Assurances

ASYMMETRIC INFORMATION IN AUTOMOBILE INSURANCE: AN OVERVIEW

Pierre-André Chiappori

Volume 66, numéro 4, 1999

URI : https://id.erudit.org/iderudit/1105238ar
DOI : https://doi.org/10.7202/1105238ar

Aller au sommaire du numéro

Éditeur(s)

HEC Montréal

ISSN

0004-6027 (imprimé)
2817-3465 (numérique)

Découvrir la revue

Citer cet article

Chiappori, P.-A. (1999). ASYMMETRIC INFORMATION IN AUTOMOBILE INSURANCE: AN OVERVIEW. *Assurances*, *66*(4), 629–644. https://doi.org/10.7202/1105238ar

Résumé de l'article

Le but de cet article est de présenter une brève revue des modèles empiriques utiles pour tester ou évaluer l'importance de l'asymétrie d'information pour les marchés d'assurance automobile. L'économie de l'assurance moderne a été influencée par les développements récents de la théorie des contrats. Notre compréhension des aspects cruciaux des formes des contrats optimaux, de la concurrence dans les marchés d'assurance ou du rôle de la réglementation publique, pour n'en nommer que quelques-uns, réfère aux concepts de base de la théorie des contrats—risque moral, antisélection, engagement, renégociation et autres. Il est juste de dire que l'assurance a été et est toujours un des champs importants et prometteurs d'application empirique des contrats d'assurances.

ASYMMETRIC INFORMATION IN AUTOMOBILE INSURANCE: AN OVERVIEW*

by Pierre-André Chiappori

ABSTRACT

The goal of this paper is to briefly review a number of empirical models that explicitly aim at testing for or evaluating the importance of asymmetric information in automobile insurance. Modern insurance economics has been deeply influenced by the recent developments of contract theory. Our understanding of such crucial aspects 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 extend still remains, one of the most important and promising field of empirical application for contract theory.

RÉSUMÉ

Le but de cet article est de présenter une brève revue des modèles empiriques utiles pour tester ou évaluer l'importance de l'asymétrie d'information pour les marchés d'assurance automobile. L'économie de l'assurance moderne a été influencée par les développements récents de la théorie des contrats. Notre compréhension des aspects cruciaux des formes des contrats optimaux, de la concurrence dans les marchés d'assurance ou du rôle de la réglementation publique, pour n'en nommer que quelques-uns, réfère aux concepts de base de la théorie des contrats—risque moral, antisélection, engagement, renégociation et autres. Il est juste de dire que l'assurance a été et est toujours un des champs importants et prometteurs d'application empirique des contrats d'assurances.

L'auteur :

Pierre-André Chiappori is professor of economics at the University of Chicago.

* Financial support from the Chaire d'Économie de l'Assurance (Paris) is gratefully aknowledged. Errors are mine.

■ INTRODUCTION

Modern insurance economics has been deeply influenced by the recent developments of contract theory. Our understanding of such crucial aspects 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 extend still remains, one of the most important and promising field of empirical application for contract theory.

By their very nature, insurance data provide nearly ideal material for testing the predictions of contract theory. As argued by Chiappori (1994) and Chiappori and Salanié (1997), most predictions of contract theory are expressed in terms of a relationship between, on the one hand, some "performance" that characterizes the outcome of the relationship under consideration, and on the other hand some transfers taking place between the parties. Under moral hazard, for instance, the transfer will be positively correlated with the outcome, but in a smoothed way, 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 latter being generally proposed by the other party. Also, the exact translation of the notions of "performance" and "transfer" obviously varies with the particular field at stake. Depending on the particular context, the "performance" may be a production, a profit, the realization 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 extend standardized, and large samples should be considered, in order to apply the usual tools of econometric analysis. As it turns out, data of this kind are quite scarce. In some contexts, the contract is essentially implicit, and its true implications are uneasy to grasp. More frequently, contracts do not present a standardized form because of the complexity of the information needed either to characterize the various (and possibly abundant) states of the world that should be considered, or to precisely describe each party's 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. A typical example is repeated contracts, where the history of past relationship may provide crucial indications that in general are not (fully) available for the purpose of empirical observation. Last but not least, the "performance" is often not recorded, or even not precisely defined. In the case of labor contracts, for instance, the employee's "performance" is often the product of a supervisor's subjective estimation, and may not be recorded on the firm's files.

In contrast, insurance contracts basically fulfill all of the previous requirements. Automobile insurance provides a typical example. Here, contracts are largely standardized. The insurer's information is accessible, and can generally be summarized through a reasonable number of quantitative or qualitative indicators. The "performance" – whether it represents the occurrence of an accident, its cost, or both – is in general 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 go. It should thus be no surprise that empirical tests of adverse selection, moral hazard or repeated contract theory on insurance data, and especially automobile insurance, has attracted considerable attention.

The goal of this paper is to briefly review a number of empirical models that explicitly aim at testing for or evaluating the importance of asymmetric information in automobile insurance. The structure of this contribution is as follows. We first review some of the main theoretical issues at stake. We argue, in particular, that while adverse selection and moral hazard are generally recognized as cornerstones of modern contract theory, empirically distinguishing between these concepts may be quite difficult, especially when only "static" (cross-sectional) data are available. Then we briefly describe several contributions explicitly aimed at testing for asymmetric information in automobile insurance. The main conclusions are outlined in the last section.

THE THEORETICAL BACKGROUND

It is by now customary to distinguish between 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.

☐ Adverse selection

Adverse selection arises when one party – generally, the subscriber – has a better information than the other party – the insurer – about some parameter that is relevant for the relationship. Most of the time, the informational advantage is linked with the level of risk; typically, the issue will be whether the client knows better her accident probability, or the (conditional) distribution of losses incurred in case of accident. An key feature is that, in such cases, the agent's informational advantage is directly related to the insurer's (expected) cost of providing the contract.

A first point that should be emphasized is that, whenever empirical applications are concerned, the agent's better knowledge of her risk is not the only possible source of asymmetry, and possibly not the most important one. There are good reasons to believe, for instance, that the insureds also know better their own preferences, and particularly their level of risk aversion – although this aspect is often disregarded in theoretical models. A possible justification for this lack of interest is that, in principle, adverse selection on preferences has negligible consequences upon the form and the outcome of the relationship, at least in a context of pure competition. Competition typically imposes that companies always charge a fair premium, unless the latter cannot be directly computed (which is precisely the case when the agent's risk is not known). Hence, the equilibrium contract should not depend on the subscriber's preferences, whether the latter are public or private. In particular, in a model of competitive insurance markets with perfect information, the introduction of hidden information on preferences will not alter the equilibrium outcome.

This conclusion should however be qualified, for at least two reasons. For one thing, perfect competition is a natural assumption within a simplified theoretical model, but much less so in reality. Fixed costs, product differentiation, price stickiness, switching costs and cross-subsidization are part of the real world; oligopoly is probably the rule rather than the exception. In this context, firms are able to make positive profits, that are related to the agents' demand elasticity; the latter, in turn, directly reflects risk aversion. To take an extreme case, it is well known that in a principal-agent framework – equivalent to some monopoly position of the insurance company – adverse selection on risk aversion does matter for the form of the optimal contract.

A second caveat is that even when adverse selection on preferences *alone* does not matter, it may still, when added to asymmetric information of a more standard form, considerably modify the properties of equilibria. In a standard Rothchild-Stiglitz (from now RS) context, for instance, heterogeneity in risk aversion may result in violations of the classical single-crossing property of indifference curves "à la Spence-Mirrlees", which in turn generates new types of competitive equilibria[2]. More generally, situations of bi- or multi-dimensional adverse selection are much more complex than the standard ones, and may require more sophisticated policies[3].

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 that emerge from the theoretical models? Considering the case of pure competition, the answer is not straightforward. It obviously depends, among other things, on the particular definition of an equilibrium that is adopted. It is fair to say, however, that no general agreement has been reached on this issue. 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 conclusion holds for equilibria à 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 clearly illustrated 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 expect to hold in more general setting. Still, one can argue, following Chiappori and Salanié (1997), that three conclusions seem fairly robust; namely:

1. under adverse selection, agents are likely to be faced with menus of contracts, among which they are free to choose;

2. contracts with more comprehensive coverage are sold at a higher (unitary) premium;

3. contracts with more comprehensive coverage are chosen by agents with higher expected accident costs.

The first prediction is essentially qualitative; note that it holds for different types of adverse selection (i.e., agents may differ by their risk, but also by their wealth, preferences, risk aversion, etc.). The second prediction, in most circumstances, essentially reflects individual rationality: if pricing is approximately fair, an agent will not choose a contract with higher deductible (or more coinsurance) unless its unitary price is lower[4]. Again, this is not specific of adverse selection à la RS, where the agent's private information is related to his riskiness. Testing for this property is an interesting perspective, that has been followed by various authors. It however requires an explicit and adequate estimation of the firm's pricing policy, which may in some case raise difficult technical problems.

In contrast, the third property can be tested without estimating the pricing policy of the firm. If agents, facing the same menu of contracts (sold at identical fares), self select on the basis of some private information they have about their riskiness, then a positive correlation between coverage and expected costs should be observed, whatever the prices that were proposed in the initial stage. It should be noted that this prediction seems quite robust. For instance, it does not require single crossing, and it holds when moral hazard or multidimensional adverse selection are introduced; also, it remains valid in a dynamical setting[5].

This claim must however be qualified, or at least clarified. What must be stressed, at this point, is that this prediction is valid within a group of *observationally identical* agents. In practice, insurance companies use observable characteristics to categorize individual risks. As far as pricing *across* the classes thus constructed is concerned, the previous conclusions are totally irrelevant. Some agents may be offered contracts entailing both higher unitary premium and larger deductible[6]; the point being that they cannot choose the class they will be categorized into. The self-selection issue applies only *within* such classes. The empirical translation is that one must systematically consider probability distributions that are *conditional on all observables*. Although this requirement is in principle straightforward, how this conditioning is actually performed on "real" data is one of the key problems of this line of empirical investigation.

☐ Moral hazard

Moral hazard occurs when accident probabilities are not exogenous, but depend 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 at least adjust the rate accordingly. 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 decrease incentives to reduce risk. In the extreme case of complete insurance (when the insured's welfare simply does not depend on the occurrence of an accident), incentives are killed, resulting in maximum accident probabilities. More generally, different contracts provide different incentives, hence result in different observed accident rates. This is the bottomline of most empirical tests of moral hazard.

Quite interestingly, the basic moral hazard story is very close to the adverse selection one, except for an inverted causality. Under adverse selection, people are characterized by different levels of risk (that will later be translated into dissimilar accident rates); because of these discrepancies, they choose different contracts. In a context of moral hazard, people first choose different contracts; then they are faced with different incentive schemes, hence adopt more or less cautious behavior, which ultimately results in heterogeneous accident probabilities. In both case, however, the conclusion is that, controlling for observables, the choice of a contract will be correlated with the accident probability – again, more comprehensive coverage being associated to higher risk. This suggests that it may be hard to distinguish between adverse selection and moral hazard in the static framework (i.e., using cross-sectional data). I may, as an econometrician, find out that, conditionally on observables, agents covered by a comprehensive automobile insurance contract are more likely to have an accident. But I hardly can say whether they chose full coverage because they knew their risk was higher, or whether, on the contrary, they became more risky because the comprehensive contract they selected for some exogenous reason killed most incentives to drive safely.

☐ Distinguishing adverse selection from moral hazard

The adverse selection versus moral hazard puzzle can be solved in different ways. One is to exploit some dynamics elements of the relationship. Whenever changes in the incentive structure can be observed on a given population, should these changes be exogenous (resulting for instance from a new regulation) or endogenous (as produced, say, by an experience rating pricing policy), then it should be possible to single out the consequences of incentives upon behavior, i.e., the moral hazard component. This path has been followed by several authors. A kind of static counterpart is when a sample of observationnally identical subscribers are faced with different incentive schemes, and it is known that the selection into the various schemes was not endogenous. An ideal situation would be a controlled experiment, where agents are randomly assigned to different schemes. The celebrated Rand study on medical expenditures (see Newhouse *et al.*) provides a perfect illustration of such a context.

Finally, the estimation of a fully specified structural model can in some cases allow to distinguish between the two aspects. In that case, however, the distinction may depend in a very fundamental way of the particular, parametric representation adopted. Then its robustness is not guaranteed.

☐ "Ex-post" moral hazard

The notion of ex-post moral hazard refers to a key feature of insurance data: what the insurer can observe are claims, not accident. In most cases, the decision to file a claim is made by the subscriber, and must be understood as a response to specific incentives. Should the costs of filing a claim exceed the expected benefits – say, because the expected cost is below the deductible, or experience rating implies that the claim will result in higher future premia – then the insured is always free not to declare.

This simple remark has two consequences. One is that the incentives to file a claim should be monitored by the insurance company, particularly when the processing of a small claim involves important fixed costs for the company. Deductibles, for instance, are often seen by insurance companies as a simple and efficient way of avoiding small claims. More related to the present topic is that fact that the empirical distribution of claims will in general be a truncation of that of accidents – since "small" accidents are typically not declared. Moreover, the truncation is endogenous; it depends on the contract (typically, on the deductible or the presence of experience rating), and also, possibly, on the individual characteristics of the insured (say, because the cost of higher future premia is generally related to the (expected) frequency of future accidents). This can potentially generate severe biases. To take an obvious example: if high deductibles discourage small claims, they lead to an automatic reduction of the number of *declared* accidents. This generates a (spurious) correlation between the choice of the contract and the observed level of risk, even in the absence of adverse selection or ex ante moral hazard. A basic problem of any empirical estimation, therefore, is to control for this potential biases.

EMPIRICAL ESTIMATIONS OF ASYMMETRIC INFORMATION IN THE STATIC FRAMEWORK

While the theoretical analysis of contracts under asymmetric infonnation began in the 70s, the empirical estimation of insurance models entailing either adverse selection or moral hazard is more recent. Among early contributions, one may mention Boyer and Dionne (1987) and Dahlby (1983), who does not reject the presence of some asymmetric information. However, Dahlby uses aggregate data only, so that it is not clear whether his results would be robust to the inclusion of more detailed individual data.

The hedonistic approach (Puelz and Snow 1994)

The field has however experienced a considerable development during the last decade. An important contribution is due to Puelz and Snow (1994), and relies upon an hedonistic model of insurance pricing. Using individual data from an automobile insurer in Georgia, they build a two-equation model of insurance contracts. The first equation represents the pricing policy adopted by the insurance firm. It takes the form:

$$P_i = g(D_i, X_i, \varepsilon_i)$$

where P_i and D_i are the premium and the deductible in the contract chosen by individual i, the X_i are individual-specific exogenous variables and ε_i is an econometric error term. This allows to directly test our second prediction – namely, that higher premia should be associated to lower deductible. This property is indeed confirmed by the data. However, as argued above, this result, per se, cannot provide a strong support to the existence of adverse selection. Whatever the reason for offering a menu of contracts, one hardly expects that rational insurees choose contracts with a higher unitary premium *and* a large deductible. More interesting is the test they propose for the third prediction – i.e., that the choice of a contract offering a more comprehensive coverage should be correlated with a higher accident probability. For this purpose, they estimate a second equation that describes the agent's choice of deductible. The latter depends on the agent's "price of deductible" $\hat{g}_D$, as estimated from a third equation not presented in the article, and on his (unobserved) accident probability. The latter is proxied by a dummy variable RT_i that equals one if the individual had an accident and zero otherwise. This leads to an equation of the form:

$$D_i = h(\hat{g}_{Di}, RT_i, X_i, \eta_i)$$

where η_i is another error term. The Rothschild-Stiglitz model predicts that higher risks buy better coverage, i.e. a lower deductible, so that h should decrease in RT. Puelz and Snow specify their first equation as a linear model and estimate it by ordinary least squares. Since there are only three levels of deductible in their data set, they estimate their second equation (again linear) by ordered logit; they find a negative coefficient for RT_i (although the choice of deductible does not vary much with the risk type).

□ Problems with the hedonistic approach

There are several problems in the Puelz-Snow approach, that provide an interesting illustration of the difficulties encountered by any attempt at testing the predictions of contract theory. A first (and somewhat technical) one is related to the approximation of the (unknown) accident probability by the dummy variable *RT*. This procedure introduces a measurement error in the second equation. In linear models, the estimates would be biased towards zero, which would reinforce the conclusion of Puelz-Snow. In an ordered logit, it is not clear which way the bias goes.

A second concern is that the data set under consideration comprises individuals of various ages and driving records. This important heterogeneity may be troublesome for two reasons. One is heteroscedasticity. Presumably, the distribution of the random shocks, and especially of η_i, will depend on the driver's seniority. Within a non linear model such as the ordered logit, this will bias the estimation. The second and more disturbing problem relates to experience rating. Insurers typically observes past driving records; these are highly informative on probabilities of accident, and, as such, are used for tarification. Omitting these variables will typically generate a bias, that tends precisely to overestimate the level of adverse selection: the corresponding information is treated by the econometrician as being private, whereas it is in fact common to both parties. However, the introduction of past experience is a quite delicate task, because it is (obviously) endogenous. Not only are panel data required, but endogeneity then raises specific (and delicate) econometric problems.

A final (and quite general) problem relates to the use of a highly constrained functional form. In the second equation, in particular, the relationship of the latent variable to the accident probability π and the price $\hat{g}_D$ is taken to be linear. This needs not be the case. To illustrate this point, Chiappori and Salanié (1996) consider the case of constant absolute risk aversion. Then the individual's choice of deductible is of the form:

$$D_i = \frac{1}{\sigma_i} \log \frac{1-\pi_i}{\pi_i} \frac{-\hat{g}_{Di}}{1+\hat{g}_{Di}}$$

which is highly nonlinear. They argue that, in fact, applying the Puelz-Snow procedure to data generated by a *symmetric* information model, according to this formula, may well result in the kind of negative estimates they get, simply because the accident term captures in fact some of the omitted nonlinearities.

A particularly elegant illustration of this fact is provided by Dionne, Gouriéroux and Vanasse (1998). Their idea is to first run an ordered probit on the "accident" variable, then to introduce the resulting *predictors* $\hat{\pi}_i$ of this ordered probit in the right-hand side of the second equation (for the choice of deductible), together with the dummy RT_i. They find that the $\hat{\pi}$ variable has a large and highly significant negative coefficient, while the RT variable is no longer significant. This, obviously, has nothing to do with adverse selection, as $\hat{\pi}_i$ is by construction a function of the *observed* variables only. If insureds have some private information, only new information contained in the agent's choice of contract, as summarized in RT, should be interpreted as an adverse selection measure. The result suggests, a contrario, that the negative influence of RT in the initial model can be spurious and due to misspecification.

☐ Correcting misspecifications

Several studies have attempted to correct these biases. Chiappori (1994) and Chiappori and Salanié (1996) propose a very general approach, that may potentially apply to most problems entailing adverse selection. The idea is to simultaneously estimate two (non linear) equations. One relates to the choice of the deductible. In the (simplest) case of a binomial decision, it takes the form

$$y_i = \mathbb{I}[f(X_i, \beta) + \varepsilon_i > 0] \qquad (1)$$

where, as above, the X_i are individual-specific exogenous variables, the β are parameters to be estimated, and ε_i is an econometric error term. Note that, contrarily to Puelz and Snow, the accident variable RT is *not* included in the right hand side. Nor is the premium; the idea, here, is that the latter is computed as a function of observables only, so that any information it conveys is already included in $f(X_i,\beta)$ – provided, of course, that the corresponding functional form is flexible enough.

The second equation takes the occurrence (and/or severity) of an accident as the dependent variable. In the simplest case, the latter is the dummy for the occurrence of an accident (our previous RT variable), and the equation takes the form:

$$RT_i = \mathbb{I}[g(X_i,\gamma) + \eta_i > 0] \qquad (2)$$

Note that this setting can easily be generalized. For instance, a recent contribution by Richaudeau (1997) takes into account the number of accident. Equation (2) is estimated using a count data model; the % are approximated by their "generalized residual" counterpart. In the same way, the distribution of accident costs (conditional on occurrence) can be introduced at that stage.

The key idea, then, is to simultaneously estimate the two equations, allowing for general correlation across the error terms. According to standard theory, asymmetric information should result in a positive correlation, under the convention that $y_i = 1$ (resp. $RT_i = 1$) corresponds to more comprehensive coverage (resp. the occurrence of an accident). One obvious advantage of this setting is that is does not require the estimation of the pricing policy followed by the firm, which is probably an extremely difficult task – and a potential source of important bias.

To circumvent the non linearity problems discussed above, as well as the issues raised by experience rating, Chiappori and Salanié consider a subsample of inexperienced drivers (which is equivalent to allowing each variable to interact with a young driver dummy); moreover, they introduce a large number of exogenous variables, allowing for crossed effects. They use both a parametric and a non parametric approach. The latter 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 adverse selection), 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.

This method can be given a fully general form. Following the presentation proposed by Dionne, Gouriéroux and Vanasse (1997) and Gouriéroux (1997), a general strategy can be summarized as follows. Let Y, X and Z respectively denote the endogenous variable under consideration (say, the occurrence of an accident), the initial exogenous variables and the decision variables at the agent's disposal (say, the choice of a particular contract within a given menu). Let $l(Y|X, Z)$ denote the probability distribution of Y conditional on X and Z. In the absence of adverse selection, the agent's choice conveys no information upon the endogenous variable. The translation is that:

$$l(Y|X, Z) = l(Y|X)$$

Obviously, this relationship can be given different, equivalent forms:

$$l(Z|X, Y) = l(Z|X)$$

or

$$l(Y, Z|X) = l(Y|X)\, l(Z|X)$$

(the latter version expressing the fact that, conditionally on X, Y and Z should be independent).

Interestingly enough, in all the empirical applications to automobile insurance just listed (with the exception of the initial paper by Puelz and Snow), independence is not rejected; in other words, these studies find no evidence of adverse selection. One remark must be stressed at this point. According to the previous arguments, the existence of a positive correlation across the residual cannot be interpreted as establishing the presence of asymmetric information without some precautions: as argued above, 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, that some empirical study does *not* reject the null (i.e., the absence of correlation). Although, in principle, this result might as well be due to a misspecification bias, this explanation is much less credible in that case; for it must be the case that, while (fully conditional) residual are actually positively correlated, there exists some bias that goes in the opposite direction with the *same* (absolute) magnitude – so that it exactly offsets the correlation.

□ Adverse selection versus moral hazard

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. Even in the static context, however, some papers have tried to disentangle the two types of asymmetries. In principle, any situation were some agents are, for exogenous reasons, faced with different incentive schemes can be used for testing for moral hazard. The problem, of course, is how to be sure that the differences in schemes are purely exogenous, and do not reflect some hidden characteristics of the agents. As an example, Chiappori and Salanié (1997) 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 probability. 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 allow to distinguish between the two types of effects. Chiappori and Salanié find evidence in favor of the second explanation: the accident rates of the "favored" young drivers are, other things equal, smaller than average by a small but significant percentage.

■ CONCLUSION

To conclude this brief overview, a few remarks are in order. First, a striking common feature of most empirical studies is their inability to detect any significant component of asymmetric information. This suggest that the corresponding problems, although systematically emphasized by the theory, may not be in fact systematically relevant. This conclusion, however, should not be pushed too far. For one thing, automobile insurance is but one particular field. In many other areas, adverse selection may well constitute a major problem; think, for instance, of unemployment insurance or the market for annuities, just to name a few. Secondly, the theoretical models remain extremely useful, in particular to predict the consequences of specific regulations. Indeed, a typical cause of adverse selection is the existence of specific rules that prohibit the use of particular variables[7]. In general, such regulations rely on the priors that discrimination based upon these variables is unethical or unfair, and should be suppressed. What theory suggests, however, is that they may well reveal counterproductive, to the extend that they replace explicit discrimination based upon observables by the indirect selection devices induced by competition in an adverse selection setting. Clearly, the importance and potential social cost of such perverse effects may not be trivial. But this is an empirical issue, for which more applied research is clearly needed.

☐ Notes

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

2. See Villeneuve (1996) or Chassagnon (1996), and Chassagnon and Chiappori (1997) for a theoretical investigation of the new equilibria.

3. Typically, they may require more instrument than in the standard models; in addition, one may have to introduce randomized contracts.

4. This needs not be true when loading is important and reflects cross-subsidies across contracts. Indeed, agents with lower risk will then typically prefer partial coverage, even at a (slightly) higher unitary price. Note, however, insurance companies are unlikely to charge a higher unitary price to less risky customers in any case.

5. The literature on repeated adverse selection clearly indicates that, while partial pooling may occur (especially in the initial stages), and although revelation mechanisms are much more complex, the positive correlation between the contract choice and expected cost is still present.

6. This is typically the case of insurance for young drivers, for instance.

7. To name a few examples: race, sex, age, ...

☐ References

BOYER, M., and G. DIONNE (1989), "An Empirical Analysis of Moral Hazard and Experience Rating", *Review of Economics and Statistics*, 71, 128-134.

CHASSAGNON, A. (1996), *Sélection adverse : modèle générique et applications*, PhD dissertation, DELTA.

CHASSAGNON, A. and P.A. CHIAPPORI (1997), *Insurance under Moral Hazard and Adverse Selection: The Case of Pure Competition*, Mimeo, DELTA.

CHIAPPORI, P.A. (1994), *Assurance et économétrie des contrats : quelques directions de recherche*, Mimeo, DELTA.

CHIAPPORI, P.A. and B. SALANIÉ (1996), "Empirical Contract Theory: The Case of Insurance Data", CREST DP9639, forthcoming in the *European Economic Review*.

CHIAPPORI, P.A. and B. SALANIÉ (1997), *Testing for Adverse Selection: the Case of Automobile Insurance Markets*, Mimeo, CREST.

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 Academic Publishers, 423-444.

DIONNE, G., ed. (1992), *Contributions to Insurance Economics*, Kluwer, Boston.

DIONNE, G., C. GOURIÉROUX and C. VANASSE (1997), *The Informational Content of Household Decisions, with an Application to Insurance under Adverse Selection*, CREST DP9701.

DIONNE, G., C. GOURIÉROUX and C. VANASSE (1998), *Evidence of Adverse Selection in Automobile Insurance Markets*, in Automobile Insurance: Road Safety, New Drivers, Risks, Insurance Fraud and Regulation, edited by Dionne G. and C. Laberge Nadeau, Kluwer Academic Press, Boston 13-46.

GOURIÉROUX, C. (1997), *The Econometrics of Insurance*, Invited Geneva Association Lecture, Conference on Risk and Uncertainty, Paris.

HELLWIG, M. (1987), "Some Recent Developments in the Theory of Competition in Markets with Adverse Selection", *European Economic Review*, 31, 154-163.

PUELZ, R. and A. SNOW (1994), "Evidence of Adverse Selection: Equilibrium Signalling and Cross-Subsidization in the Insurance Market", *Journal of Political Economy*, 102, 236-257.

RICHAUDEAU, D. (1997), *Contrat d'assurance automobile et risque routier : analyse théorique et empirique sur données individuelles françaises 1991-1995*, thèse de doctorat, Université de Paris I Pantheon-Sorbonne, 331 pages.

VILLENEUVE, B. (1996), *Essais en économie de l'assurance*, PhD dissertation, DELTA-CREST.