a claim is much more likely to be filed in any case[19]. A more restrictive version is to exclusively consider accidents involving bodily injuries, since reporting is mandatory in that case; but this implies a drastic reduction in the number of accidents in the data.

Alternatively, one can explicitly model the filing decision as part of the accident process. For any accident, the agent computes the net benefit of filing a claim, and reports the accident only when this benefit is positive (or above some threshold). Although accidents involving no claims are generally not observed[20], adequate econometric techniques can be used. Note, however, that these require estimating a complete structural model.

2.3 Is the correlation positive?

The existence of a positive correlation between risk and coverage (appropriately measured) cannot be interpreted as establishing the presence of asymmetric information without some precautions: any misspecification can indeed lead to a spurious correlation. Parametric approaches, in particular, are highly vulnerable to this type of flaws, especially when they rely upon some simple, linear form. But the argument is not symmetric. Suppose, indeed,

---

[19] In principle, the two drivers may agree on some bilateral transfer and thus avoid the penalties arising from experience rating. Such a "street-settled" deal is however quite difficult to implement between agents who meet randomly, will probably never meet again, and cannot commit in any legally enforceable way (since declaration is in general compulsory according to insurance contracts). We follow the general opinion in the profession that such bilateral agreements can be neglected.

[20] Some data sets do, however, record accidents that did not result in claims. Usually, such data sets have been collected independently of insurance companies. See Richaudeau (1999).

that some empirical study does *not* reject the null hypothesis of zero correlation. In principle, this result might also be due to a misspecification bias; but this explanation is not very credible as it would require that while (fully conditional) residuals are actually positively correlated, the bias goes in the opposite direction with the *same* magnitude—so that it exactly offsets the correlation. In other words, misspecifications are much more likely to bias the results *in favor* of a finding of asymmetric information.

Moreover, a positive correlation may come from variables that are observed by insurers but not used in pricing. There are many instances of such "unused observables": regulation may forbid price-discrimination based on some easily observed characteristics such as race, or insurers may voluntarily forgo using some variables for pricing. Finkelstein and Poterba (2006) show for instance that British insurers do not use residential address in pricing annuities, even though it is clearly informative on mortality risk. The theoretical arguments that led us to the positive correlation property of Theorem 2 extend to such cases, as long as the list of "pricing-relevant" variables excludes unused unobservables. A positive correlation then may be entirely due to these unused observables; and since they are public information it would be abusing terms to conclude to the presence of asymmetric information.

Given these remarks, it may come as a surprise that the estimated correlation is often close to zero. The case of automobile insurance is emblematic. Using three different empirical approaches, Chiappori and Salanié (1997, 2000) could not find evidence of a nonzero correlation; and most later work has confirmed their findings. A few studies of automobile insurance have estimated a positive correlation; but it was often due to special features of a local market. As an example, Cohen (2005) found that Israeli drivers who learned that they were bad risks tended to change insurers and buy underpriced coverage, an opportunistic behavior that was facilitated at the time by local regulations concerning information on past driver records.

The evidence on health insurance is more mixed, with some papers finding positive correlation, zero correlation, or negative correlation. The Medigap insurance market[21] is especially interesting since Fang, Keane and Silverman (2008) found robust evidence that risk and coverage are negatively correlated. They show that individuals with higher cognitive ability are both more likely

---

[21] Medigap insurance is private, supplementary insurance targeted at Medicare recipients in the US.

to purchase Medigap and have lower expected claims.

This points towards the fact that asymmetric information may bear on several dimensions—not only risk. As we explained in section 1.2, with perfectly competitive markets the positive correlation property should hold irrespective of the dimensionality of privately known characteristics. This is often misunderstood. For instance, Cutler, Finkelstein and McGarry (2008) argue that much of the variation in test results across markets can be explained by the role of heterogeneous risk-aversion; but variations in the market power of insurers are also necessary, and can be evaluated using the variation of profits with coverage. In Chiappori et al (2006), we used this approach and we found clear deviations from perfect competition.

The findings by Fang et al (2008) stress the importance of taking into account the cognitive limitations of insurees; we return to this in the conclusion.

2.4 Adverse selection versus moral hazard: the static context

As argued above, the previous tests are not specific of adverse selection. Moral hazard would typically lead to the same kind of correlation, although with a different causality. In order to distinguish between adverse selection and moral hazard, one need some additional structure.

In some cases, one explanation may seem more plausible. For instance, it has often been argued that asymmetric information in annuity contracts was mostly due to selection: individuals are unlikely to die younger *because* of a lower annuity payment. Sometimes the data contain variables that can be used to directly assess adverse selection. For instance, Finkelstein and McGarry (2006), studying the long-term care insurance market, use individual-level survey data from the Asset and Health Dynamics (AHEAD) cohort of the Health and Retirement Study (HRS). A crucial feature of this data is that it provides a measure of individual beliefs about future nursing home use—an information to which insurers have obviously no access. They find that these self-assessed risk estimates are indeed informative of actual, subsequent nursing home utilization, and also of the person's long-term care insurance holdings—a clear indicator of adverse selection. In addition, they can then analyze the determinants of the demand for long term care insurance; they conclude that these determinants are typically multi-dimensional.

Other papers have relied either on natural or quasi-natural experiments, or on the fact that moral hazard and adverse selection generate different predictions for the dynamics of contracts and claims. We discuss here the natural experiments approach, reserving dynamics for section 3.

2.4.1 Natural experiments

Assume that, for some exogenous reason (say, a change in regulation), a given, exogenously selected set of agents experiences a modification in the incentive structure they are facing. Then the changes in the incentives that agents are facing can reasonably be assumed exogenous in the statistical sense (i.e., uncorrelated with unobserved heterogeneity.) The resulting changes in their behavior can be directly studied; and adverse selection is no longer a problem, since it is possible to concentrate upon agents that remained insured throughout the process[22].

The first, and arguably most influential study of moral hazard is the celebrated Rand study on medical expenditures (Manning et al. 1987), in which individuals were randomly allocated to different coverage schemes. While such examples, involving explicit randomization, are actually quite rare (if only because of their cost), the basic idea may in some occasions apply even in the absence of an actual experiment of this kind. Any context where similar individuals are facing different incentive schemes can do, provided one can be sure that the selection into the various schemes is not related to risk-relevant characteristics. Clearly, the key issue in this literature is the validity of this exogeneity assumption.

A typical example is provided by the changes in automobile insurance regulation in Québec, where a “no fault” system was introduced in 1978, then deeply modified in 1992. Dionne and Vanasse (1996) provide a careful investigation of the effects of these changes. They show that the new system introduced strong incentives to increase prevention, and that the average accident frequency dropped significantly during the years that followed its introduction. Given both the magnitude of the drop in accident rate and the absence of other major changes that could account for it during the period under consideration, they conclude that the reduction is claims is indeed due to the change in incentives [23].

---

[22] In addition, analyzing the resulting attrition (if any) may in some cases convey interesting information on selection issues.

[23] See Browne and Puelz (1998) for a similar study on US data.

An ideal experiment would also have a randomly assigned control group that is not affected by the change, allowing for a differences-of-differences approach. A paper by Dionne and St-Michel (1991) provides a good illustration of this idea. They study the impact of a regulatory variation of coinsurance level in the Quebec public insurance plan on the demand for days of compensation. Now it is much easier for a physician to detect a fracture than, say, lower back pain. If moral hazard is more prevalent when the information asymmetry is larger, theory predicts that the regulatory change will have more significant effects on the number of days of compensation for those cases where the diagnosis is more problematic. This prediction is clearly confirmed by empirical evidence. Note that the effect thus identified is ex post moral hazard. The reform is unlikely to have triggered significant changes in prevention; and, in any case, such changes would have affected all types of accidents.

Additional evidence is provided by Fortin et al. (1994), who examine how the Canadian Workers' Compensation (WC) and the Unemployment Insurance (UI) programs interact to influence the duration of workplace accidents. Here, the duration is estimated from a mixed proportional hazard model, where the baseline hazard is estimated nonparametrically, and unobserved heterogeneity is taken into account using a gamma distribution. They show that an increase in the generosity of Workers' Compensation in Quebec leads to an increase in the duration of accidents. In addition, a reduction in the generosity of Unemployment Insurance is, as in Dionne and St-Michel, associated with an increase in the duration of accidents that are difficult to diagnose. The underlying intuition is that workers' compensation can be used as a substitute to unemployment insurance. When a worker goes back to the labor market, he may be unemployed and entitled to UI payments for a certain period. Whenever workers' compensation is more generous than unemployment insurance, there will be strong incentives to delay the return to the market. In particular, the authors show that the hazard of leaving WC is 27% lower when an accident occurs at the end of the construction season, when unemployment is seasonally maximum[24].

Chiappori, Durand and Geoffard (1998) use data on health insurance that display similar features. Following a change in regulation in 1993, French health insurance companies modified the coverage offered by their contracts

---

[24] See also Fortin and Lanoie (1992), Bolduc et al. (1997), and the survey by Fortin and Lanoie (1998).

in a non uniform way. Some of them increased the level of deductible, while others did not. The tests use a panel of clients belonging to different companies, who were faced with different changes in coverage, and whose demand for health services are observed before and after the change in regulation. In order to concentrate upon those decisions that are essentially made by consumers themselves (as opposed to those partially induced by the physician), the authors study the occurrence of a physician visit, distinguishing between general practitioner (GP) office visits, GP home visits and specialist visits. They find that the number of GP home visits significantly decreased for the agents who experienced a change of coverage, but not for those for which the coverage remained constant. They argue that this difference is unlikely to result from selection, since the two populations are employed by similar firms, display similar characteristics, and participation in the health insurance scheme was mandatory.

Finally, a recent paper by Weisburd (2011) uses an intriguing quasi-experiment in which a large Israeli firm covered car insurance premia for some of its employees. These lucky employees only had to pay the deductible if they filed a claim, and the firm would also cover the increase in premia that resulted from experience rating. As a result, those employees who did not benefit from the scheme faced steeper incentives; and to the extent that employees were randomly assigned between the two groups, differences in claims isolate the incidence of moral hazard. Weisburd argues that this is indeed the case; she finds that as expected, employer-paid premia are associated with more claims.

2.4.2 Quasi-natural experiments

Natural experiments are valuable but scarce. In some cases, however, one finds situations that keep the flavor of a natural experiment, although no exogenous *change* of the incentive structure can be observed. The key remark is that any situation were identical agents are, for *exogenous* reasons, faced with different incentive schemes can be used for testing for moral hazard. The problem, of course, is to check that the differences in schemes are purely exogenous, and do not reflect some hidden characteristics of the agents. For instance, Chiappori and Salanié (2000) consider the case of French automobile insurance, where young drivers whose parents have low past accident rates can benefit from a reduction in premium. Given the particular properties of the French experience rating system, it turns out that the marginal cost of accident is reduced for these drivers. In a moral hazard context, this should result in less cautious behavior and higher accident probabilities. If, on the contrary, the parents' and children's driving abilities are (positively) correlated, a lower premium should signal a better driver, hence translate into less accidents. The specific features of the French situation thus allows to distinguish between the two types of effects. Chiappori and Salanié find evidence in favor of the second explanation: other things equal, "favored" young drivers have slightly fewer claims.

A contribution by Cardon and Hendel (1998) uses similar ideas in a very stimulating way. They consider a set of individuals who face different menus of employer-based health insurance policies, under the assumption that there is no selection bias in the allocation of individuals across employers. Two types of behavior can then be observed. First, agents choose one particular policy within the menu at their disposal; second, they decide on the level of health expenditures. The authors identify a fully structural model, which allows them to simultaneously estimate a selection equation that describe the policy choice, and estimate the price elasticity of demand controlling for selection bias. The key ingredient for identifying the specific effects of moral hazard is that while people are free to choose any contract in the menu they face, they cannot choose the menu itself; and different menus involve different coinsurance levels. The "quasi-experimental" features stem precisely from this random assignment of people to different choice sets. Even if less risky people always choose the contract with minimum coinsurance, the corresponding coinsurance rates will differ across firms. In other words, it is still the case that identical people in different firms face different contracts (i.e., different coinsurance rates) for exogenous reasons (i.e., because of the choice made by their employer). Interestingly enough, the authors find no evidence of adverse selection, while price elasticities are negative and very close to those obtained in the Rand HIE survey. This suggests that moral hazard, rather than adverse selection, may be the main source of asymmetric information in that case.

3 Dynamic models of information asymmetries

Tests based on the dynamics of the contractual relationship can throw light on the predictions of models of asymmetric information. In addition, moral hazard and adverse selection models have quite different predictions in dynamic situations; therefore dynamic studies offer an opportunity to disentangle them.

Empirical studies exploiting dynamics can be gathered into two broad categories. First, some work assumes that observed contracts are optimal, and compares their qualitative features with theoretical predictions in both a moral hazard and an adverse selection framework. While the derivation of diverging predictions is not always easy in a static context, the introduction of dynamic considerations greatly improves the picture.

Natural as it seems, the assumption that contracts are always optimal may not be warranted in some applications. For one thing, theory is often inconclusive. Little is known, for instance, on the form of optimal contracts in a repeated moral hazard framework, at least in the (realistic) case where the agent can freely save. And the few results we have either require utterly restrictive assumptions (CARA utilities, monetary cost of effort) or exhibit features (randomized contracts, for instance) that sharply contrast with real life observations. Even skeptics of bounded rationality theories may accept that such very sophisticated constructs, that can hardly be understood by the average insurance salesman (let alone the average consumer), are unlikely to be implemented on a large scale[25].

Another potential deviation from optimality comes from the existence of regulations, if only because regulations often impose very simple rules that fail to reproduce the complexity of optimal contracts. An interesting example is provided by the regulation on experience rating by automobile insurance companies, as implemented in some countries. A very popular rule is the "bonus/malus" scheme, whereby the premium is multiplied by some constant

---

[25] A more technical problem with the optimality assumption is that it tends to generate complex endogeneity problems. Typically, one would like to compare the features of the various existing contracts. The optimality approach requires that each contract is understood as the optimal response to a specific context, so that differences in contracts simply reflect differences between the "environments" of the various firms. In econometric terms, contracts are thus, by assumption, endogenous to some (probably unobserved) heterogeneity across firms, a fact that may, if not corrected, generate biases in the estimations.

larger (resp. smaller) than one for each year with (resp. without) accident. Theory strongly suggests that this scheme is too simple in a number of ways. In principle, the malus coefficient should not be uniform, but should vary with the current premium and the driver's characteristics; the deductible should vary as well; etc.[26].

3.1 Tests assuming optimal contracts

Only a few empirical studies consider the dynamics of insurance relationships. An important contribution is due to Dionne and Doherty (1994), who use a model of repeated adverse selection with one-sided commitment. Their main purpose is to test the "highballing" prediction, according to which the insurance company should make positive profits in the first period, compensated by low, below-cost second period prices. They test this property on Californian automobile insurance data. According to the theory, when various types of contracts are available, low risk agents are more likely to choose the experience rated policies. Since these are characterized by highballing, the loss to premium ratio should rise with the cohort age. If insurance companies are classified according to their average loss per vehicle (which reflects the "quality" of their portfolio), one expects the premium growth to be negative for the best quality portfolios; in addition, the corresponding slope should be larger for firms with higher average loss ratios. This prediction is confirmed by the data: the "highballing" prediction is not rejected. Interestingly, this prediction contrasts with those of a standard model involving switching costs, in which insurers would actively compete in the first period, typically resulting in below-cost initial premium, and overcharge the clients thus acquired in the following periods.

Hendel and Lizzeri (2003) have provided very convincing tests of a symmetric learning model à la Harris and Holmstrom (1982) on life insurance data. Theory tells us that contracts involving commitment from the insurer, in the sense that the dynamics of future premium is fixed in advance and cannot depend on the evolution of the insuree's situation, should entail front loading, representing the cost of the insurance against the classification risk. Some contracts involve commitment from the insurer, in the sense that the dynamics of future premium is fixed in advance and cannot depend on the

---

[26] Of course, the precise form of the optimal scheme depends on the type of model. It is however basically impossible to find a model for which the existing scheme is optimal.

evolution of the insuree's health. For other contracts, however, future premia are contingent on health. Specifically, the premium increases sharply unless the insured is still in good health (as certified, for instance, by a medical examination). In this context, the symmetric learning model generates very precise predictions on the comparison between contracts with and without commitment. Contracts with non contingent future premia should entail front loading, representing the cost of the insurance against the classification risk. They should also lock-in a larger fraction of the consumers, hence exhibit a lower lapsation rate; in addition, only better risk types are likely to lapse, so that the average quality of the insurer's client portfolio should be worse, which implies a higher present value of premiums for a fixed period of coverage. Hendel and Lizzeri show that all of these predictions are satisfied by existing contracts[27]. Finally, the authors study accidental death contracts, i.e. life insurance contracts that only pay if death is accidental. Strikingly enough, these contracts, where learning is probably much less prevalent, exhibit none of the above features.

Another characteristic feature of the symmetric learning model is that any friction reducing the clients' mobility, although ex post inefficient, is often ex ante beneficial, because it increases the agents' ability to (implicitly) commit and allows for a larger coverage of the classification risk. Using this result, Crocker and Moran (1998) study employment-based health insurance contracts. They derive and test two main predictions. One is that when employers offer the same contract to all of their workers, the coverage limitation should be inversely proportional to the degree of worker commitment, as measured by his level of firm-specific human capital. Secondly, some contracts offer "cafeteria plans", whereby the employee can choose among a menu of options. This self-selection device allows the contract to change in response to interim heterogeneity of insurees. In this case, the authors show that the optimal (separating) contract should exhibit more complete coverage, but that the premiums should partially reflect the health status.

---

[27] The main puzzle raised by these findings is that a significant fraction of the population does not choose commitment contracts, i.e., does not insure against the classification risk. The natural explanation suggested by theory (credit rationing) is not very convincing in that case, since differences in premiums between commitment and no commitment contracts are small (less than $300 per year), especially for a client pool that includes executives, doctors, businessmen and other high income individuals. Heterogeneous risk perception across individuals is a better story, but formal tests still have to be developpped. Obviously, more research is needed on this issue.

Both predictions turn out to be confirmed by the data. Together with the results obtained by Hendel and Lizzeri, this fact that strongly suggests the symmetric learning model is particularly adequate in this context.

3.2 Behavioral dynamics under existing contracts

Another branch of research investigates, for given (not necessarily optimal) insurance contracts, the joint *dynamics* of contractual choices and accident occurrence. This approach is based on the insight that these properties are largely different under moral hazard, and that these differences lead to powerful tests. A classical example involves the type of experience rating typical of automobile insurance, whereby the occurrence of an accident at date t has an impact on future premia (at date $t+1$ and beyond). In general, existing experience rating schedules are highly non linear; the cost of the marginal accident in terms of future increases in premium is not constant, and often actually non monotonic.[28] In a moral hazard framework, these changes in costs result in changes in incentives, and ultimately in variations in accident probabilities; under pure adverse selection, on the contrary, the accident probabilities should either remain constant or change is a systematic way (e.g., through aging), irrespective of the accident history.

One idea, therefore, is to use theory to derive the main testable features of individual behavior for the various models at stake. Abbring et al. (2003) and Abbring, Chiappori and Pinquet (2003) develop a test of this type. The test is based on the so-called "negative contagion" effect. With many experience rating schemes, the occurrence of an accident increases the cost of the next one, therefore the insuree's incentives to avoid it. Under moral hazard, a reduction of its probability of occurrence should result. In principle, the variations in individual accident probabilities that follow the occurrence of an accident can be statistically detected using adequate techniques. The main empirical challenge, however, is to disentangle such fairly small fluctuations from the general, background noise due to unobserved heterogeneity. Should one simply look at the intertemporal correlation of accident occurrences among agents, the dominant phenomenon by far reflects some time-invariant (or time-correlated) unobserved heterogeneity: good drivers are less likely both to have had an accident in the past and to have one in the future.

---

[28] Typically, the cost of the first accident is low; marginal costs then increase, peak and drop sharply. See for instance Abbring et al (2009).

Technically, the "negative contagion" property holds only *conditionally* on agents' characteristics, including unobserved ones; any empirical test must therefore control for the latter.

This problem, which is quite similar to the distinction between state dependence and unobserved heterogeneity in the labor literature (see Heckman and Borjas, 1980, and Heckman, 1981) can in principle be solved when panel data are available. In practice, the authors use French data, for which regulation imposes that insurers increase the premium by 25% in case an accident occurs; conversely, in the absence of any accident during one year, the premium drops by 5%. The technique they suggest can be intuitively summarized as follows. Assume the system is malus only (i.e., the premium increases after each accident, but does not decrease subsequently), and consider two sequences of 4-year records, $A = (1, 0, 0, 0)$ and $B = (0, 0, 0, 1)$, where 1 (resp. 0) corresponds to the occurrence of an accident (resp. no accident) during the year. In the absence of moral hazard, and assuming away learning phenomena, the probability of the two sequences should be exactly identical; in both cases, the observed accident frequency is 25%. Under moral hazard, however, the first sequence is more probable than the second : in A, the sequence of three years without accident happens after an accident, hence when the premium, and consequently the marginal cost of future accidents and the incentives to take care are maximum.

One can actually exploit this idea to construct general, non parametric tests of the presence of moral hazard. The intuition goes as follows. Take a population of drivers observed over a given period; some drivers have no accident over the period, others have one, two or more. Assume for simplicity a proportional hazard model, and let H_1 be the distribution of the first claim time T_1 in the subpopulation with exactly one claim over the period. Note that H_1 needs not be uniform; with learning, for instance, claims are more likely to occur sooner than later. Similarly, define H_2 to be the distribution of the second claim time T_2 in the subpopulation with exactly two claims in the contract year. In the absence of moral hazard, it must be the case that

$$H_1(t)^2 = H_2(t),$$

a property that can readily be tested non parametrically.

These initial ideas have recently been extended by Abbring, Chiappori and Zavadil (2008), Dionne et al. (2009) and Dionne, Michau and Dahchour (forthcoming). The first paper, in particular, explicitly models the forward-looking behavior of an agent in the Dutch automobile insurance market, which exhibits a highly non linear bonus-malus scheme; they then use this model to compute the dynamic incentives faced by an agent and to construct a structural test that exploits these computations in detail. Their framework explicitly distinguishes ex ante and ex post moral hazard and models both claim occurrences and claim sizes. Interestingly, all three papers (using respectively Dutch, Canadian and French data) find evidence of (ex ante and ex post) moral hazard, at least for part of the population, and compute the magnitude of the resulting effect.

Conclusion

As argued in the introduction, empirical applications of contract theory have become a *bona fide* subfield; and insurance data has played a leading role in these developments. This literature has already contributed to a better knowledge of the impact of adverse selection and moral hazard in various markets. The practical importance of information asymmetries has been found to vary considerably across markets. In particular, there exists clear and convincing evidence that some insurance markets are indeed affect by asymmetric information problems, and that the magnitude of these problems may in some cases be significant.

There exist a number of crucial normative issues where our theoretical and empirical knowledge of asymmetric information are likely to play a central role. To take but one example, a critical feature of the recent reform of the US health insurance system (PPACA, 2010) is the prohibition of the use of preexisting conditions in the underwriting process. While the benefits of such a measure are clear in terms of ex ante welfare and coverage of the "classification risk", some of its potential costs have been largely underanalyzed. In particular, the prohibition would introduce a massive amount of adverse selection (since agents have a detailed knowledge of the preexisting conditions that insurers are not allowed to use), into a system that remains essentially market-oriented. From a theoretical viewpoint, the consequences may (but need not) be dramatic; for instance, should a RS equilibrium prevail, one might expect the persistence of a de facto price discrimination based on preexisting conditions, coupled with huge welfare losses due to severe restrictions of coverage for low risk individuals.[29] Whether such problems will ultimately arise, with which magnitude, and which regulation would be

---

[29] See Chiappori (2006) for a preliminary investigation of these effects.

needed to mitigate them remains a largely unexplored empirical question; even the basic information needed to attempt a preliminary welfare evaluation (e.g., the joint distribution of income, health risk and risk aversion) is only very partially known. Some pioneering studies have taken steps in this direction however (see Einav, Finkelstein, and Levin 2010), and one can only hope that our ability to simulate the effect of such reforms will improve in the near future.

Finally, a better understanding of actual behavior is likely to require new theoretical tools. The perception of accident probabilities by the insurees, for instance, is a very difficult problem on which little is known presently. Existing results, however, strongly suggest that standard theoretical models relying on expected utility maximization using the "true" probability distribution may fail to capture some key aspects of many real-life situations. Our analysis in section 1.2 shows that the positive correlation property should hold on perfectly competitive markets under fairly weak conditions on the rationality of agents; but with market power there is much we still need to learn. The confrontation of new ideas from behavioral economics with insurance data is likely to be a very promising research direction in coming years.

References

Abbring, J., P. A. Chiappori, J. Heckman and J. Pinquet (2003), "Testing for Moral Hazard on Dynamic Insurance Data", *Journal of the European Economic Association*, Papers and Proceedings, 1, 2-3, 512-521

Abbring, J., P. A. Chiappori and J. Pinquet (2003) "Moral Hazard and Dynamic Insurance Data", *Journal of the European Economic Association* 1, 4, 767-820.

Abbring, J., P. A. Chiappori and T. Zavadil (2008) "Better Safe than Sorry? Ex Ante and Ex Post Moral Hazard in Dynamic Insurance Data", Mimeo, Columbia University.

Bolduc, D., Fortin, B., F. Labrecque and P. Lanoie (1997), "Incentive Effects of Public Insurance Programs on the Occurrence and the Composition of Workplace Injuries", CIRANO Scientific Series, Montreal, 97s-24.

Boyer, M., et G. Dionne (1989) : "An Empirical Analysis of Moral Hazard and Experience Rating", *Review of Economics and Statistics*, 71, 128-34

Browne, M., and R. Puelz (1998): "The Effect of Legal Rules on the Value of Economic and Non-Economic Damages and the Decision to File", mimeo, University of Wisconsin-Madison

Brugiavini, A. (1990) : "Longevity risk and the life cycle", PhD Dissertation, LSE, London

Cardon, J., and I. Hendel (2001): "Asymmetric Information in Health Insurance : Evidence From the National Health Expenditure Survey", *RAND Journal of Economics*, v. 32, iss. 3, pp. 408-27

Cawley. J. and T. Philipson (1999) : "An Empirical Examination of Information Barriers to Trade in Insurance", *American Economic Review*, v. 89, iss. 4, pp. 827-46

Chassagnon, A., and P.A. Chiappori (1997): "Insurance under moral hazard and adverse selection: the competitive case", mimeo, DELTA.

Chiappori, P. A. (1994): "Assurance et économétrie des contrats : quelques directions de recherche", mimeo, DELTA.

Chiappori, P. A. (2006): "The Welfare Effects of Predictive Medicine", in *Insurance: Theoretical Analysis and Policy Implications*, CESifo conference volume, P.A. Chiappori and C. Gollier, Ed., MIT Press, Boston, p. 55-80.

Chiappori, P. A., F. Durand and P.Y. Geoffard (1998), "Moral Hazard and the Demand for Physician Services : First Lessons from a French Natural Experiment", *European Economic Review*, 42, 499-511

Chiappori, P. A., B. Jullien, B. Salanié and F. Salanié (2006), "Asymmetric Information in Insurance: General Testable Implications", *Rand Journal of Economics*, 37 4.

Chiappori, P. A., I. Macho, P. Rey and B. Salanié (1994), "Repeated moral hazard : memory, commitment, and the access to credit markets", *European Economic Review,* 1527-53.

Chiappori, P. A. and B. Salanié (1997), "Empirical Contract Theory: The Case of Insurance Data", *European Economic Review*, 41, 943-51

Chiappori, P. A. and B. Salanié (2000), "Testing for Asymmetric Information in Insurance Markets", *Journal of Political Economy*, 108, 56-78.

Chiappori, P. A., B. Salanié and J. Valentin (1999), "Early Starters vs Late Beginners", *Journal of Political Economy*, 107, 731-60.

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

Cohen, A. and P. Siegelman (2010, "Testing for Adverse Selection in Insurance Markets", *Journal of Risk and Insurance*, 77, 39–84.

Cooper, R., and B. Hayes (1987) : "Multi-period Insurance Policies", *International Journal of Industrial Organization*, 5, 211–231.

Crocker, K. and J. Moran (2003) : "Contracting with Limited Commitment : Evidence from Employment-Based Life Insurance Contracts", *RAND Journal of Economics*, v. 34, iss. 4, pp. 694-718

Cutler, D., A. Finkelstein, and K. McGarry (2008): "Preference Heterogeneity and Insurance Markets: Explaining a Puzzle of Insurance", *American Economic Review*, 98, 157–162.

Cutler, D. and R. Zeckhauser (2000): "The Anatomy of Health Insurance", in *Handbook of Health Economics*, A. J. Culver and J,. Newhouse eds, 563–643, North Holland.

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

Dahlby, B. (1992), "Testing for Asymmetric Information in Canadian Automobile Insurance", in *Contributions to Insurance Economics*, G. Dionne ed, Kluwer.

de Garidel-Thoron, T. (2005): ???? SCRATCH???? "Welfare-improving Asymmetric Information in a Dynamic Insurance Market", *Journal of Political Economy*, 113, 121–150.

de Meza, D, and D. Webb (2001): "Advantageous Selection in Insurance Markets", *Rand Journal of Economics*, 32, 249–262.

Dionne, G, Ed. (1992): *Contributions to Insurance Economics*, Kluwer, Boston.

Dionne, G., and N. Doherty (1994), "Adverse Selection, Commitment and Renegotiation: Extension to and Evidence from Insurance Markets", *Journal of Political Economy*, 102-2, 210-35.

Dionne, G., C. Gouriéroux and C. Vanasse (1997), "Testing for Evidence of Adverse Selection in the Automobile Insurance Market: A Comment", *Journal of Political Economy*, 109, 444–453.

Dionne, G., Michaud, P.C., Dahchour, M. (forthcoming), "Separating Moral Hazard from Adverse Selection and Learning in Automobile Insurance: Longitudinal Evidence from France", *Journal of the European Economic Association.*

Dionne, G., J. L. Pinquet, C. Vanasse and M. Maurice (2009): "Incentive Mechanisms for Safe Driving: A Comparative Analysis with Dynamic Data", Working paper, Economix, Nanterre.

Dionne, G, and P. St-Michel (1991): "Workers' Compensation and Moral hazard', *Review of Economics and Statistics*, LXXIII, 236-44.

Dionne, G. and C. Vanasse (1992), "Automobile Insurance Ratemaking in the Presence of Asymmetrical Information", *Journal of Applied Econometrics*, 7, 149-65

Dionne, G. and C. Vanasse (1996), "Une évaluation empirique de la nouvelle tarification de l'assurance automobile au Québec", Mimeo, Universite de Montreal.

Einav, L., A. Finkelstein, and J. Levin (2010), "Beyond Testing: Empirical Models of Insurance Markets", *Annual Review of Economics*, 2, 311–336.

Fang, H., M. Keane, and D. Silverman (2008), "Sources of Advantageous Selection: Evidence From the Medigap Insurance Market", *Journal of Political Economy*, 116, 303–350.

Finkelstein, A., and K. McGarry (2006), "Multiple Dimensions of Private Information: Evidence From the Long-Term Care Insurance Market", *American Economic Review*, 96, 938–958.

Finkelstein, A., and J. Poterba (2004), "Adverse Selection in Insurance Markets: Policy-holder Evidence From the U.K. Annuity Market", *Journal of Political Economy*, 112, 183–208.

Finkelstein, A., and J. Poterba (2006), "Testing for Adverse Selection with 'Unused Observables'," NBER Working Paper 12112.

Fortin, B. and P. Lanoie (1992), "Substitution between Unemployment Insurance and Workers' Compensation", *Journal of Public Economics*, 49, 287-312.

Fortin, B., P. Lanoie and C. Laporte (1995), "Is Workers' Compensation Disguised Unemployment Insurance", CIRANO Scientific Series, Montreal, 95s-48.

Fortin, B and P. Lanoie, "Effects of Workers' Compensation: a Survey" (1998), CIRANO Scientific Series, Montreal, 98s-04.

Friedman, B.M., and M.J. Warshawski (1990) : "The cost of annuities : implications for savings behavior and bequests", *Quarterly Journal of Economics*, 420, 135-54

Fudenberg, D., B. Holmstrom and P. Milgrom (1990), "Short-Term Contracts and Long-Term Agency Relationship", *Journal of Economic*

Theory, 51, 1-31

Glied, S. (2000):, “Managed Care”, in *Handbook of Health Economics*, A. J. Culver and J,. Newhouse eds, 707–753, North Holland.

Gouriéroux, C. (1999): “The Econometrics of Risk Classification in Insurance’, *Geneva Papers on Risk and Insurance Theory*, 24, 119-139.

Handel, B. (2011): “Adverse Selection and Switching Costs in Health Insurance Markets: When Nudging Hurts”, mimeo Berkeley.

Harris M. and B. Holmstrom (1982), “A Theory of Wage Dynamics”, *Review of Economic Studies*, 49, 315-333.

Hart, O. and J. Tirole (1988) : “Contract Renegotiation and Coasian Dynamics”, *Review of Economic Studies*, 55, 509-40.

He, D. (2009) : “The Life Insurance Market: Asymmetric Information Revisited”, *Journal of Public Economics*, 93, 1090–1097.

Heckman, J.J. (1978): “Simple statistical models for discrete panel data developed and applied to test the hypothesis of true state dependence against the hypothesis of spurious state dependence, *Annales de l’INSEE*, 30-31, p. 227-269.

Heckman, J. J. (1981): "Heterogeneity and state dependence", in S. Rosen (Ed.), *Studies in Labor Markets*. University of Chicago Press. 3, 32

Heckman, J. J. and G. Borjas (1980): "Does unemployment cause future unemployment? Definitions, questions and answers from a continuous time model of heterogeneity and state dependence", *Economica* 47, 247-283. 3, 32

Hellwig, M. (1987) : “Some recent developments in the theory of competition in markets with adverse selection”, *European Economic Review*, 31:154–63.

Hemenway, D. (1990), “Propitious Selection”, *Quarterly Journal of Economics*, 105, 1063–1069.

Hendel, I. and A. Lizzeri (2003): “The Role of Commitment in Dynamic Contracts : Evidence from Life Insurance”, *The Quarterly Journal of Economics*, MIT Press, vol. 118(1), pages 299-327.

Hirshleifer, J. (1971): “The Private and Social Value of Information and the Reward to Inventive Activity,” *American Economic Review*, 61, 561-574.

Jullien, B., B. Salanié and F. Salanié (1999): “Should More Risk-Averse Agents Exert More Effort?”, *Geneva Papers on Risk and Insurance Theory*, 24, 19-28.

Jullien, B., B. Salanié and F. Salanié (2007): "Screening Risk-averse Agents under Moral Hazard: Single-crossing and the CARA Case", *Economic Theory*, 30, 151–169.

Laffont, J.-J. and J. Tirole (1993) : *A Theory of Incentives in Procurement and Regulation*, Cambridge, MIT Press.

Manning, W. et al. (1987), "Health insurance and the demand for medical care : evidence from a randomized experiment," *American Economic Review*, 77(3) : 251-277.

Pouyet, J., B. Salanié and F. Salanié (2008), "On Competitive Equilibria with Asymmetric Information", *The B.E. Journal of Theoretical Economics, Topics*, 8, article 13.

Puelz, R. and A. Snow (1994), "Evidence on Adverse Selection : Equilibrium Signalling and Cross-Subsidization in the Insurance Market", *Journal of Political Economy*, 102, 236-57.

Richaudeau, D. (1999): "Automobile insurance contracts and risk of accident: An Empirical Test Using French Individual Data", *The Geneva Papers on Risk and Insurance Theory,* 24, 97-114

Rochet, J.C., and L. Stole (2003): "The Economics of Multidimensional Screening", in *Advances in Economics and Econometrics: Theory and Applications—Eight World Congress*, Cambridge University Press.

Rogerson, W. (1985): "Repeated Moral Hazard", *Econometrica*, 53, 69-76.

Rothschild, M. and J. Stiglitz (1976), "Equilibrium in Competitive Insurance Markets", *Quarterly Journal of Economics*, 90, 629-649.

Rubinstein, A., and M. Yaari (1983): "Insurance and Moral Hazard", *Journal of Economic Theory*, 14, 441-52.

Salanié, B. (2005), *The Economics of Contracts: A Primer*, MIT Press.

Villeneuve, B. (2003), "Concurrence et antisélection multidimensionnelle en assurance", *Annales d'Economie et Statistique*, 69, 119-142.

Weisburd, S. (2011), "Identifying Moral Hazard in Car Insurance Contracts", mimeo Tel Aviv.