

The Welfare Effects of Predictive Medicine*

P.A. Chiappori[†]

Columbia University

Forthcoming in:

“Insurance: Theoretical Analysis and Policy Implications”,

Edited by Pierre-André CHIAPPORI and Christian GOLLIER

MIT Press

Abstract

The paper tries to estimate the welfare loss due to early revelation of relevant information about individual risk (the 'Hirschleifer effect') in the case of predictive medicine. Using CRRA preferences with the range of

I received useful comments from Georges Dionne, an anonymous referee and the workshop participants. Errors are mine. Financial support from the NSF (Grant No. 0096516) is gratefully acknowledged.

Department of Economics, University of Chicago, and Department of Economics, Columbia University. Email: pc2167@columbia.edu.

coefficients that can be found in the literature, we argue that for 'large' risks, the loss can be major. We briefly discuss some of the problems raised by the regulatory interventions aimed at alleviating the loss.

1. Introduction

One of the more striking achievements of modern science is the progress of medical diagnosis and predictive medicine. Most infections can be detected long before the apparition of the first symptoms of actual illness (HIV testing being only one spectacular example). Even more remarkably, recent improvements in our understanding of genetic mechanisms allow to unveil predispositions to a number of diseases. In some cases (monogenic diseases typically), a genetic test enables physician to predict with quasi certainty the future occurrence (or not) of the pathology. Even in more difficult cases, in which the disease results from complex interactions between genetic and environmental factors, evidence of a genetic predisposition can significantly modify the *ex ante* probability of occurrence, with potentially major consequences on markets for health and life insurance. Clearly, these progresses will continue at a rapid pace. As a consequence, the context of insurance markets ten or twenty years from now may markedly differ from the current situation. The *ex ante* heterogeneity of individual status regarding insurance may dramatically increase, leading to a much wider dispersion in premium and coverage; some individuals may even be unable to acquire any coverage at all.

For an economist, the resulting dispersion can be analyzed from two different

viewpoints. Ex post, i.e. once tests are available, inter-individual differences can be viewed as inequalities, that governments may (or may not) be willing to alleviate. A more interesting perspective relies on an ex ante analysis. The prospect of future improvements in insurers' ability to predict raises for all agents a *classification risk* - the risk of being found to be a 'bad' risk. Risk averse agents may be willing to buy an insurance against this risk. In the absence of predictive tests, a standard insurance policy, which charges a similar premium to all agents, provides coverage against both the risk of illness and the classification risk. However, the availability of a test may, under circumstances that will be discussed below, restrict (or destroy) the feasibility of a coverage against the classification risk. In a world of risk averse agents, such restrictions of the scope of insurance unambiguously decrease welfare. This is the well-known 'Hirschleifer effect', whereby the availability of more accurate information may be harmful from a Pareto point of view.¹

The analysis just sketched is familiar to economists. Clearly enough, it describes only one side of the consequences of predictive medicine, although probably an important one. While better information may be detrimental to risk averse

¹The intuition that information availability may decrease welfare was initially mentioned by Drèze (1960). See Hirschleifer (1971) for a formal definition and a detailed discussion.

agents by restricting the scope of efficient risk sharing contracts, it may also improve decision making in a number of ways. Often enough, early knowledge of a risk allows to alleviate or eliminate its consequences through adequate prevention. The discovery a predisposition to breast cancer may lead to systematic investigations and/or preventive treatments, ultimately reducing mortality. Another example is provided by hemochromatosis, the most frequent genetic disease affecting Caucasian populations.² When undetected, hemochromatosis causes serious lesions to several organs, including liver, brain and lungs, and may generate lethal cancers. However, simple treatments are available, that allow to fully control the evolution of the disease and eliminate all harmful consequences. In the latter case, early information is unambiguously good.

In some cases, however, no preventive measures or treatments are available. A typical example is provided by Huntington disease, a lethal degeneracy of the nervous system. Huntington is a monogenic disease; a genetic test that has been available for several years allows to detect the presence of the gene, hence to predict either that the illness will certainly not occur, or that it will occur with probability almost one.³ Even in this extreme situation, though, the availability of

²About one person in ten carries at least one gene, resulting in one in four hundred developing the disease.

³The age of occurrence cannot be fully predicted, which explains the residual uncertainty. In

the test enable individuals to make more accurate choices regarding intertemporal allocation of resources. Major decisions regarding marriage, fertility, investment and others may be dramatically improved. The choice of a spouse, the decision to have children should take into account the genetic risk involved.⁴ Similarly, the optimal investment in human capital crucially depends on the time period during which returns will be received; again, a more precise knowledge of the risk can significantly ameliorate the efficiency of the decision to attend college or invest in specialized training, leading to potentially important gains in welfare.

Theory thus suggests that the development of predictive medicine will have a range of opposite consequences, with an ambiguous final impact on welfare. Hence, any welfare analysis must ultimately rely on an evaluation of the magnitude of the various effects at stake. The goal of the present paper is to propose a first and partial attempt at such a quantification. As it will be clear below, some key parameters (such as the joint distribution of income and risk aversions) are largely unknown, and some aspects of the decision processes (such as the subjective costs of learning one's exposure to a lethal risk) have not (yet) been fully explored. As

the vast majority of cases, the disease occurs between the ages of 30 and 45. Although the disease itself cannot be cured, the disabilities it provokes require expensive care.

⁴The number of cases of Tay-Sachs disease, a lethal illness that affects mostly members of the Ashkenazi Jewish community, has been considerably reduced through a careful monitoring of marriages by the community. The explicit aim of the monitoring was to avoid situations in which spouses would both carry the (recessive) gene.

such, the computations proposed here provide at best a partial and preliminary analysis of this difficult issue. We believe, however, that this attempt, imperfect as it may be, still constitutes a step in the right direction. The development of predictive medicine may have a deep impact on our lives, and raise difficult problems regarding regulation and health policy. It is unlikely that much progress can be made from a normative point of view without a more precise understanding of the main issues at stake, including the size of the welfare effects involved.

2. The costs of information availability: how large is the Hirschleifer effect?

Our focus, throughout the paper, is on the impact of predictive medicine on the market for insurance. Early knowledge of unfavorable predispositions may affect other aspects of economic life, most notably employment and housing; these issues will not be studied here. We consider the typical situation of an individual who, in the absence of the test, would be considered by insurance companies as 'average' in her risk class; however, an unfavorable test outcome increases the predicted risk of this individual relative to her class. Note that, in practice, individual risk can often be assessed even in the absence of sophisticated tests, although less

precisely. Insurance companies routinely ask questions about the health status of the subscriber's parents; should the latter suffer from health problems that are known to involve some genetic component, this information will be used for the underwriting process. In that sense, genetic testing is an old practice.⁵ Even when the disease has no known genetic component (AIDS being a typical case), proxy can be used to detect higher risk individuals.⁶ For these reasons, what predictive medicine provides is simply a more precise assessment of probabilities, resulting (potentially) in an exacerbation of existing inequalities. The analysis below should thus be understood as *conditional* on the individual's risk class, i.e. on all relevant observable characteristics; it thus applies within a 'cell', as defined by insurers from observable characteristics for underwriting purposes.

Finally, we assume in this section that any information resulting from a medical test is publicly revealed, and that in particular insurance companies can freely require and use this information for underwriting. This needs not be the case. The information could be revealed exclusively to the individual, and regulation

⁵During the recent debate on the access of insurance companies to the outcome of genetic tests of Huntington disease, an argument often used in support of free access was precisely that discrimination already existed, based on family history. The test, it was argued, would allow children of Huntington patients who did not carry the gene to access insurance in normal conditions.

⁶Anecdotal evidence suggests that in situations where HIV tests were either unavailable or prohibited for underwriting purposes, single males aged between 25 and 45 and working in specific professions were charged an extremely high price for life insurance.

may prohibit insurance companies from imposing the test, asking about privately undertaken tests, or simply using the outcome of any such test for underwriting purposes. The scope and the problems of such a regulation will be discussed in Section 3.

2.1. The Hirschleifer effect: theoretical background

Consider a disease that affects a proportion π of a given population.⁷ In the absence of predictive testing, all individuals in the population are facing the same probability π of getting the disease. As argued above, this assumption does not contradict the fact that some factors (age, gender, etc.) may influence the probability of getting the disease, since the population may be defined in a narrow sense (say, all Caucasian females aged between 20 and 30), and the analysis below can be repeated for any such cell within the total population. In the initial situation, we assume that agents can purchase full insurance coverage, for a premium $\gamma\pi$ proportional to the probability of occurrence (γ denoting the corresponding 'loading' coefficient).⁸

⁷For a detailed presentation, see Hirschleifer (1971) and Croker and Snow (2000)

⁸Note that, in theory, whenever the insurance is priced above the fair rate, agents should be willing to purchase only partial coverage. We thus disregard the various limitations (deductible, copayments,...) which would lead to some risk being born by the insured agent. This assumption, made for simplicity, allows to concentrate on the main issue while avoiding tedious 'background risk' computations.

We are interested in a situation in which a test allows to precisely assess the risk of each individual. Assume, for simplicity, that there are only two classes: 'high risk' agents, in proportion λ , face a probability of occurrence P , while 'low risk' agents, in proportion $1 - \lambda$, face a smaller probability p (with $\lambda P + (1 - \lambda)p = \pi$). Since the information is publicly revealed, high risk individual will be charged a fair rate, corresponding to P , instead of the average rate based on the ex ante probability π . Note that even when π is small, the conditional probability P may be large (it is close to one in the case of Huntington disease), resulting in a major increase of the premium. This is exactly the definition of the classification risk.

Assuming that the premium is always proportional to the probability of occurrence (γ denoting the corresponding 'loading' coefficient), the initial, sure premium $\gamma\pi$ is thus replaced with a random premium that equals either γp (with probability $1 - \lambda$) or γP (with probability λ). While, in expected terms, the change is revenue neutral, the introduction of a risk implies for risk averse agents a welfare loss, the magnitude of which depends on the (relative) size of the potential loss and the degree of risk aversion. A standard measure of the loss is its risk premium, defined as the certain amount the agent would be willing to give up, in the initial situation (i.e. before the test is introduced) in order to avoid the risk.⁹

⁹Other measures could be used (for instance, one could compute the certain amount that

Technically, the risk premium r is defined by

$$u(W - \gamma\pi - r) = \lambda u(W - \gamma P) + (1 - \lambda) u(W - \gamma p) \quad (1)$$

where u is the agent's VNM utility and W her wealth.

It should be noted that this formula relies on two implicit assumptions. One is that the agent would still be willing to pay the insurance premium, even if she was found to be of the high risk type. This needs not be the case: when found at risk, she may prefer to give up the insurance coverage altogether. The decision, again, depends on the agent's risk aversion, but also on the efficiency of the treatment that the insurance would cover. Specifically, the agent may, on the one hand, decide not to buy the insurance and pay for the required treatment out of her own pocket. This attitude is however unlikely from a risk averse agent, at least for reasonable values of the loading factor. On the other hand, she may decide, should the disease occur, to opt for an alternative and cheaper (although less efficient) treatment, or even to give up the treatment altogether. The analysis, here, varies with the details of the context at stake, and should in particular depend on the type of disease, its consequences on life expectancy and the quality

should be paid to the agent, once the test is available, to compensate her for the uncertainty).

of life, the range of treatments available, etc. A specific investigation devoted to a particular disease should obviously consider these possibilities. Given the general scope of the present paper, we stick to the assumption that the agent will always prefer to purchase a coverage. Our analysis, thus, fits particularly well cases in which either an expensive but (reasonably) efficient treatment exists, as for AIDS, or the disabilities provoked by the disease require expensive care, as for Huntington.

More generally, equation (1) assumes that the only impact of the classification risk is on the premium. This implies in particular that the level of coverage is not affected - say, because the coverage is complete in all cases. As before, this assumption, made for simplicity, does not sound particularly unrealistic.¹⁰ The possible distortions due to partial coverage are anyway of second order with respect to the main issue, namely the welfare loss due to the classification risk.

2.2. The distribution of risk aversion: empirical evidence

In formula (1) above, a key parameter is the form of the utility function, and in particular the resulting degree of risk aversion. Any tentative assessment of the

¹⁰A well-known puzzle of insurance is precisely that people tend to buy full insurance even when the price is above the fair rate. This effect is magnified when agents can self-insure through accumulated savings - see Gollier (2003).

cost of early information, as measured by the risk premium r , will heavily rely on the choice made at this level. A standard conclusion of most empirical studies, in insurance as well as finance, is that Constant Relative Risk Aversion (CRRA) utilities tend to provide a good fit to existing data. However, the value of the coefficient of relative risk aversion is not clear. Different sources have been used to estimate this crucial parameter.

1. Many experiments have been designed to study risk aversion. These attempts, however, have recently been questioned. Rabin (2000) argues in particular that small-scale lotteries cannot provide adequate estimates of people's reactions when faced with significant risks.¹¹ Perhaps more convincing are the various attempts based on 'natural experiments', an interesting example being provided by Beetsma and Schotman's (2000) study of the television game 'Lingo'. Despite all the flaws inherent to these approaches (specificity of the context, selection of participants,...), we believe that these 'real life' situations do provide useful information. In the Beetsma and Schotman paper mentioned above, the authors find that CRRA preferences provide a good fit, and their estimate of the coefficient of relative risk aversion is

¹¹Rabin's claim is actually stronger: he argues that expected utility theory is not adequate for modeling small bets.

around 7.

2. Another source comes from the analysis of portfolio composition, in particular the division of total wealth between risky and non risky assets. Under CRRA preferences, the share of wealth invested in risky assets does not depend on wealth. This assumption fits existing data reasonably well, at least for portfolio with non zero risky assets.¹² Several existing works (for instance Campbell 1996 and Blake 1996) find a high value for the coefficient of relative risk aversion (between 8 and 30).
3. The returns on risky assets can also be used to measure risk aversion: according to standard theory, the expected returns of stocks should exceed that on government bonds by a risk premium that directly reflects risk aversion. However, standard estimates using historical data find an implausibly high value for the coefficient of relative risk aversion (50 or more). This 'equity premium puzzle' has generated a huge literature, surveyed for instance in Cochrane (2000).
4. The analysis of saving behavior, based on the estimation of individual Euler equations, provides a joint estimation of risk aversion and intertemporal

¹²A significant proportion of low wealth portfolios do not include risky assets at all, a feature that is usually explained by the existence of fixed information and transaction costs.

substitutability. A first generation of models were using intertemporally separable preferences, of the CRRA (or constant elasticity) family. A standard criticism addressed to this literature was that the two determinants of intertemporal choices - risk aversion and income smoothing - were represented by the same coefficient, characterizing the concavity of the utility function. A second generation of models, using a more flexible specification, try to disentangle the two effects. The estimation of such models leads to complex identification problems. Still, Attanasio and Weber (1989) find a coefficient of relative risk aversion varying from 5 to 25 depending on the functional form.

5. In principle, insurance behavior should provide an ideal context for studying risk aversion. In practice, however, such an investigation faces important technical difficulties. A few studies, starting with Drèze (1981), try to recover risk aversion from the observation of individual choices among various contracts offering different premium and coverages. Usually, the CRRA assumption is not rejected, and the estimated coefficient is high (around 8). One problem with these results is that they require the strong assumption that people are aware of their *true* accident probability; the realism of this assumption may vary with the context.

6. Finally, some surveys ask questions about choices in hypothetical lotteries ('would you accept a new job that would ultimately, with equal probabilities, either increase your wage by 50% or decrease it by 30%?'). Despite the well known problems affecting this type of data, such studies present a significant advantage, since they allow a tentative estimation of the *distribution* of the coefficient of relative risk aversion in the population. One of the most widely used estimation was provided by Barsky et al (1997). Table 1 summarizes its main findings.

Class	I	II	III	IV
% of total population	65	11	11	13
Expected relative risk aversion*	15.7	7.2	5.7	3.8

* conditional on survey response

Table 1 - Risk premium for various RRA coefficients (L/W=5%, p=1%)

Here, the 'classes' are defined from the answers to two hypothetical questions, and expected relative risk aversions are computed assuming a lognormal distribution of risk tolerance with additional noise. According to these estimates, about two third of the population exhibit a level of risk aversion above a lower bound of 3.8, for an average over this class estimated at 16, while only 13% of the responses

are compatible with a 'low' level. Moreover, average risk aversion does not seem to vary much between different levels of wealth, suggesting that the distributions of wealth and risk aversion are roughly independent.¹³

2.3. The distribution of risk aversion: a simple calibration

A remarkable feature of the previous studies is that despite the diversity of the approaches adopted, the estimations of the coefficient of relative risk aversion are consistently quite high (above 5). A simple calibration exercise confirms that such values should not be considered unrealistic. Consider the benchmark case of a 1% probability of losing an amount L equal to 5% of the person's total wealth; one may think, for instance, of the risk of car theft for a middle class household. The following table summarizes the risk premium, as a percentage of the *ex ante* fair rate π , for CRRA preferences and various values of the coefficient of relative risk aversion:

Coefficient of relative risk aversion	1	2	3	5	10
Risk premium (% of ex ante fair price)	3%	5%	8%	15%	31%

¹³However, Guiso and Paiella (this volume), in a similar context, question the adequacy of the CCRA form.

Table 2 - Risk premium for various RRA coefficients (L/W=5%, p=1%)

A remarkable feature of this table is that the risk premium exceeds the standard loading factor on theft insurance (usually between 20 and 30%) only for high levels of relative risk aversion (close to 10). Moreover, this result is largely robust to changes in the occurrence probability, provided that it remains reasonably small. For instance, if the probability is increased to 10%, the table becomes:

Coefficient of relative risk aversion	1	2	3	5	10
Risk premium (% of ex ante fair price)	2%	5%	7%	12%	26%

Table 2 bis - Risk premium for various RRA coefficients (L/W=5%, p=1%)

The results are however sensitive to the magnitude of the loss. Keeping the probability at 1% but increasing the relative loss to 10% gives the following:

Coefficient of relative risk aversion	1	2	3	5	10
Risk premium (% of ex ante fair price)	5%	11%	17%	31%	74%

Table 2 ter - Risk premium for various RRA coefficients (L/W=10%, p=1%)

This is in line with standard intuition about risk aversion: the benefits of insurance increase rapidly with the size of the loss - a property that plays a key role in what follows. Still, even in the latter case insurance coverage with a 30% loading factor will not be purchased by individuals whose coefficient of risk aversion is smaller than 5. Altogether, the purchase of insurance against a risk of this size - a commonly observed behavior - indicates a 'large' coefficient of risk aversion. This conclusion is pretty much in line with the previous results in suggesting that a significant fraction of the population may exhibit a 'large' (i.e., above 5) degree of relative risk aversion.

2.4. The Hirschleifer cost: a first approximation

Using the results described above, it is possible to propose a preliminary calibration of the size of the loss resulting from the absence of insurance against the classification risk. A key parameter is the size of the potential loss, i.e. the ex post premium P charged to a high risk individual, as a proportion of the individual's wealth. At this point, a calibration requires more precise hypothesis. If

the amount at stake is relatively 'small' - say, less than 10% of total wealth (or, equivalently, if the annual cost of insurance for high risk individuals is 'small' with respect to their annual income), the Tables above indicate that the welfare loss resulting from inability to cover the classification risk, as measured by the risk premium, remains low (certainly less than the fair ex ante premium for reasonable levels of risk aversion). The picture is totally different in the case of 'large' risks, i.e. when the ex post cost of insurance for high risk agents represents a significant proportion of disposable income. Table 3 is based on the same parameter values as Table 2, except that the size of the annual loss is now assumed to be 50% of disposable income.

Coefficient of relative risk aversion	1	2	3	5	10
Risk premium (% of ex ante fair price)	39%	98%	194%	590%	3560%

Table 3 - Risk premium for various RRA coefficients (L/W=50%, p=1%)

The risk premium is now considerably higher than the fair premium - three times as much for a RRA coefficient of 3, seven time as much for a coefficient

of 5, more than thirty six time as much for a coefficient of 10. Not surprisingly, these numbers increase even further for larger relative losses. Even when the RRA coefficient is fixed at the modest level of 3, the risk premium is represents about thirty times the fair premium when the potential loss represents 80% of income.

How relevant are these calibrations? Unfortunately, for many diseases and most income levels, treatment costs representing of 50% or even 80% of disposable income are by no means unrealistic. The annual cost of treatment of an HIV positive patient receiving a preventive tritherapy is of the same order of magnitude as the median per capita net disposable income in the US. Anecdotal evidence abounds of cases in which non covered individuals lose a large fraction or even the totality of their wealth after a severe illness. Accepting the conclusion put forth by Barsky et al. that the correlation between wealth and risk aversion is weak, it is probably fair to say that the financial risk linked to the inability to cover the classification risk, for serious diseases such as AIDS, represents for a large fraction of the population a very significant proportion of disposable income.¹⁴

A more difficult issue is the relevance of CRRA with a relatively large coefficient for such 'large' risks. Although we tend to believe that the estimates

¹⁴The Medicaid system has sometimes been described as 'an impicit insurance contract with a deductible equal to your wealth'.

provided represent a reasonable 'best guess', it is fair to say that the question is largely open; not much exists on the empirical estimation of risk aversion in the case of large financial risks.

Assuming the calibration above provides a good first approximation of the phenomenon at stake, then the amount paid on insurance premium, a number often used to quantify the importance of the insurance business, may provide a dramatically biased estimation of the magnitudes at stake. In the case of major financial risks, our estimates suggest that premiums may underestimate the true welfare cost by a factor of twenty (or more). Just to provide an order of magnitude, assume, following Barsky et al., that the two third of the population has a 'large' level of relative risk aversion, and let us adopt for them the (very) conservative average value of 8 (remember that Barsky et al. estimate their mean relative risk aversion at 15). Assume that treatment costs, relative to the income distribution, are such that for 15% of this sample, the cost of the treatment would average (or exceed) 50% of the household's disposable income - again, a conservative assumption in the case of AIDS. We thus isolate a sub-population, representing 10% of the total, which is particularly vulnerable to the classification risk. For these agents alone, the welfare loss resulting from the loss of coverage against the classification risk - as estimated by the risk premium - would represent about

\$2,000 per household, hence an aggregate welfare loss exceeding probably twenty billions dollars. If one further assumes that the AIDS risk is actually larger than average for this poorer fraction of the population, the number could be even higher; and, again, we are only considering a small fraction of the total population. The striking conclusion is that the potential losses due to the Hirschleifer effect may in some cases be huge.

2.5. The benefits of early knowledge

As stressed in Introduction, early information about a predisposition to some disease may, and will in many cases, lead to efficient prevention or treatment, thus considerably increasing welfare. In such contexts, the Hirschleifer effect vanishes (with the notations above, p and P are equally small) and early information has a strictly positive value. In this subsection, we consider the opposite extreme situation of a pathology for which early detection has no direct medical benefit. Huntington disease is a typical example of such a context; the presence of the genetic anomaly causing Huntington simply predicts a largely ineluctable and almost always fatal evolution. Even in this case, early knowledge can affect long term behavior, and result in significant economic gains.

Stoler (2004) has recently proposed a preliminary investigation of the impact

of such early knowledge on investment in human capital, with an emphasis on the decision to attend college. Clearly, the benefits of human capital accumulation increase with the length of the horizon during which the returns will be received. Huntington disease reduces the expected length by one half on average, with an additional uncertainty on the age of occurrence. A key characteristic of many genetic diseases, including Huntington, is that almost all patients know early (i.e., when one of their parents start developing the disease) that they are at risk. If one parent carries the gene, the probability that each child will also carry it, hence develop the disease, is exactly 50%. The impact of the test is to replace this 50% probability with a certainty, in one direction or the other. Stoler computes the benefit of this shift as a function of the individual-specific college premium, i.e. the expected gain in income resulting from college attendance for the individual under consideration. The general shape of the gain is given in Figure 1 below.

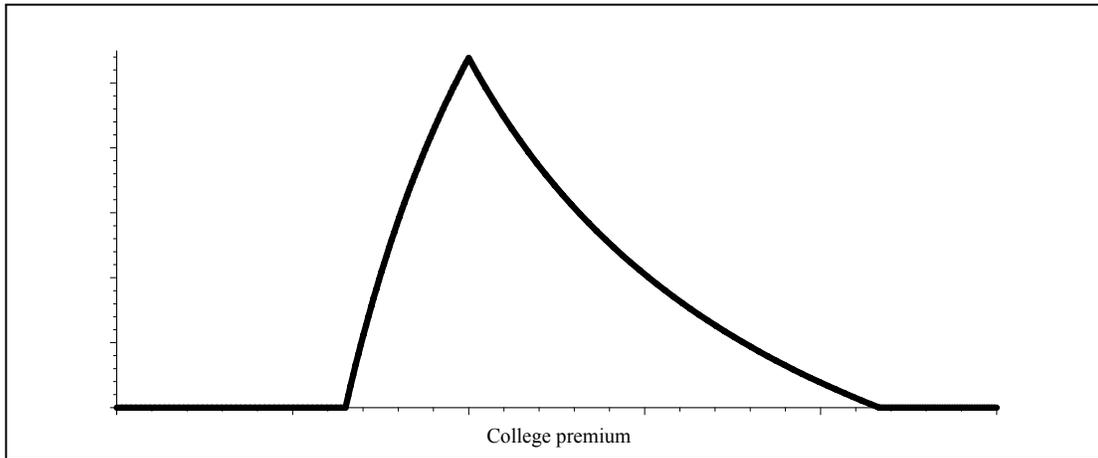


Figure 1: ex ante benefits of early detection of Huntington disease, as a function of the college premium
(CES utility, perfect financial markets)

Source: Stoler (2003)

For low level of the premium, college attendance is not profitable even for an healthy agent, hence the benefit of early knowledge is null. At a higher premium, a healthy individual would choose to attend college, but the 50% uncertainty is sufficient to discourage college attendance. The healthy children of a Huntington patient thus benefit from the test by deciding to attend college, and their gain increases with the premium. Then a threshold is reached, beyond which college

attendance is the optimal decision even under the 50% uncertainty, although not for an individual knowing for sure she will develop the disease. The benefit of the test goes only to agents who are found to carry the gene and avoid therefore to attend college; note that this benefit decreases with the premium. Finally, for very high levels of the college premium, agents will attend college irrespective of the test outcome, and the benefit is again null.

Simplistic as this Figure may be, it still emphasizes a crucial aspect, namely that the benefits of early detection are linked with a key, individual-specific parameter, namely the potential benefits of human capital accumulation. In general, these are modeled as determined by agents' ability, a characteristic that is (partially) unobservable by the econometrician and heterogeneously distributed in the population. Again, a cost-benefit analysis of early detection must involve the joint distribution of ability and risk aversion, on which little is known. Empirical work currently in progress on Huntington patients may help clarifying this complex issue.

3. Regulating discrimination: the curse of adverse selection?

3.1. Regulation: a strategic analysis

Given the potential magnitude of the Hirschleifer effect, a natural question is whether government intervention could alleviate the resulting loss in welfare. It has often been argued, in particular, that the access of insurance companies to the results of genetic or other tests should be strictly regulated. Specifically, various levels of regulation have been proposed, that can be broadly summarized as follows:

- Level 1: insurance companies are allowed to require any test to be undertaken by the subscriber during the underwriting process.
- Level 2: insurance companies are not allowed to require a test, but may ask questions about any test that may have been privately undertaken by the subscriber in the past; the subscriber cannot decline to answer.
- Level 3: explicit questions about past tests are prohibited; however, subscribers are allowed to voluntarily communicate to their insurer any test result they may have, and insurers are then allowed to use this information

for underwriting.

- Level 4: insurance companies are not allowed to use test information for underwriting, even when this information has voluntarily been supplied by the subscriber.

Various versions of these four levels have been proposed and sometimes implemented. In particular, a basic principle governing medical ethics is that no information about health status, including the outcome of a test, should be disclosed (including to an insurance company) without the patient's informed consent.

From an economic viewpoint, however, many of the distinctions sketched above appear spurious. A simple, game-theoretic analysis of this issue relies on two preliminary remarks.¹⁵

Remark 1: *if free, anonymous tests are feasible, then an agent always benefits from **privately** acquiring information about her health status.*

The operative word, here, is 'anonymous'. Indeed, a crucial issue is whether the mere fact that the agent has undertaken the test is public information or can be kept secret by the agent. In the former case, an agent can readily avoid the

¹⁵The analysis sketched below summarizes classical results of the literature, including contributions by Hoy (1989), Tabarrok (1994), Doherty and Thistle (1996) and recently Doherty and Posey (1998), Hoy et al. (2003) and Eisen (this volume). For a recent survey, see for instance Dionne, Doherty and Fombaron (2000).

Hirschleifer effect by deciding not to get the information. In the hypothetical case in which the test has no benefit (i.e., no treatment is available, prevention is vain, and the agent's investment decisions are inelastic to her life expectancy), then no agent will ever choose to take the test, and the competitive equilibrium is ex ante efficient. Note that such an outcome obtains under one condition only: namely, the agent must be able to prove that she did not undertake the test.

Let us now consider the alternative case, in which such a proof does not exist; i.e., the claim that the test has not been taken cannot be verified by the insurance company. Then the Hirschleifer effect becomes an Hirschleifer curse. Indeed, an agent cannot lose by *secretly* acquiring the relevant information, since she can either use it, discard it, or even claim she never acquired it. The curse comes from the fact that this private benefit for each participant may ultimately result in a public loss for all agents, because of the externality induced by adverse selection.

Remark 2: *In a competitive market, a patient who finds she belongs to the low risk type always gains in freely revealing this information to the insurance company.*

In particular, if free, anonymous tests are feasible and if the patient can, if she wants, freely communicate 'hard' evidence of their result to insurers, then regulation levels 1, 2 and 3 above are equivalent, and result in full revelation of

information (associated to the corresponding contract discrimination) at equilibrium.

The economic intuition is clear. In a competitive market, an agent who is known to belong to the low risk type will be charged a smaller premium than both high risk agents and agents of unknown type. It follows that, at equilibrium, all agents will privately acquire the information, and those for whom news are good will forward the result to their insurer. Insurance companies, in turn, anticipate this pattern and will tend to treat similarly the agents who were found to be high risk and those who (claim they) did not undertake the test.¹⁶ Ultimately, this treatment will justify both the decision to privately undertake the test and to forward the favorable result to the insurer.

How empirically relevant is this game-theoretic analysis? In practice, anonymous tests are hard to eliminate. In the case of AIDS, for instance, such tests are paramount.¹⁷ Assuming that costs are negligible may be more questionable,

¹⁶The latter claim should however be qualified in the case in which insurers can implement full separation through the offer of an adequate menu of contracts. In that case, three contracts will be offered, corresponding to the three categories of agents (tested positive, tested negative, untested). See Doherty and Posey (1998) for a very careful analysis of this model. In practice, however, full separation may be quite difficult to obtain when adverse selection is in fact multidimensional (i.e., agents differ in risk, but also risk aversion, wealth,...).

¹⁷The case of AIDS is specific because of the contamination risk: uninformed agents are a danger for their potential partners. This externality may well justify the availability of any-

especially when the psychological costs of learning about one's true condition are taken into account. In practice, patients are often reluctant to undertake a test, and cite as their main motivation the fact that 'living with such a knowledge can be an excessively heavy burden'.¹⁸ However, in the equilibrium structure just described, the potential gain from learning may be significant (if the result is a low risk classification), with a zero downside (since no difference is made between high risk and unknown agents). In the end, the answer will depend, among other things, on the magnitude of the potential gain from being found to be a low risk with respect to the (financial and psychological) cost of undertaking the test. Empirical evidence suggests that, in some cases, the impact is indeed minimal.¹⁹ In other situations, AIDS being an obvious but by no means unique example, the opposite logic prevails; then it is safe to consider, at least for diseases that are not too rare, that the patient's free access to the result of the test constitute a natural benchmark case to evaluate the impact of regulation.

The most problematic assumption is probably the free revelation of information to the insurer. A situation were no agent chooses to disclose the information

mous test despite the induced Hirschleifer curse. Obviously, such arguments do not apply to genetic tests.

¹⁸This conclusion, for instance, comes out very clearly in Stoler's (2004) survey.

¹⁹For an example dealing with breast cancer genetic testing and life insurance markets, see Subramanian et al. (1999).

relative to her health status is in principle not stable (in the usual, Nash equilibrium sense), because any low risk agent would have an incentive to privately inform her insurer. Assume, however, that the proportion of high risk agents, λ , is very small. The difference in risk between a low risk insuree and the average individual in the population may then be negligible, or at least insufficient to justify a new contract; then the incentives to privately undertake the test and communicate its results (if favorable) disappears altogether. In technical terms, for 'rare' diseases, even small transaction cost may stabilize a pooling equilibrium.

It should however be stressed that this point holds true only insofar as the proportion of high risk individuals *within the population under consideration* is small - which is often not the case in practice, because insurers use existing information to isolate particular subpopulations of 'potentially at risk' individuals. Testing each American citizen for Huntington would not make much sense, because the proportion of individuals carrying the gene is so small that costs would be proportionally excessive. However, the problem typically arise for the specific subpopulation of persons who have a parent suffering from the disease. Then the probability is large (50%), and the argument just sketched fully applies - as it may apply to the subpopulation of single males working in artistic professions for AIDS, and other examples can readily be found.

At any rate, the previous arguments indicate that regulation of information diffusion is a complex issue, because of the externality inherent to any adverse selection context. It suggests in particular that the emphasis generally put on voluntary disclosure of information is partly misplaced. Forwarding an information against the will of the patient is clearly unacceptable in most cases,²⁰ and everyone agrees on necessary enforcement of this prohibition. The previous analysis shows, however, that such a prohibition is generally not sufficient: a world in which information can only be voluntarily supplied (our 'level 3') is likely to result in exactly the same discrimination as complete disclosure of information, because insurance companies are likely to consider any agent as a high risk unless proved otherwise. In order to be effective, a regulation must probably prohibit the use of test outcomes altogether, even when they are voluntarily supplied by the subscriber. This corresponds to our 'level 4' regulation, which can be called 'comprehensive'.

3.2. Comprehensive regulation and adverse selection

We now consider the case of 'level 4' regulation, whereby insurers are not allowed to take into account any information. Such a ban creates a typical situation

²⁰A possible exception could be health policy consideration in the case of severe epidemics.

of adverse selection on the insurance markets: a relevant information (the test outcome) is known by the agent, but cannot be used by the insurance company. It is useful, at this point, to briefly review the standard analysis of competition under adverse selection. A basic distinction should be established between exclusive and non exclusive contracts. In a word of exclusive contracts (health insurance being a typical illustration), insurees can freely choose the company from which they buy coverage, but then their contractual relationship with this company is exclusive, in the sense that they cannot simultaneously buy insurance against the same risk from another provider. Annuity contracts provide the opposite example of non exclusive relationship: any agent can simultaneously buy annuities from different firms.

Exclusive contracts involve a stronger commitment from one of the parties; not surprisingly, it tends to generate more efficient relationships, since some contracts are simply not available in the absence of exclusivity (see Bennardo and Chiappori 2003 for a precise discussion). Technically, a crucial difference is that exclusivity is necessary to implement *convex pricing*, a natural tool in a context of asymmetric information.²¹ The consequences of adverse selection on competitive

²¹Convex pricing implies that the unit price increases with the quantity purchased. Clearly, such a feature cannot be implemented if agents can independently buy small quantities from different providers.

markets markedly differ in the two contexts.

Non exclusive contracts In the absence of exclusivity, contracts typically involve linear pricing, the unit price being identical for all agents.²² The consequences of adverse selection in this context are well known.²³ High risk agents will, everything equal, purchase larger amounts of coverage, driving up the unit price. If the proportion of high risk agents is small and low risk agents are sufficiently risk averse, an equilibrium exists in which low risk individuals subsidize high risk agents - a property that can actually improve welfare from an ex ante point of view. However, with many high risk individuals and/or insufficient risk aversion, low risk agents may stop purchasing insurance altogether, which results in a collapse of the market: only high risk individuals are left, and they pay their fair rate.

The comparison with the full information disclosure context is thus delicate, and depends on the parameters of the model at stake. The welfare analysis will typically rely on the following remarks:

²²Life insurance is a good example: according to Cawley and Philipson (1998), price schedules are actually slightly concave, probably reflecting a technology in which fixed costs play an important role, while under adverse selection optimal pricing tend typically to be convex (although this property may depend on the form of preferences).

²³See Chiappori and Rochet (2003) for a general analysis involving multidimensional adverse selection

1. When the proportion of high risk agents is low, and risk aversion is important (say, because the risk at stake is large), then the prohibition may be ex ante welfare increasing by alleviating the ex post discrimination between high and low risk agents
2. In the opposite situation (a significant number of high risk individuals, and lower degree of risk aversion in general), the regulation is unambiguously harmful, since it cannot avoid discrimination (high risk types are charged an ex post fair price), and implies the additional cost that low risk agents are not covered at all.
3. In both cases, an additional regulation imposing a ceiling on the amount purchased can be useful. In complex cases of multidimensional adverse selection, optimal contracts are typically more complex, and may involve full disclosure for higher levels of purchase.

A final remark is that the previous discussion implicitly assumes that agents are committed to their contract and cannot renegotiate or resell it once purchased. In the life insurance framework, for instance, this implies that the purchased contract will be carried over until it expires (or the patient dies). However, recent innovations contradict this assumption. The huge development of the market for

viaticals, whereby agents sell the rights stemming from their life insurance contract to a third party at a fair price minus a fee, has dramatically altered the picture. The theoretical analysis of viaticals is particularly interesting because it provides a typical example of a contractual innovation that is both remarkably efficient ex post and remarkably inefficient ex ante. Ex post efficiency comes from the fact that an agent who simultaneously learns (say, from an HIV or a genetic test) that her life expectancy is much shorter than expected and that large sums will have to be spent on health in the next future may greatly benefit from immediate availability of (a fraction of) her life insurance coverage. However, the availability of viaticals essentially forbids ex ante any significant cross-subsidization between risk classes. Given the existence of a resale market on which the insurance contract can be sold *at a fair price*, an initial premium significantly smaller than the fair amount would result in huge arbitrage opportunities and the predictable collapse of the market altogether. From an ex ante perspective, thus, viaticals destroys the possibility of a coverage against the classification risk.

Exclusive contracts The case of exclusive contracts is more difficult to analyze, if only because no general agreement has been reached so far about the optimal equilibrium concept. The general intuition stemming from the classical

(Rothschild and Stiglitz 1976, Hellwig 1984) as well as more recent (Bisin and Gottardi 2002, Geanakoplos et al. 2003) literature is that competition typically result in the apparition of a range of contracts which may de facto implement full separation of types. The basic idea is that companies will try to attract particular subsegments of the subscriber population through an adequate design of the proposed contracts. For instance, a basic contract may offer full coverage at a high unit price, but clients may choose much cheaper options involving specific limitations of coverage; for a relevant definition of the limitations, the options may attract the low risk fraction of the population, while high risk individuals will remain covered by the basic contract, probably at a fair ex post rate.²⁴ In other words, such menus, which are standard in insurance and seem difficult to prohibit, are likely to result in full discrimination, at least insofar as adequately designed menus can be used to fully separate the various types.

Clearly, should such an outcome be predicted, then the judgment one can formulate about the regulation would be negative: not only discrimination would be just as strong as in the case of complete disclosure of information, but in addition low risk agents would face strong limitations of coverage. Whether this separation

²⁴Typical examples include contracts imposing a ceiling on lifetime expenditures for some specific diseases, such as AIDS.

can be achieved as an *equilibrium* outcome is however unclear, and depends on the particular equilibrium concept used. In the initial Rothschild-Stiglitz contribution, no equilibrium exists when the proportion of high risk agents is small - although the empirical meaning of such a conclusion is somewhat ambiguous. Various alternative concepts have been developed, leading to different predictions (the equilibrium is always separating for some concepts, but may be pooling for others) - although it can be noted that several recent approaches, based either on game theoretical refinements or on a more 'Walrasian' perspective, seem to support separation.

4. Conclusion

The first conclusion emerging from this brief presentation is that many questions remain open, especially from an empirical viewpoint. The consequences of competition under adverse selection are crucially relevant for any normative analysis of these issues. It is fair to say that the *theoretical* understanding of these problems has not yet reached a state of general agreement. More damaging is the fact that the empirical knowledge of these issues is scarce.²⁵

²⁵One can mention among others Cutler and Zeckhauser (1997), Browne and Frees (1999), and Buchmueller and DiNardo (1999).

A second conclusion is that any cost benefit analysis of early information disclosure must adopt a case-by-case approach, and consider each disease as specific. Clearly, diseases for which an early diagnosis leads to an improved treatment deeply differ from those, such as Huntington, for which no cure exists. From a regulatory perspective, 'scarce' pathologies should also be distinguished from 'frequent' ones.

A third implication is that in any case, regulation is a complex issue, for which an explicit and precise economic analysis is necessary. One pattern emerging from the above discussion is for instance that regulation, in order to be effective, must be quite comprehensive. Prohibiting the use of genetic against the will of the patient is certainly necessary, if only to protect individual rights, but is probably not sufficient to avoid discrimination. Even a general ban on the use of test outcomes for underwriting may in some cases be counterproductive, resulting in full scale discrimination plus severe restrictions on access to coverage by low risk agents. Again, more work is needed, especially from an empirical perspective.

Various policies have been proposed, from a lax regulation that acknowledges discrimination as unavoidable but tries to limit its most shocking excesses, to compulsory public coverage of all 'large' risks. Intermediate solutions have been proposed. Some specialists advocate a lump sum payment to people carrying ad-

verse genes. Diamond (1992) has proposed a system of group insurance, whereby a government agency would divide the population into large groups and private firms would bid for the coverage of an entire group. The recently introduced Swiss regulation provides a very interesting example. In particular, it explicitly restricts the firms' ability to offer options and menus of contracts - a feature clearly aimed at avoiding the implementation of option-based revelation mechanisms.

Finally, it should be stressed that the classification risk issue goes well beyond the impact of predictive medicine. Classification into a 'high risk' group may also result from exterior events, such as the occurrence of a serious health problem.²⁶ One of the basic challenges facing any health coverage system is precisely its ability to provide long term contracts covering these types of risks. In a paper in this Volume, Mark Pauly (2005) shows how the market developed adequate responses to this problem. For instance, all *individual* health insurance contracts in the U.S. contain guaranteed renewability provisions (required now by law) which protect insureds against selective increases in premiums based on their health state or health insurance. The case of group health (or life) insurance is more complex. On the one hand, contracts are usually such that premiums do not increase for

²⁶For instance, it is well known that an agent who suffered from a heart attack faces a significantly higher risk of future health problems.

an individual based on a test result or drugs in health state. On the other hand, a worker who loses his job after a significant increase in his health risk is likely to pay a significant higher premium for his health coverage in the future. In another 'worst case' scenario, the worker belongs to a firm which is small, so that his increased risk raises significantly the total cost of health coverage to the employer, who reacts either by dropping the coverage altogether or by switching to a new contract entailing low caps on some treatments.²⁷ And, of course, millions of individuals are currently not covered. This suggests classification risk and the Hirschleifer effect will remain major issues for policymakers in the future.

²⁷In December 1987, John McGann's was diagnosed with AIDS. His (small) employer opted out its previous HMO coverage, and adopts a self-financed plan. The new plan entailed a \$5,000 ceiling for lifetime expenditures related to AIDS (as opposed to a ceiling of \$1,000,000 in the initial plan). All court rulings were favorable to the employer.

References

Attanasio, O., and G. Weber (1989), 'Intertemporal Substitution, Risk Aversion and the Euler Equation for Consumption', *Economic Journal*, vol. 99, no. 395, pp. 59-73

Barsky, Robert, B., Thomas F. Juster, Miles S. Kimball and Matthew D. Shapiro (1997), "Preference Parameters and Behavioral Heterogeneity: an Experimental Approach in the health and Retirement Study", *Quarterly Journal of Economics*, CXII, 2, 537-580.

Beetsma, R., and P. Schotman (2001), 'Measuring Risk Attitudes in a Natural Experiment: Data from the Television Game Show Lingo', *Economic Journal*, vol. 111, no. 474, pp. 821-48

Blake, D. (1996), 'Efficiency, Risk Aversion and Portfolio Insurance: An Analysis of Financial Asset Portfolios Held by Investors in the United Kingdom', *Economic Journal*, Vol. 106, No. 438, pp. 1175-1192.

Campbell, J. (1996), 'Understanding Risk and Return', *Journal of Political Economy*, vol. 104, no. 2, pp. 298-345

Campbell, J., and J. Cochrane (2000), 'Explaining the Poor Performance of Consumption-Based Asset Pricing Models', *Journal of Finance*, vol. 55, no. 6, pp. 2863-78

Crocker, K., and A. Snow (1986), 'The Efficiency Effect of Categorical Discrimination in the Insurance Industry', *Journal of Political Economy*, 94, 321-44.

Crocker, K., and A. Snow (2000), 'The Theory of Risk Classification', in *Handbook of Insurance*, G. Dionne, Ed., Kluwer, London.

Diamond, P. (1992), 'Organizing the Health Insurance Market', *Econometrica*, vol. 60, no. 6, pp. 1233-54

Dionne, G., N. Doherty and N. Fombaron (2000), 'Adverse Selection in Insurance Markets', in *Handbook of Insurance*, G. Dionne, Ed., Kluwer, London.

Doherty, N., and P. Thistle (1996), 'Adverse Selection with Endogenous Information in Insurance Markets', *Journal of Public Economics*, vol. 63, no. 1, pp. 83-102

Doherty, N., and L. Posey (1998), 'On the Value of a Checkup: Adverse Selection, Moral Hazard and the Value of Information', *Journal of Risk and Insurance*, vol. 65, no. 2, June 1998, pp. 189-211

Drèze, J. (1960), 'Le paradoxe de l'information', *Economie Appliquée*, 13, pp. 71-80

Eekhoudt Louis, and Miles Kimball (1992), "Background Risk, Prudence and the Demand for Insurance", in *Contributions to Insurance Economics*, G. Dionne ed. London: Kluwer Academic Press.

Eekhoudt, L, C. Gollier, and H. Schlesinger (1996), "Changes in Background Risk and Risk Taking Behavior", *Econometrica* 3, 64, 683-689.

Eisen, R. (2005), 'Adverse Selection in the Health Insurance Market after Genetic Tests', this volume.

Gollier, C. (2003), 'To Insure or Not to Insure?: An Insurance Puzzle', *The Geneva Papers on Risk and Insurance Theory*, 28 (1): 5-24,

Guiso, L., and M. Paiella (2001), 'Risk Aversion, Wealth and Background Risk', mimeo, Bank of Italy

Hirshleifer, J. (1971), 'The private and social value of information and the reward of inventive activity', *American Economic Review* 61 , 561–574.

Hoy, M. (1989), 'The Value of Screening Mechanisms under Alternative Insurance Possibilities', *Journal of Public Economics*, vol. 39, no. 2, pp. 177-206

Hoy, M., F. Orsi, F. Eisinger and J. P. Moatti (2003), 'The Impact of Genetic Testing on Healthcare Insurance', *Geneva Papers on Risk and Insurance: Issues and Practice*, vol. 28, no. 2, pp. 203-21

Pauly, M. (2005), 'Time, Risk, Precommitment, And Adverse Selection In Competitive Insurance Markets', this volume.

Rabin, Matthew (2000), "Risk Aversion and Expected Utility Theory: A Calibration Theorem", *Econometrica*, 68, 1281-1292.

Stoler, A. (2004), 'Economic implication of genetic testing and mortality risk', mimeo, University of Chicago.

Subramanian, K., et al (1999), 'Estimating Adverse Selection Costs from Genetic Testing for Breast and Ovarian Cancer: The Case of Life Insurance', *Journal of Risk and Insurance*, vol. 66, no. 4, pp. 531-50

Tabarrok, A. (1994), 'Genetic Testing: An Economic and Contractarian Analysis', *Journal of Health Economics*, vol. 13, no. 1, pp. 75-91