The Empirical Measure of Information Problems with Emphasis on Insurance Fraud

by Georges Dionne

Working Paper 00-04
March 2000
ISSN : 1206-3304

The author would like to thank Marie-Gloriose Ingabire for her help with bibliographical research, FCAR-Quebec and CRSH-Canada for their financial support. He would also like to mention the researchers who helped him develop several of the ideas on the subject over the years: P.A. Chiappori, K. Dachraoui, N. Doherty, C. Fluet, N. Fombaron, R. Gagné, C. Gouriéroux, P. Lasserre, P. Picard, B. Salanié, P. St-Michel, C. Vanasse, P. Viala. He thanks P. Lanoie, C. Fluet and B. Villeneuve who proposed interesting improvements in the first version of this article.

Electronic versions (pdf files) of Working papers are available on our Web site: http://www.hec.ca/gestiondesrisques/papers.html.
The Empirical Measure of Information Problems with Emphasis on Insurance Fraud

Georges Dionne

Georges Dionne holds the Risk Management Chair and is professor of finance at École des HEC.
The Empirical Measure of Information Problems with Emphasis on Insurance Fraud

Abstract

We discuss the difficult question of measuring the effects of asymmetric information problems on resource allocation. Two of them are retained: moral hazard and adverse selection. One theoretical conclusion, shared by many authors, is that information problems may introduce significant distortions into the economy. However, we can verify, in different markets, that efficient mechanisms have been introduced in order to reduce these distortions and even eliminate, at the margin, some residual information problems. This conclusion is stronger for adverse selection. One explanation is that adverse selection is related to exogenous characteristics while moral hazard is due to endogenous actions that may change at any point in time.

Keywords: Empirical measure, information problem, moral hazard, adverse selection, insurance fraud.

JEL numbers: D80, G22, C25, G11.

Résumé

Nous abordons la difficile question de la mesure empirique des effets des problèmes d'information sur l'allocation des ressources. Deux problèmes retiennent notre attention; le risque moral et l'antisélection. Une conclusion, acceptée par la plupart des auteurs, est que les problèmes d'information créent des distorsions importantes dans l'économie. Mais nous pouvons vérifier, dans certains marchés, que des mécanismes efficaces ont été mis en place pour réduire ces distorsions et même éliminer, à la marge, des problèmes résiduels d'information. Cette conclusion semble plus forte avec l'antisélection qu'avec le risque moral. Une explication est le fait que l'antisélection concerne des caractéristiques exogènes, alors que le risque moral est expliqué par des actions endogènes qui peuvent être modifiées en tout temps.

Mots clés : Mesure empirique, problème d’information, risque moral, antisélection, fraude à l’assurance.

Classification JEL : D80, G22, C25, G11.
1. Introduction

The study of information problems in economics began in the early 1960s. The two best known problems, moral hazard and adverse selection, were introduced into the literature in 1963 by Kenneth Arrow in a classic article published in the American Economic Review. In 1970, Akerlof came up with the first analysis of a market equilibrium in the presence on adverse selection. Optimal contracts were first characterized for adverse selection in articles by Pauly (1974), Rothschild and Stiglitz (1976), and Wilson (1977), and for ex-ante moral hazard by Holmstrom (1979) and Shavell (1979). Even if the problem of ex-post moral hazard was defined early on by Pauly (1968), it was later formalized by Townsend (1979) and Gale and Hellwig (1985).

In the early 1980s, several theoretical developments were advanced to account for different facts observed in several markets. Specifically, dealing only with models of two-party contracts, multi-period contractual relations were introduced; the renegotiation of contracts was formalized; the problem of contractual commitments was analyzed; and simultaneous treatment of several problems became a consideration. Other noteworthy proposals were developed to explain hierarchical relations in firms and in organizations (see the references).

The contracts most often studied are insurance contracts, banking contracts, work and sharecropping contracts, and types of auctions, etc. Several forms of contracts observed in these different markets were catalogued in various theoretical contributions. The best known are partial insurance coverage (co-insurance and deductibles), compensation based on hours worked and performance, compensation of executives with stock purchase options, debt, bonus-malus, temporal deductibles, venture capital contracts with warrants, etc. There was also rationalization of several corporate organizational practices such as the use of foremen, internal and external controls, decentralization of certain decisions, and the centralization of more difficult-to-control decisions.

The empirical study of information problems began much later. The main motivation was to distinguish the stylized (qualitative) facts used to construct certain models from real or more quantitative facts. For example, in classroom and theoretical journals, different automobile insurance deductibles can very well be used to justify adverse selection, but there is no evidence that insurers established this partial coverage for that reason. It can also be argued that labor contracts with performance compensation are used to reduce moral hazard in firms, but it has not necessarily been empirically demonstrated that there is less moral hazard in firms using this form of compensation than in other firms that use fixed compensation but set up other incentives or other control mechanisms to deal with the information problem.

Another strong motivation for empirically verifying the effects of information problems is the search for ways to reduce their negative impact on resource allocation. For example, we know that partial insurance is effective in reducing ex-ante moral hazard, as it exposes the insured person to risk. On the other hand, this mechanism is not effective against ex-post moral hazard, as the accident has already occurred. Partial insurance may even have pernicious effects and encourage the padding of costs (Dionne and Gagné,
The audit is the most effective instrument against ex-post moral hazard. This shows the importance of identifying the real problem when attempting to correct imperfections.

When it comes to empirically measuring information problems and assessing the effectiveness of mechanisms set up to correct them (relationship between the nature of contracts and their performance), a number of complications soon arise. For one thing, several information problems may be present, simultaneously, in the data base studied; the theoretical predictions must then be carefully defined so as to distinguish the effects of different information problems on the parameters of the contracts to be estimated. Moreover, firms have a whole range of mechanisms (substitutes or complementary) at their disposal and they may be selected for reasons other than information problems or for information problems other than the ones to be taken up in a particular study. In other words, the information problems under consideration are often neither a necessary nor a sufficient condition to justify the existence of certain mechanisms.

Treating several information problems simultaneously is difficult, as the literature does not offer many theoretical predictions, even when available range of contributions is reviewed. But if we simply limit ourselves to verify whether a market contains any residual information asymmetry, regardless of its origin, it is easier to demonstrate its absence, since there is no need to distinguish between the different forms of information asymmetry. Otherwise, we have to ascertain which form is still present and document its cause in order to analyze the instruments which could mitigate or eliminate it.

As a rule, the distinction between moral hazard and adverse selection can be brought down to a problem of causality (Chiappori, 1994, 2000). With moral hazard, the non-observable actions of individuals that affect the way contracts work are consequences of the forms of contracts. For example, a contract may increases the risk of the activity, because it reduces the incentives to act with prudence. With pure adverse selection, the nature of different risks already exists before the contract is written. The contracts selected will flow from the risks present. There is thus a form of reverse causality between the two information problems. When an exogenous change occurs in an insurance contract, we can limit our test to the way it affects existing policy holders and isolate a moral hazard effect. Or, we could make comparisons to see whether the chance of catastrophe differs between new and old policy holders and check for any bias caused by adverse selection.

Another difficulty in the empirical measurement of information problems is the fact that researchers are not privy to any more information than decision makers. Two solutions have been adopted to make up for that difficulty: (1) use of confidential polls and (2) development of econometric strategies capable of isolating the desired effect. The experimental approach is a third avenue that I shall not deal with in detail (see, however, Section 4 for an example).

The polling method has the advantage of providing direct access to private information not available to the other parties to the contracts. Such information makes it possible to measure directly motivations for choosing specific contractual clauses as well as the
behaviour of agents. The drawback of this method is that it is very costly. It can also be biased, because it is very difficult to explain all the complexity of the problem studied to respondents and because several alternative explanations might have been overlooked in the questionnaires.

The development of econometric strategies requires a good knowledge of the theoretical problem under study and of the econometric methods suitable to the project. This is why the most productive research teams are composed of theoreticians and econometricians. The objective is to isolate effects that are not directly observable by both parties to the contract but which are taken into account by certain variables or combination of variables. As discussed by Chiappori (1994), econometric work consists in distinguishing between two types of information. The first type is composed of variables observable by the two parties to the contract. These variables can be used to make estimates conditional on the characteristics observed. The second type is linked to that which is not observable by econometricians (and by at least one contractual party), but which may explain choices of contracts or behaviours. In the case of adverse selection, choices of contract can be interpreted by econometricians as being a bias of endogenous selection. One way of taking this into account is to estimate simultaneously the decisions of agents by introducing hidden connections (or informational asymmetries) between the decisions. One known form is the non-null correlation between the random terms of the different equations (Chiappori and Salanié, 1997).

Quality of data is a determining factor in the measurement of desired effects. The data must correspond directly to the contractual relations studied and to the duration of the contractual periods. There must also be access to data broken down contract by contract. The work of formulating raw data for the purposes of research should not be underestimated. Raw data are used in the day-to-day operations of firms which are not concerned with research problems and do not always contain the direct information on variables needed for the problem studied.

Econometric specifications must correspond to the theoretical models under consideration, if erroneous conclusions are to be avoided. Often, we choose (or are forced) to use only part of the information available to decision-makers, and thus bias the effects of certain variables so that they capture the effects of other forgotten or inaccessible variables.

Finally, the agents party to different contracts are often risk averse and display different levels of such aversion. This last characteristic is also difficult to observe and can itself be a source of asymmetric information. Some authors have recently proposed models taking into account the varying degrees of aversion to risk, but there are very few predictions capable of isolating the effects of information problems as they relate to varying degrees of risk aversion among agents (see Dionne, Doherty and Fombaron, 2000, for a longer discussion in relation to adverse selection).

The rest of my exposé will take up examples of the empirical verification of the presence or absence of a residual information problem in a market. These examples highlight the various difficulties which are not always well understood by those who tackle the
empirical measurement of information problems. The first is a test for the presence of adverse selection in the portfolio of a private insurer. The question to ask is the following: Are the choices of deductibles explained by this information problem or not?

The second example deals with labor contracts and methods of compensation. Methods of compensation are often observable by econometricians, whereas individual effort is not. Furthermore, individual output can hardly be used to deduce effort, because it depends on several other factors, such as the outcome of a random variable or other non-observable staffing practices.

We next treat ex-post moral hazard in markets covering work accidents and medical services. The main difficulty is assigning variations in demand to price effects, moral hazard, and adverse selection. Many studies show that a change in coverage will affect consumption, but few are capable of determining whether the cause is a problem of moral hazard, for example. A section on insurance fraud will also be presented. We will see how parameters of standard insurance contracts may affect incentives to defraud.

Finally, we shall discuss market equilibrium in reference to adverse selection in markets for used cars. Can the price differences observed for the same quality be explained by adverse selection?

2. Measurement of Residual Adverse Selection in the Portfolio of an Insurer

Adverse selection has been dealt with in several theoretical essays (for example, see Dionne, Doherty and Fombaron, 2000). In this section, we limit ourselves to insurance contracts. Two mechanisms have been proposed in the literature to account for this resource allocation problem: deductibles and classification of risks. The two are complementary and the empirical questions with which we are concerned are the following:

Does the effective use of risk classification suffice to account for this information problem?

Or :

Do we need additional self-selection mechanisms? In other words, is there any residual adverse selection in classes of risk that justify the use of deductibles?

Before answering these questions, we should summarize the relevant theoretical contributions associated with them. Crocker and Snow (1985, 1986, 2000) proposed models showing that the classification of risks does improve the welfare of all individuals if two conditions are respected. The variables used to evaluate the individual risks must be easily observable (or observable at low cost). They must also be correlated with the individual risks.

We can easily certify that most of the variables involved in the classification of risks for automobile insurance contracts are easily observed by insurers. To check the second condition, we need to estimate individual frequencies of accidents in terms of these same
variables of ratemaking. This is why it is so important to have high quality data on an insurer's portfolio.

The next step is to check whether deductibles in different classes of risk, are chosen in terms of individual risks. The model constructed by Rothschild and Stiglitz (1976) and Wilson (1977) predicts that high risks will choose lower deductible than low risks. Puelz and Snow (1994) used accidents at the end of the contractual period to approximate individual risks. They found that those who were the most accident prone chose the lowest deductible.

This finding is not convincing, for it is subject to an econometric specification error. The authors estimated two equations: one equation dealing with insurance pricing and the other equation dealing with choice of deductible. They used the second equation to test for the presence of adverse selection. As their choice-of-deductible equation contained only a few explanatory variables, the coefficient of the 'accidents' variable may capture information other than that related to residual adverse selection.

The standard method for correcting this specification problem is to introduce the mathematical expectation of the number of accidents (or its predicted value obtained from the estimates of the accidents distribution) in the choice-of-deductible equation (Dionne, Gouriéroux, and Vanasse, 1998; or Chiappori and Salanié, 1997, for an equivalent approach; see also Section 5 of this chapter for more details). In doing this second regression, we check to see if the accident variable is still significant. If not, this means that there is no residual information in the risk classes. If the predicted variable is significant and bears the same sign as the accident variable in the first step, we cannot conclude that it measures adverse selection, since its prediction was obtained with variables observable by the insurer. The fact that it is significant is usually due to non-linearities not modeled in the equation. These non-linearities can be eliminated by increasing the interactions between variables in the choice-of-deductible equation, as do insurers when setting their premia.

Finally, we may conclude that there is a residual information problem in the portfolio when there is still a statistical link between the deductible variable and the accident variable in a model well specified. For example, the presence of residual adverse selection might have prevented the standard econometric specification method from completely correcting the problem. But the true residual information problem may be other than adverse selection. Other tests are necessary to isolate the true information problem.

There are numerous lessons to be drawn from this example. On this point, the theoretical environment has been well documented. The theoretical predictions of Rothschild and Stiglitz have had currency for more than twenty years and have been taught in microeconomics courses for a good many years. Some authors before Puelz and Snow had proposed tests for the theory, but the data used were not always adequate and often too aggregate.
Puelz and Snow had access to a good quality data base. They rather successfully isolated the relevant empirical questions, but they did not consider all of the instruments an insurer could use to take adverse selections into account effectively. Moreover, they failed to correctly interpret their econometric results and, most unfortunately, they never suspected that their conclusion on the residual adverse selection measured in the portfolio might be the result of an econometric specification.

This does not mean that there is no adverse selection in the automobile insurance markets. The fact that insurers classify risks is in large part explained by adverse selection. But, the absence of residual information asymmetry in the classes of risk shows that, when this classification is correctly done, the choice of deductible is not needed to treat adverse selection. In other words, the Rothschild and Stiglitz (1976) model is not useful in this portfolio.

Others will want to point out that moral hazard may also be present in this portfolio and that we probably did not screen for all the factors capable of explaining how deductibles relate to the differing degrees of policy holders’ risk aversion. The second criticism is easier to handle. Let’s start with that one.

Though rare, works treating differences in aversion to risk conclude that, in cases of adverse selection, good risks, who have a stronger risk aversion, may ask for coverage other than the one imposed by the self-selection constraint of the high risk (more expensive than actuarial, better than that of the good-to-weak aversion category, but still partial) (Villeneuve, 1996 and Smart, 1998).

Risk aversion cannot be directly observed. To screen for it, as in Puelz and Snow, we (Dionne, Gouriéroux, Vanasse, 1998) used the amounts of insurance coverage chosen by individuals as protection against potential civil liability losses. Some of these variables are significant and of the right sign when we calculate the choice-of-deductible equation for damages to the car combined with the predicted accident variable. But we also show that it is possible to make these variables non-significant by increasing the number of variables and the number of interactions between the variables insurers use in setting their rates. This finding implies that the methods for classifying policy holders can take into account not only the differences between individual risks but also the differences in risk aversion.

To adequately account for moral hazard in insurance contracts along with adverse selection, we must have access to a model capable of making theoretical predictions in an environment where the two information problems are simultaneously present. This exercise was dealt with by Chassagnon and Chiappori (1995) in a competitive market context. They found that agents who are less worried about protection choose contracts with the broadest coverage and the lowest deductibles (see also Dionne and Lasserre, 1987).

If we are limited to static contracts with data covering just one period, it is difficult to ascertain where the causality of moral hazard and adverse selection is heading. Panel data and experiments can help define the two information problems. The data of Dionne,
Gouriéroux and Vanasse (1998) contained information on the bonus-malus of the company's clients. This information can here be considered as another good instrument for taking moral hazard into account. The preliminary results show that use of these variables have no impact on conclusions concerning the presence of residual adverse selection (see also Dionne and Gagné, 2000, for discrimination between information problems. We shall come back on this contribution in Section 5).

Another test concerns the type of commitment we may observe in dynamics insurance contracts. Dionne and Doherty (1994) proposed such a test by analyzing the variation of insurers loss premium ratio as a function of the premium rate. They verified that some automobile insurers use commitment to attract selective portfolios with disproportionate numbers of low risks. These results are consistent with the commitment and renegotiation model and reject both the no-commitment and the full commitment models. However, we must emphasize that these preliminary results represent an indirect test of the theory since the authors did not have access to the more accurate data. As mentioned by the authors, a direct test would require that data on different risk groups or cohorts be available as well data on the insurance prices faced by the different cohorts over time.

3. Ex-Ante Moral Hazard and Choices of Work Contracts

There is, by definition, ex-ante moral hazard if one of the parties to a contract can affect the results of the contractual relation by non-observable actions before realization of the random variables (Holmstrom, 1979; Shavell, 1979) (see Arnott, 1992, and Winter, 2000, for reviews of the insurance literature with moral hazard). In the simple model that we shall now treat, the realized output is observable but we do not know whether its value is due to the agent's effort or to the outcome of a random variable. We thus have a problem of identification to solve, if we want to check for the presence of residual moral hazard. (For other applications, see Dionne, Gagné, Gagnon and Vanasse, 1997, and Dionne and Vanasse, 1997.)

One useful prediction that models with moral hazard have made for the labour market is that forms of compensation can have an impact on work incentive: a worker paid based on performance should work harder than a worker paid an hourly wage. In other words, there should be less moral hazard when workers are paid based on performance, since their compensation is exposed to risks whose impact they can vary by their efforts.

Empirically, the hardest factor to measure in the model is the worker's effort, as this means gaining access to a variable the employer cannot observe and which can still be used to see whether methods of compensation have any impact on effort. Foster and Rosenzweig (1994) used calories consumed by workers as an approximation of the effort they expend.

They propose a simple theoretical model of workers' health in which body mass (kg/square meter) is affected by food intake, illness, and work effort. They show that it is possible, for the types of jobs studied, to make a direct connection between forms of compensation and the calories consumed. More specifically, in periods where workers have access to methods of compensation that reward more high powered performance,
they work harder and consume more calories, thus justifying the direct theoretical link
between method of compensation and consumption of calories.

To test their model, they used panel data containing information on 448 farming families
in the Philippines; the members of these families may work either for themselves or for
outsiders, under different forms of compensation. These individuals were interviewed
four times concerning their wages, their modes of compensation, the type of work done,
and the quantity of calories consumed over the previous 24 hours. A period of four
months separated the interviews.

The results from estimation of the health function indicate that self-employment and
piece work significantly reduce the body mass index as compared with unemployment,
whereas work compensated on an hourly basis shows no significant effect. This seems to
indicate either less effort or a measurable presence of moral hazard on the part of those
who are paid with an hourly rate.

Now, what about the link between methods of payment and the performance rate per
calorie consumed? They found that the calories consumed are associated with higher pay
and performance in self-employment and piece work. Consequently, workers receiving
these modes of payment consume more calories and, thus, can be said to work harder.

The next important question we must ask is the following one: Is this a test for moral
hazard or for adverse selection? In other words, do workers themselves choose their type
of work and mode of compensation?

The authors tried to answer this question by checking to see whether their data contained
any sample selection effect. They used two methods to do this: Heckman's two-step
Probit selection (1979) and Lee's multinomial Logit selection (1983). Both models render
identical results.

It should be pointed out that 47.1% of the subjects worked under different regimes during
the same period. But this statistic does not suffice to qualify the choices as random, since
only 28% worked for hourly wages in all four periods.

Taking explicitly into account workers' choices of types of compensation tends to
strengthen rather than weaken the results. Modes of compensation actually have a bigger
impact on the use of calories with the selection model. This implies that those who
choose incentive pay at the margin do so because they truly want to work harder. But,
unlike what the authors suggest, the model tested is not a pure moral-hazard model. It is
rather a mixed model containing aspects of adverse selection and moral hazard. The best
physically endowed and most highly motivated will choose the highest paying but most
demanding work.

In fact, to isolate a pure moral-hazard effect, it practically takes an exogenous change in a
compensation regime or in some other parameter impinging on all the agents. We are
now going to study changes of this nature as we turn to ex-post moral hazard.
4. Ex-Post Moral Hazard, Demand for Medical Services, and Duration of Work Leaves

In our applications, ex-post moral hazard deals with non-observable actions on the part of agents, actions which occur during or after the outcome of the random variable or accident (Townsend, 1979, and Gale and Hellwig, 1985). For example, an accident can be falsified to obtain better insurance compensation. This form of moral hazard is often associated with fraud or falsification (Crocker and Morgan, 1998; Crocker and Tennyson, 1998; Bujold, Dionne, and Gagné, 1997; Picard, 2000). Partial insurance of agents is not optimal in reducing this form of moral hazard, for the agent knows the state of the world when he makes his decision. Claims auditing is more appropriate, but it is costly, resulting in the potential presence of this moral hazard in different markets.

The main difficulty in isolating the ex-post moral hazard effect in different levels of insurance coverage is separating the effects of price and income variations from the effects of asymmetric information. Contrary to what is often read in the literature, not every variation in consumption following upon a variation in insurance coverage can be tied to ex-post moral hazard. When compared with full-coverage regimes, it is perfectly conceivable that a health insurance regime with partial coverage might be explained by transaction costs and patients' decision to curtail consumption of certain services because they must share in the cost. If for some reason, the transaction costs drop and the insurance coverage expands, the consumption of medical services will increase, since their price will be cheaper. But this increase will not be due to moral hazard. It will simply be a classic effect of demand. There are still too many articles in the literature which confuse variations in demand with moral hazard.

Another big difficulty in isolating moral hazard is linked to the possibility that potential policy holders, better informed than the insurer about the state of their health over the next period of the contract, will make an endogenous choice of insurance regime. As a rule, those expecting health problems choose more generous insurance regimes, even if the per unit cost is higher. This is a well-known adverse selection effect.

In the famous Rand corporation study (Manning et al., 1987 and Newhouse, 1987) dealing with the effects of changes in insurance coverage on the demand for medical services, the experimental method used was capable of isolating the elasticity of the demand from the effects of adverse selection by random selection of families who might be subject to exogenous changes in insurance coverage but who were not free to choose their insurance coverage ex-ante. They thus successfully calculated elasticities of demand much lower than those obtained in other studies that did not screen for the effect of endogenous choices of insurance regimes (adverse selection).

Their measurement of the elasticity of demand for medical services is not a measurement of ex-post moral hazard. It is, in fact, very unlikely that there is any moral hazard in their data, considering all the screening done.

Let us now consider work accidents. As we indicated above, using an exogenous change in an insurance regime can isolate moral hazard. An exogenous change in an insurance
regime can be interpreted as a laboratory experiment, if certain conditions are met. As for laboratory animals, it is possible to restrict the choices of insurance available to the subjects.

It is also important to have a control group which undergoes the same insurance changes, but which does not have the same information problems as those expected. For example, if we suspect that some workers with specific medical diagnoses (hard to diagnose and verify) have greater information asymmetry with the insurer, there have to be other workers having undergone the same insurance changes at the same time but whose information asymmetry is weaker (easy to diagnose and verify). The reason for this is that it is hard to isolate an absolute effect with real economic data, because other factors not screened for may lead to changes in behaviour. The control group allows us to isolate a relative effect arising from the information problem, all things being equal. To simplify the analysis, it is preferable that the period under study should be short enough to get around having to screen for several changes at once.

Dionne and St-Michel (1991) managed to bring together all these conditions in a study of change in coverage for salary losses associated with work accidents (see B. Fortin and P. Lanoie, 1992, 1995, 2000, for similar studies and for a survey of different issues associated to workers compensation). The change in insurance coverage studied was exogenous for all the workers. Other forms of insurance were not really available, even if, in theory, it is always possible to buy extra insurance in the private sector if one is not satisfied with the public regime. But very few individuals do so in Quebec for this type of compensation. The fact that there are state monopolies over several types of insurance coverage in Quebec, makes it easier for us to meet this condition.

Dionne and St-Michel (1991) showed, first of all, that the increase in insurance coverage had a significant positive effect on the duration of absence from work. But this effect cannot be interpreted as being moral hazard, for it may simply be associated with an increase in the demand for days off due to their lower cost. Next, the authors checked to see whether this effect was only significant for diagnoses with greater asymmetry of information (hard to diagnose) between the worker and the insurer as represented by a doctor. This second finding confirms that the only effect observed on the duration of absences was that of moral hazard, since the workers of the control group (those without information asymmetry, easy to diagnose) did not modify their behaviour. Moreover, the change-of-regime variable without interaction with diagnostics, is no longer significant when the diagnostic-change-of-insurance variables are adjusted. This implies that there is no demand effect. However, the change of regime achieved the desired redistribution effects sought after by allowing poorer workers to have access to more insurance.

We may conclude that an ex-post moral hazard effect has been isolated (see Cummins and Tennyson, 1996, Butler et al., 1996a, Ruser, 1998, and Dionne, St-Michel and Vanasse, 1995, for similar results). It is, in fact, highly unlikely that the change in regime studied had any impact on ex-ante prevention activities which might affect the seriousness of work accidents. There is no reason to think that the average worker can practice such selective prevention as to influence diagnostics ex-ante. But, ex-post, when he knows his diagnosis, he can take undue advantage of the situation of asymmetric
information. Some workers might be more tempted to provoke accidents or to say falsely that they had an accident in order to have access to more compensation, when the rates are more generous. These activities were not distinguished from other forms of moral hazard by Dionne and St-Michel, since they can be interpreted as ex-post moral hazard.

It is also difficult to find the link between this result and adverse selection. On the one hand, workers could not choose their insurance coverage in this market and, on the other hand, it is highly unlikely that the change in insurance regime had any short-term effect on workers' choice of more or less risky jobs.

Bernard Fortin and Paul Lanoie (2000) present a review of the literature on the incentive effects of work accident compensation. They use the classification of different forms of moral hazard proposed by Viscusi (1992). The form of ex-post moral hazard we just described can be classified moral hazard as duration of claims, which they distinguish from moral hazard as substitution hazard. This distinction can be explained, for example, by the fact that compensation for work accidents are more generous than those for unemployment insurance. Activities resulting in accidents are called causality moral hazard, which is ex-post moral hazard (bordering on ex-ante moral hazard), since the action takes place at the time of the accident. The result obtained by Dionne and St-Michel captures these three forms of ex-post moral hazard. It is even possible that workers may have substituted workers’ compensations for unemployment insurance.

Can we now perform closer analysis and distinguish between the three forms of ex-post moral hazard: incentives provoking hard-to-verify accidents; decisions to prolong length of absence in hard-to-check diagnoses; or decisions to substitute accident compensations for unemployment insurance, or even falsification? This distinction would be important as it is not obvious that the mechanisms for correcting the situation would be the same for each of these forms of asymmetric information.

The last three forms are difficult to distinguish, since they belong to the same market. However, it is possible to separate new accidents from older ones using indicative variables. We know, for example, that the accidents provoked occur early on Monday mornings (see also Fortin and Lanoie, 1998, and Derrig, 1997) and that, among seasonal workers, requests to extend work absences pick up with the approach of unemployment insurance periods. Further research must be done on this subject.

5. Insurance Fraud

Insurance fraud has become an important economic problem. In the Québec automobile insurance market, the cost of fraud was estimated at $100 million in 1994, just under 10% of total claims (Caron and Dionne, 1997). The Insurance Bureau of Canada has estimated that the total annual cost of liability insurance fraud was about $2 billion in Canada (Medza, 1998), while it is estimated to be nearly $70 billion per year in the United States for all types of claims (Foppert, 1994).

The causes of the rapid growth of insurance fraud are numerous: changes in morality, increased poverty, modifications in the behaviour of the intermediaries (medical doctors
or mechanics for instance), attitude of insurers, etc. (Dionne, Gibbens and St-Michel, 1993). In two papers, Dionne and Gagné (1999, 2000) highlight the nature of insurance contracts. In both cases, they use the theoretical model proposed by Picard (1996) to obtain an equilibrium without commitment of the parties. In the second one (2000), they test whether the presence of a replacement cost endorsement can be a cause of fraudulent claims for automobile theft. This endorsement was introduced in the automobile insurance market to increase the protection of the insureds against depreciation.

Traditional insurance markets do not offer protection against the replacement value of an automobile. Rather, they cover current market value, and when a theft occurs, the insurance coverage is largely partial with respect to the market value of a new automobile. A replacement cost endorsement gives the opportunity to get a new vehicle in the case of theft or in the case of total destruction of the car in a collision, usually if the theft or the collision occurs in the first two years of ownership of a new automobile. In case of total theft, there is no deductible. Ex-ante and without asymmetric information, this type of contract can be optimal. The only major difference is the expected coverage cost which can easily be reflected in the insurance premium.

Intuitively, a replacement cost endorsement may decrease the incentives toward self-protection since it can be interpreted as more than full insurance when the market value of the insured car is lower than the market value of a new car. The presence of a replacement cost endorsement in the insurance contract may also increase the incentives to defraud for the same reason. For example, the insured may have an incentive to set up a fraudulent theft because of the additional protection given by the replacement cost endorsement. This particular type of fraud is known as opportunistic fraud since it occurs when an opportunity occurs and usually not when an insurance contract for a new vehicle is signed. Alternatively, under adverse selection, an individual may choose to include in his coverage a replacement cost endorsement because he knows he will be more at risk.

A first objective of the study by Dionne and Gagné (2000) was to test how the introduction of a replacement cost endorsement affects the distribution of thefts in the automobile insurance market. Another significant objective was to propose an empirical procedure allowing the distinction between the two forms of moral hazard. In other words, they seek to determine whether an increase in the probability of theft may be explained by a decrease in self-protection activities or by an increase in opportunistic fraud. They also took into account the adverse selection possibility since the insured ex-ante decision to add a replacement cost endorsement to the insurance policy might be explained by unobservable characteristics that also explain higher risks.

As discussed in Section 2, Dionne, Gouriéroux and Vanasse (1998) proposed a method that was applied to adverse selection. In their article, Dionne and Gagné (2000) extend this method in order to take into account both forms of moral hazard simultaneously. Furthermore, their approach makes it possible to isolate adverse selection.

Let us first consider \( y \), an endogeneous binary variable indicating the occurrence of a theft. The decision or contract choice variable \( z \) (in this case the presence of a replacement cost endorsement; in Section 2, the choice of a particular deductible) will
provide no additional information on the distribution of \( y \) if the prediction of \( y \) based on \( z \) and other initial exogenous variables \( x \) coincides with that based on \( x \) alone. Under this condition, we can write the conditional distribution of \( y \) as

\[
\phi_y(y \mid x, z) = \phi_y(y \mid x),
\]

where \( \phi(\bullet \mid \bullet) \) denotes a conditional probability density function. A more appropriate but equivalent form for different applications is

\[
\phi_z(z \mid x, y) = \phi_z(z \mid x).
\]

In that case, the distribution of \( z \) is estimated and when condition (2) holds, this distribution is independent of \( y \) which means that the distribution of theft is independent of the decision variable \( z \), here the replacement cost endorsement, since (1) and (2) are equivalent. Their empirical investigation relies on the indirect characterization as defined by (2). It can be interpreted as the description of how the individual’s decision affects his future risks (moral hazard) or of what his decision would be if he knew his future risks (adverse selection).

This type of conditional dependence analysis is usually performed in a parametric framework where the model is a priori constrained by a linear function of \( x \) and \( y \), that is

\[
\phi_z(z \mid x, y) = \phi_z(z \mid x' a + by).
\]

This practice may induce spurious conclusions, since it is difficult to distinguish between the informational content of a decision variable and an omitted nonlinear effect of the initial exogenous variables. A simple and pragmatic way of taking into account these potential nonlinear effects of \( x \) is to consider a more general form

\[
\phi_z(z \mid x, y) = \phi_z(z \mid x' a + by + cE(y \mid x)),
\]

where \( E(y \mid x) \) is an approximated regressor of the expected value of \( y \) computed from the initial exogenous information. Assuming normality, \( E(y \mid x) \) is computed with the parameters obtained from the estimation of \( y \) using the \textit{Probit} method.

The above framework can be applied to test for different types of information asymmetries. The failure of condition (2) to hold may allow a distinction between different types of information problems depending on how \( y \) is defined. Dionne and Gagné (2000) defined \( y \) using 5 different contexts or sub-samples \((s)\):

- \( s = 0 \) when no theft occurred;
- \( s = 1 \) if a partial theft occurred at the beginning of the cost endorsement contract;
- \( s = 2 \) if a partial theft occurred near the end of the cost endorsement contract;
- \( s = 3 \) if a total theft occurred at the beginning of the cost endorsement contract;
- \( s = 4 \) if a total theft occurred near the end of the cost endorsement contract.
Using such a categorization, they identified the different types of information problems: adverse selection, ex-ante moral hazard and ex-post moral hazard or opportunistic fraud.

If we are in presence of a pure adverse selection effect, the time dimension (that is, the proximity of the expiration of the replacement cost endorsement in the contract, since it is valid for only two years after buying a new car) would not have any importance. In other words, the effect of pure adverse selection would be significant and of approximately the same size whether it is a new contract or an old one. However, the effects may not be of the same magnitude. Therefore, with a pure adverse selection effect, condition (2) should not hold in all sub-samples considered (i.e. \( s = 1, 2, 3 \) and 4).

Assuming that the same self-protection activities are involved in the reduction of the probabilities of both types of theft (partial and total), condition (2) should not hold under ex-ante moral hazard for both types of theft. In that case, the presence of a replacement cost endorsement in the insurance contract reduces self-protection activities leading to an increase in the probabilities of partial and total theft. In addition, since the benefits of prevention are decreasing over time, ex-ante moral hazard increases over time. Thus, as for adverse selection, ex-ante moral hazard implies that condition (2) does not hold in all sub-samples considered, but with a stronger effect near the end of the contract (i.e. sub-samples 2 and 4) than at the beginning (i.e. sub-samples 1 and 3).

In the case of opportunistic fraud, the pattern of effects is different. Because the incentives to defraud are very small or even nil in the case of a partial theft, condition (2) should hold in both sub-samples 1 and 2. Also, because the benefits of fraud for total theft are small at the beginning of the contract but increasing over time with a replacement cost endorsement, condition (2) should also hold in the case of a total theft at the beginning of the contract (\( s = 3 \)). However, near the end the contract, the incentives to defraud reach a maximum only in the case of a total theft when the insurance contract includes a replacement cost endorsement. It follows that with a fraud effect, condition (2) would not be verified in sub-sample 4.

Their empirical results show that the total theft occurrence is a significant factor in the explanation of the presence of a replacement cost endorsement in an automobile insurance contract only when this endorsement is about to expire. The total theft occurrence is not a significant factor neither at the beginning of the contract, nor at a middle stage.

As suggested by Chiappori (1998), one possibility to obtain separation from claim data is to use a dynamic model. The data of Dionne and Gagné (2000) did not allow them to go in that direction. The originality of their methodology, although in the spirit of Chiappori (1998), was to use different contracting dates for the replacement cost endorsement but claims over one period. Consequently, Dionne and Gagné (2000) were first able to separate moral hazard from adverse selection since the latter should have the same effect at each period according to the theory. Finally, they were able to separate between the two forms of moral hazard by using partial and total thefts and by assuming that the same preventive actions affect both distributions. Their results do not reject the presence of
opportunistic fraud in the data which means that the studied endorsement has a direct significant effect on the total number of car thefts in the analyzed market.

In their 1997 article, Dionne and Gagné discuss the effect of higher deductible on the costs of claims explained by falsification. Since the significant contribution of Townsend (1979), an insurance contract with a deductible is described as an optimal contract in the presence of costly state verification problems. In order to minimize auditing costs and guarantee insurance protection against large losses to risk averse policy-holders, this optimal contract reimburses the total reported loss less the deductible when the reported loss is above the deductible and pays nothing otherwise. The contract specifies that the insurer commits itself to audit all claims with probability one and this deductible contract is optimal only for the class of deterministic mechanisms. Consequently, we should not observe any fraud, notably in the form of build-up, in markets with deductible contracts, since the benefits of such activity are nil. However, fraud is now a significant problem in automobile insurance markets for property damages where deductible contracts are often observed.

The recent literature on security design has proposed different extensions to take into account different issues regarding the optimal insurance contracts. Three main issues related to the empirical model of Dionne and Gagné (1997) are discussed in this literature. First, the deductible model implies that the principal fully commits to the contract in the sense that he will always audit all claims even if the perceived probability of lying is nil. It is clear that this contract is not renegotiation proof: at least for small losses above the deductible, the insurer has an incentive not to audit the claim and save the auditing cost. However, if the client anticipates such a behaviour from the insurer, he or she will not necessarily tell the truth when filing the claim!

One extension to the basic model was to suggest that random audits are more appropriate to reduce auditing costs. However, the optimal insurance contract is no longer a deductible contract and the above commitment issue remains relevant. Another extension is to suggest that costly state falsification is more pertinent than costly state verification for insurance contracting with ex-post moral hazard. The optimal contract under costly state falsification leads to insurance overpayments for small losses and under-compensation for severe accidents. We do not yet observe such contracts for property damages in automobile insurance markets, although they seem to be present for bodily injuries in some states or provinces (Crocker and Tennyson, 1996).

The empirical hypothesis of Dionne and Gagné, 1997, is as follows: when there is a sufficient high probability the fraud will succeed, the observed loss following an accident is higher when the deductible of the insurance contract is higher. Because they only have access to reported losses, a higher deductible also implies a lower probability of reporting small losses to the insurer. In order to isolate the fraud effect related to the presence of a deductible in the contract, they introduce some corrections in the data to eliminate the potential bias explained by incomplete information.

Their results are quite significant. They imply that when there are no witnesses (other than the driver and his or her passengers) on the site of the accident, the losses reported to
the insurance companies are somewhere between 24.6% and 31.8% higher for those insured with a $500 deductible relatively to those with a $250 deductible. Furthermore, they are confident that this increase corresponds to build-up, because their result is closely related to the presence of witnesses. Since the mean loss reported in their sample is $2552.65, these increases correspond to increases of the reported losses from $628 to $812, which is far more than the difference between the two deductibles ($250). Thus, it seems than when an insured decides to defraud, not only does he or she try to recover the deductible, but also to increase his or her net wealth (for instance, by increasing the net value of the automobile).

It may be argued that the choice of the deductible is the consequence of an extension of the traditional adverse selection problem because the insured anticipates higher expected losses. However, if this ex-ante argument were right, we should observe a significant effect of the deductible on reported losses even when the presence of witnesses is more likely, which was not the case. It would be surprising to obtain such an ex-ante effect only in the case of accidents without witnesses, because it is difficult to anticipate the type of accident and its severity when choosing the deductible ex-ante.

It may also be the case that insurers can affect the probability of successful falsification by increasing the frequency of audits in the case of claims for which no witnesses are involved and for which the policy bears a high deductible. In other words, insurers may use the presence of witnesses as a fraud indicator. If it is the case, the results show that insurers are not fully efficient in their investigations since there is still a significant effect associated with the deductible in the reported loss equation. This interpretation is supported by the fact that insurers only detect 33% of fraud when they audit (Caron and Dionne, 1997).

Recent contributions (Crocker and Morgan, 1998; Crocker and Tennyson, 1996) tend to show that other types of contracts are more effective than deductible contracts in reducing this type of ex-post moral hazard when falsification activities are potentially present. However, they limit the behaviour of the insurer to full commitment. The full characterization of an optimal contract in presence of ex-post moral hazard is then an open question in the literature.

6. Adverse Selection and the Quality of the Product in a Market

Akerlof (1970) was the first to propose a model with asymmetric information on the quality of products. This pathbreaking article has motivated many researchers to study the second-hand markets for durable goods. In general, owners of used goods know better the quality of their good than a potential buyer. Kim (1985) proposed a model suggesting that traded used cars should be of higher quality. Bond (1982) did test a similar proposition but did not find any evidence of adverse selection on the market for used pick up trucks. However Lacko (1986) did report some evidence for older cars only, a result also obtained by Genesove (1993). We now consider in detail this paper.

The main hypotheses to be considered for testing the presence of adverse selection are the following:
(a) During the transaction, one party is better informed than the other about the product's quality: usually the seller.

(b) Both of the parties involved in the transaction value quality.

(c) The price is not determined by either party but by the market.

(d) There is no market mechanism such as guarantees or reputation to eliminate adverse selection.

To test for residual adverse selection, Genesove (1993) analyzed the market for used cars sold by auction in the United States, where buyers have only a few moments to look at the cars and cannot take them for a test drive before purchase. The auction is simple: a series of ascending bids where the seller has the option of accepting or refusing the second highest bid. Sixty per cent of the sellers accept to relinquish their cars. The auction lasts one minute and a half, including the time to put the car up for auction and the time to remove it once the last bid is made! As a rule, the second price should correspond to the average quality of the cars offered, and buyers are supposed to be aware of this level of quality.

Genesove wanted to test whether any observable characteristic of the seller could be used to predict the average quality of the cars sold. In the presence of perfect information on the quality of the product, the characteristics of the seller would be of no importance. Only the quality of the product would count in explaining price equilibrium.

He thus considered two types of sellers participating in these auctions: those who sold only used cars (UC) and those who sold used cars and new cars (NC). Each seller participates in two markets: the auction market where the buyer makes no distinction in quality and a more traditional market where the real quality is more likely to be observed by the buyer.

It can be shown that the equilibrium price will be equal to the price matching the average quality each type of seller will offer. Thus, a seller whose cars are of superior quality to the average quality offered by this type will not put them up for auction unless there is a surplus in stock. In this case, he may offer some for auction, starting with those of lower quality. Moreover, the average quality of the two types may vary, as sellers may have different stock management systems. The author, in fact, shows that those who offer the two types of cars (used and new) have cars whose average quality is higher.

The motive behind stock management is important in finding an equilibrium. If the only motive for putting used cars up for auction is to take advantage of information asymmetry as shown in Akerlof's model, it is hard to obtain an equilibrium in a market where buyers are ready to pay for average quality and sellers are motivated to offer cars of only inferior or average quality. However, during a period of surplus stock, some sellers may have cars worth less than market value that they may be motivated to sell at the average-quality price, in order to gain a bonus. In other words, buyers in this type of market would have
to value cars more highly than sellers to obtain an equilibrium. Gibbons and Katz (1991) have used this type of argument to obtain an equilibrium in the work market with specific human capital.

Empirically, a positive bonus in an auction market is only possible in a situation of asymmetric information where the buyer pays the average-quality price associated with the type of seller. Thus a seller who is more likely to sell in this market because he often has surpluses will usually sell better quality cars and obtain, at the equilibrium, a higher average price for the same quality of car.

The author verified that, though the data covered cars from 1988 to 1984 and earlier, there is a significant bonus only for 1984 cars. This allows him to conclude that residual adverse selection is weak in this kind of market. This implies that enough information circulates by other mechanisms to reduce the informational bonus to zero. These mechanisms are reputation and guarantees. Sellers are not truly anonymous in the auction market. The seller must be present to accept or refuse the second price. Furthermore, there are limited guarantees protecting buyers during the first hour following the auction. So, as for the automobile insurance example, in Section 2, private markets use effective mechanisms for reducing residual adverse selection.

Two extensions are now discussed in the literature. The first one proposes to use price and quantity profiles overtime across brands of cars in order to isolate evidence of adverse selection (Hendel and Lizzeri, 1999). There will be evidence of adverse selection if the car that has a steeper price decline overtime also has the lower trade volume. This contrasts with the depreciation story where the faster price decline should correspond to a larger volume of trade. The second extension is to show that leasing can solve the lemons problem (Guha and Waldman, 1996; Hendel and Lizzeri, 1998).

7. Conclusion

We have taken up the difficult question of the empirical measurement of the effects of information problems on the allocation of resources. Two problems drew our attention: moral hazard and adverse selection.

One conclusion which seems to be accepted by a number of authors is that information problems may create considerable distortions in the economy in contrast with a situation of full and perfect information. But we have also found that effective mechanisms have been established to reduce these distortions and to eliminate residual problems at the margin.

This conclusion seems stronger for adverse selection than for moral hazard, at least in the markets studied. One possible explanation, which should be investigated in detail, is that adverse selection concerns exogenous factors, whereas moral hazard hinges on endogenous actions which are always open to modification.
Finally, given the specific nature of the problems studied —lack of information— we must be always prudent in our conclusions, since the effect measured cannot be 100% verified. There will always be a lingering doubt!

8. References


25


