On Privacy, Adverse Selection and Genetic Databases

Completed Research

Word Count: 4991

Track: Society, Policy, & Regulation

Abstract

Stored genetic samples will provide more information in the future on medical risks due to better genetic technology, than was imagined when the samples were originally collected. This poses a threat to individual privacy due to secondary use of the collected genetic information. In this paper we study the impact of genetic testing on health insurance markets. Consumers who have taken a genetic test are assumed to be informed about their risk of a genetic disease. The informed low risk consumers have an incentive to reveal genetic information to the insurer so as not to be clubbed with the high risk types. We characterize the existence and nature of insurance contracts when individuals can reveal genetic information to insurers but where revelation of genetic information is associated with a loss of privacy. We then examine the welfare implications of different policy proposals regarding genetic testing, with the decision of the consumer to take a genetic test and to reveal genetic information, being endogenous.
1. Introduction

Government and private medical institutions are increasingly using electronic databases to manage individual medical records (Lehrman 2000). An example of such databases is the vast amounts of data created by the Human Genome Project. Genetic information stored in these databases can be used to estimate the likelihood that an individual will suffer from a variety of conditions in the future (Roche et al 1996). A slew of new genetic methods, tests and knowledge promise to revolutionize medicine now that scientists have cracked the human genome map. But those same advances could land health-care providers and insurance companies in a quagmire of sticky issues, including conflicts with patient privacy (Wired.com 2000). The ethical, legal and social implication (ELSI), of mapping and sequencing the human genome, has identified privacy as one of the high-priority research areas (Meslin et al 1997).

Genetic tests attempt to detect the presence of genes that are associated with disease or predispose those who inherit the genes to disease (Meslin et al 1997). Genetic testing exacerbates the adverse selection existing in insurance markets if the results of the test are not made accessible to the insurers. However the insurers’ request for a level playing field contrasts with efforts of consumer groups to increase the privacy protection of genetic information (Subramaniam et al 1999). Since the DNA molecule is stable once removed from a person’s body and stored, it can become the source of an increasing amount of information as more is learned about how to read the genetic code (Troy 1997). In other words stored genetic samples will provide more genetic information than was imagined when the samples were originally collected. Threats to individual privacy may therefore be presented due to secondary use of the collected information (Roche et al 1996). Because of the influence that genetic knowledge can have on an individual’s right to privacy, the Genetic Privacy Act (GPA) suggests that the acquisition, use, and disposition of genetic knowledge is best placed in the hands of the individual.
In this paper we study the impact of genetic testing on health insurance markets. Consumers who have taken a genetic test are assumed to be informed about their risk of a genetic disease. The informed low risk consumers have an incentive to reveal genetic information to the insurer so as not to be clubbed with the high risk types. We characterize the existence and nature of insurance contracts when individuals can reveal genetic information to insurers but where revelation of genetic information is associated with a loss of privacy. We then examine the welfare implications of different policy proposals regarding genetic testing, with the decision of the consumer to take a genetic test and to reveal genetic information, being endogenous. The rest of the paper is organized as follows: Section 2 covers the background literature on privacy and genetic testing. Section 3 covers the assumptions, model framework and main results. Implications of the results obtained are summarized, and limitations and directions for future research are discussed in Section 4.

2. Background

Privacy is defined as the moral claim of the individuals to be left alone and to control the flow of information about themselves (Coase 1960; Westin 1967). Revealing genetic information to insurers constitutes a loss of privacy because of the inherent loss in control over the use of this information after it is released. The loss of privacy could be the fear of secondary use of the genetic information revealed or the fear of discrimination at a later date due to some information present in the genetic information which the consumer is not aware of currently.

It is well known that insurance markets are characterized by adverse selection (see Rothschild and Stiglitz 1976) - asymmetry in information on risk types between the informed or partially informed consumers and uninformed insurers has an adverse impact on the profits of insurers. Genetic testing increases the asymmetry in information if the results are available only to consumers, thus making the adverse selection more acute. There is a lot of debate on how and whether the results of the genetic tests should be used (see Tabarrok 1994). There are four broad
policy proposals regarding the use of genetic tests for insurance purposes. Under the first proposal, insurers do not inquire, and are not permitted to inquire whether applicants have been tested. In the second proposal, genetic information is revealed to the insurer only with the consent of the consumer. The third policy alternative is to permit insurers to inquire whether applicants have been tested and to get genetic information/test results with the consent of the consumer. The fourth policy alternative is to require or permit insurers to require that applicants be tested, and allow insurers to use the test results—i.e. mandatory release of genetic information for insurance purposes. (Doherty and Thistle 1996).

Hirshleifer (1971) distinguishes between private information (available only to a single individual) and public information (available to everyone). Our work is closer to the private information stream of literature – Cremer and Khalil (1992), Crocker and Snow (1992) and Doherty and Thistle (1996). Agents receive a private signal in our model and they choose whether to make the signal public or not.

Crocker and Snow (1992) have shown that if the insurers know whether the consumers are informed or uninformed about their risk type, then the private value of information on risk type for the consumer is negative. Doherty and Thistle (1996) have looked at the case where insurers cannot distinguish between informed and uninformed agents and the case where information on risk type can be concealed or revealed to the insurers. In their setting uninformed can take a test (both costless and costly) to get informed about their risk type and can reveal information to insurers are zero cost. The policy holders’ knowledge of their risk type is therefore endogenous. They characterize the nature of equilibrium in all these cases and find primarily that with costly information some uninformed may choose to remain uninformed. They also examine the welfare implications of the endogenous information model for public policy regarding genetic testing.

In our setting, revelation of information has positive costs associated with it in contrast to Doherty and Thistle (1996). We characterize this loss to be the loss of privacy – the fear of

---

1 We also consider a purely public signal in the last section.
discrimination because of being classified as a high-risk type in a later period, due to better genetic tests. All genetic testing policy proposals except the first one where insurers are neither permitted to inquire nor use genetic information are affected if the loss of privacy is taken into account. As far as the impact on consent law (Proposal 2), we need not consider the privacy concerns of the high risk type since presumably only the low risk types are affected. However in the case of mandatory testing and revelation of information, the loss in privacy of the high risk type is also taken into account.

3. Model

The insurance market setting is competitive with risk neutral insurers. There is a continuum of risk-averse consumers, identical in all respects except for accident probabilities. They possess identical von Neumann-Morgenstern state independent utility functions $U(.)$, which are twice continuously differentiable and strictly concave. Their wealth is $W$ state 1 (no loss) and $W - D$ in the loss state 2, where $W > D$. $\pi_i; i = H, L$ are the probabilities of loss of the high and low risk consumers. We also assume that $\pi_H > \pi_L$, $0 < \pi_H, \pi_L < 1$, and that these probabilities are out of the agent’s control, so that no moral hazard problem arises. We assume that each agent knows her risk type. Informed consumers in our setting are consumers who have taken a genetic test in a previous period for some non-insurance related reason e.g. job screening.

The proportion of low risk (high risk) consumers in the economy is $\theta_L$ ($\theta_H$). If agents do not purchase insurance they obtain expected utility, $V(0, \pi_i)$ for $i = H, L$; where

$$V(0, \pi_i) = \pi_i U(W - D) + (1 - \pi_i)U(W); i = H, L.$$  

The insurance contract company specifies the premium $P$ and the net indemnity $I$ paid incase of a loss. The use of coinsurance rates and/or deductibles is not permitted. We use the notation

---

2 Moral hazard problem arises due to the fact that insurers cannot observe the consumers actions – how much care she takes to avoid a loss. By assuming that consumers cannot affect the probability of an accident by their actions we assume away the moral hazard.

3 Net Indemnity is the payment made to the consumer in case of loss minus the premium.
\( \delta_i = \{P_i, I_i\} \) to denote the single period insurance contract offered to type \( i = H, L \). The expected utility of a consumer with the probability of loss \( \pi_i \) under the insurance policy \( \delta_i \) is

\[
V(\delta_i, \pi_i) = \pi_i U(W - D + I_i) + (1 - \pi_i) U(W - P_i); i = H, L
\]

If consumer releases genetic information to the insurer then the consumer faces a loss \( \gamma \) due to loss of privacy. This loss being a non-monetary loss is separable from the utility from the insurance contract. The net utility is

\[
V(\delta_i, \pi_i) - \gamma
\]

Initially let us assume that consumers cannot reveal information to the insurers, i.e. \( \gamma = 0 \). Consumers choose the policy that maximizes expected utility. Competition drives down insurance premiums to an actuarially fair rate.

In the first best case (no asymmetry in information on risk types), both risk types get complete insurance. The contract in the first best case is

\[
\delta_i^* = \{\pi_i, D_i, (1 - \pi_i)D_i\}; i = H, L
\]

However under asymmetric information (Consumers have a priori knowledge of risk types while insurers don’t), the first best contract can’t be achieved (Rothschild and Stiglitz 1976). The second best contract satisfies the constraints

\[
V(\delta_i, \pi_i) \geq V(\delta_j, \pi_i); i = (H, L), i \neq j,
\]

\[
V(\delta_i, \pi_i) \geq V(0, \pi_i); i = H, L
\]

The separating contracts are \( \{\delta_i^*, \delta_i^+\} \) (see Fig. 1). The high risk consumer get complete insurance while the low risk agent get less than full insurance.
We now assume that a proportion \( \lambda_U \) \( 0 < \lambda_U \leq 1 \) of the consumers does not have a priori knowledge of its risk type. Out of these uninformed consumers a proportion \( p_H \) (\( p_L \)) are high risk (low risk). The proportion of informed high risk type (low risk types) is \( \lambda_H \) (\( \lambda_L \)), \( 0 < \lambda_H, \lambda_L < 1 \)\(^4\).

\[
\theta_H = \lambda_H + p_H \lambda_U
\]

\[
\theta_L = \lambda_L + p_L \lambda_U
\]

All these proportions are assumed to be common knowledge. The prior probability of loss for the uninformed is \( \pi_U = \sum_{i=H,L} p_i \pi_i \). The timing is as follows, the insurers first choose the set of contracts to offer based on their beliefs on the actions by the uninformed (take genetic test or not) and the informed low risk (reveal information or not). The uninformed then decide whether to take a genetic test or not, and the informed consumer decide whether to reveal information to the insurer or not. The choice at each node affects the payoffs of all agents.

---

\(^4\) Strict inequalities are assumed here for simplicity. Thus there are always informed high and low risk consumers.
Insurance contracts for the different risk types under different information structures are shown in Fig. 1. \( W_1 \) and \( W_2 \) are the wealth in the no loss and loss state respectively. The 45° line is the complete insurance line – equal wealth in both states. \( EP_H \), \( EP_U \) and \( EP_L \) are the fair odds line (zero profit for insurer) for the high risk, uninformed and the low risk consumer. \( V_H \), \( V_U \), and \( V_L \) are representative indifference curves for the high risk, uninformed and the low risk agents\(^5\). Under full information the contracts offered are \((\delta_H^*, \delta_U^*, \delta_L^*)\). Under asymmetric information, with no uninformed the contracts are \((\delta_H^*, \delta_L^*)\), while with uninformed consumers the contracts are \((\delta_H^*, \delta_U^*, \delta_L^*)\). With asymmetric information, only the high risk consumers get full insurance. It is easy to see that the following constraints should be satisfied.

\[
V(\delta_L^*, \pi_L) > V(\delta_U^*, \pi_L) \quad V(\delta_L^*, \pi_U) > V(\delta_H^*, \pi_L)
\]

\[
V(\delta_U^*, \pi_U) = V(\delta_L^*, \pi_U) \quad V(\delta_U^*, \pi_U) > V(\delta_H^*, \pi_U)
\]

\[
V(\delta_H^*, \pi_H) = V(\delta_U^*, \pi_H) \quad V(\delta_H^*, \pi_H) > V(\delta_L^*, \pi_H)
\]

\[
V(\delta_H^*, \pi_H) = V(\delta_L^*, \pi_H) \quad V(\delta_L^*, \pi_L) > V(\delta_H^*, \pi_L)
\]

We consider two cases, one where the insurers cannot observe whether the consumers are informed about their risk types or not and the second where the insurers can see the consumers information status. Doherty and Thistle (1996) give the example of a specific diagnostic test e.g. an HIV test where the consumers can get to know for sure whether they have tested positive or not. If the insurers always get to know if the consumer has taken a test, then they know for sure that the consumer knows his risk type. In contrast to this specific test, physical exams give no such specific result and so insurers never get to observe the information status of the consumers. Crocker and Snow (1992) have looked at the observable information status case while Doherty

\(^5\) The slope of the indifference curves of low risk, uninformed and high risk consumers at the full insurance line are \(-\frac{(1-\pi_L)}{\pi_L} > -\frac{(1-\pi_U)}{\pi_U} > -\frac{(1-\pi_H)}{\pi_H}\)
and Thistle (1996) have looked at the unobservable information status case. Both of these papers fail to address the issue of privacy of genetic information. In our model we model the loss of privacy due to revelation of information to insurers and this differentiates our work from theirs – so from now on $\gamma > 0$. In section 3.1 we consider the case where all consumers are informed. In section 3.2 we look at the case where the information status is not directly observable but there is an incentive for consumers to reveal verifiable information (negative test result or genetic information that reveals that the consumer is a low risk consumer) and testing is costly for the uninformed. The observable information status case with costly revelation and costly testing is considered in section 3.3. The mandatory testing and revelation of genetic information is considered in section 3.4.

3.1 Unreported positives, verifiable negatives and costly revelation.

We first consider the case where all consumers are informed about their risk type i.e. $\lambda_U = 0$.

**Proposition 1** Assuming all consumers are informed about their risk types, insurers cannot observe risk type and consumers can report verifiable information to the insurers at non zero cost $\gamma$, the equilibrium contract are $(\delta^*_H, \delta^*_L)$ and $(\delta^*_H, \delta^*_L)$. The equilibrium $(\delta^*_H, \delta^*_L)$ only exists if $\gamma < V(\delta^*_L, \pi_L) - V(\delta^*_H, \pi_L)$ and Pareto dominates $(\delta^*_H, \delta^*_L)$ if $\gamma < V(\delta^*_L, \pi_L) - V(\delta^*_L, \pi_L)$.

**Proof.** See Appendix

Denoting $V(\delta^*_L, \pi_L) - V(\delta^*_H, \pi_L) = \gamma^*$ and $V(\delta^*_L, \pi_L) - V(\delta^*_L, \pi_L) = \gamma'$. $\gamma^*$ and $\gamma'$ can be considered to be different levels of disutilities due to loss of privacy with $\gamma^* > \gamma'$. The contract $(\delta^*_H, \delta^*_L)$ exists only when $\gamma < \gamma^*$ i.e. for small disutilities due to loss of privacy. For high losses due to revelation of information, $\gamma \geq \gamma^*$, the second best contract $(\delta^*_H, \delta^*_L)$ is the only equilibrium. It should be noted that the second best contract is always an option to the

---

6 The existence of this equilibrium requires the Rothschild-Stiglitz condition of there being a sufficient number of low risk consumers. We assume for simplicity that this condition is always satisfied where relevant.
insurer but the first best contract can only exist for small value of losses due to revelation. If \( \gamma < \gamma' \) then the contract \((\delta^*_H, \delta^*_L)\) Pareto dominates \((\delta^*_H, \delta_L)\). The following conclusions can be drawn from proposition 1

a) For \( 0 \leq \gamma < \gamma' \), \((\delta^*_H, \delta^*_L)\) is the Pareto optimal equilibrium

b) For \( \gamma' \leq \gamma < \gamma'' \), both \((\delta^*_H, \delta^*_L)\) and \((\delta^*_H, \delta^*_L)\) are possible equilibrium.

c) For \( \gamma \geq \gamma'' \), \((\delta^*_H, \delta^*_L)\) is the only possible equilibrium.

Now we consider the case \( \lambda_U > 0 \). If we assume that testing is costless, then testing is a dominant strategy for the uninformed. All uninformed consumers would take the test and become informed. Thus with costless testing, there is no impact on these results if we assume that there are some consumers who are uninformed.

3.2 Unreported positives, verifiable negatives, costly revelation, uninformed consumers and unobservable information status.

Uninformed consumers can take a costly test to know their risk type. Also by consent law they can choose who gets to see the results of the test. It is clear that only the low risk agents have an incentive to release information to the insurers.

**Proposition 2.** Assume uninformed consumers can observe their risk type at cost \( c \), informed and uninformed consumers cannot be distinguished, risk type is not directly observed by insurers and consumers can report verifiable information to insurers at non-zero cost \( \gamma \). Denoting \( \frac{c}{p_L} \) by \( c' \),

four contracts are feasible:

i) \((\delta^*_H, \delta^*_L)\) if \( c' < V(\delta^*_L, \pi_L) - V(\delta^*_H, \pi_L) \)

ii) \((\delta^*_H, \delta^*_L)\) if \( \gamma < V(\delta^*_L, \pi_L) - V(\delta^*_H, \pi_L) - c' \)

iii) \((\delta^*_H, \delta^*_U, \delta^*_L)\) if \( V(\delta^*_L, \pi_L) - V(\delta^*_U, \pi_L) - c' < \gamma < V(\delta^*_L, \pi_L) - V(\delta^*_U, \pi_L) \)

iv) \((\delta^*_H, \delta^*_U, \delta^*_L)\) if \( \gamma > V(\delta^*_L, \pi_L) - V(\delta^*_U, \pi_L) - c' \).
Proof. See Appendix

Denote \(V(\delta^*_L, \pi_L) - V(\delta^*_U, \pi_L) = \gamma_1, V(\delta^*_L, \pi_L) - V(\delta^*_U, \pi_L) = \gamma_2,\)

\(V(\delta^*_L, \pi_L) - V(\delta^*_H, \pi_L) = \gamma_3\) and \(V(\delta^*_L, \pi_L) - V(\delta^*_H, \pi_L) = \gamma_4.\) Assuming \(\gamma_3 > \gamma_2,\) it is easy to see that \(0 < \gamma_1 < \gamma_2 < \gamma_3 < \gamma_4 < \infty\) where \(\gamma_i, i = 1..4\) are different levels of aggregate disutilities due to loss of privacy and due to the cost of testing. Also representing contracts \((\delta^*_H, \delta^*_L) \equiv \delta^1,\)

\((\delta^*_H, \delta^*_L) \equiv \delta^2, (\delta^*_H, \delta^*_U, \delta^*_L) \equiv \delta^3\) and \((\delta^*_H, \delta^*_U, \delta^*_L) \equiv \delta^4.\) Fig. 2 shows the region where these contracts would exist.

**Fig. 2 Existence of different Insurance Contracts by Region.**

The disutility due to loss of privacy is represented on the X-axis while the scaled cost of testing is plotted on the Y-axis\(^7.\) For small values of aggregate disutilities (disutility due to loss of privacy + disutility due to testing), none of the uninformed stays uninformed. The contracts offered are \(\delta^1\) and \(\delta^2.\) For moderate values of aggregate disutility, there are uninformed who prefer to remain uninformed and based on specific values three or four contracts could exist. For

---

\(^7\) The appearance of Fig. 2 varies based on the relative magnitudes of aggregate disutility but the basic intuition remains the same.
high values of aggregate utilities both low risk informed prefer not to reveal information and uninformed prefer to remain uninformed or uninformed become informed while informed low risk don’t reveal information or uninformed stay uninformed and informed low risk reveal information. Wherever $\delta^1(\delta^3)$ and $\delta^2(\delta^4)$ are offered simultaneously and are the only contracts, $\delta^2(\delta^3)$ Pareto dominates $\delta^1(\delta^4)$.

3.3 Unreported positives, verifiable negatives, costly revelation, uninformed consumers and observable information status.

Everything is same as in section 3.2 except that now the information status can be observed by the insurers i.e. the insurers know which consumers are informed and which consumers are uninformed about their risk type.

Proposition 3. Assume uninformed consumers can observe their risk type at cost $c$, informed and uninformed consumers can be distinguished, risk type is not directly observed by insurers and consumers can report verifiable information to insurers at non-zero cost $\gamma$. Denoting $\frac{c}{p_L} = c^*$, four contracts are feasible:

i) $(\delta^*_H, \delta^*_L)$ if $c^* < V(\delta^*_L, \pi_L) - V(\delta^*_H, \pi_L)$

ii) $(\delta^*_H, \delta^*_L)$ if $\gamma < V(\delta^*_L, \pi_L) - V(\delta^*_H, \pi_L) - c^*$

iii) $(\delta^*_H, \delta^*_U, \delta^*_L)$ if $\gamma < V(\delta^*_L, \pi_L) - V(\delta^*_U, \pi_L)$

iv) $(\delta^*_H, \delta^*_U, \delta^*_L)$

Proof: See Appendix

Denoting $V(\delta^*_L, \pi_L) - V(\delta^*_U, \pi_L) = \gamma^*_2$, and $\gamma_3, \gamma_4$ same as before and assuming that $\gamma^*_2 < \gamma_3$, we have $0 < \gamma^*_2 < \gamma_3 < \gamma_4 < \infty$. Also denoting $(\delta^*_H, \delta^*_U, \delta^*_L) \equiv \delta^5$, $(\delta^*_H, \delta^*_U, \delta^*_L) \equiv \delta^6$ and $\delta^1$ and $\delta^2$ are same as in section 3.2, the regions of existence of the
various contracts are shown in Fig. 3. Where $\delta^5$ and $\delta^6$ are offered simultaneously, and are the only contracts offered, the contracts $\delta^5$ Pareto dominates $\delta^6$.

Fig. 3 Existence of different Insurance Contracts by Region.

3.4 Mandatory testing and release of information for insurance

The privacy loss of the high risk (low risk) consumer due to release of information is $\gamma_H(\gamma_L)$ where $\gamma_H > \gamma_L$. All consumers have to reveal genetic information to the insurers to get insured. This implies that all uninformed have to get informed in order to get insurance.

**Proposition 4.** Assume uninformed consumers can observe their risk type at cost $c$, risk type is not directly observed by insurers and high risk (low risk) consumers can report verifiable information to insurers at non-zero cost $\gamma_H$ ($\gamma_L$) where $\gamma_H > \gamma_L$ and reporting information is mandatory for insurance, then denoting $V_i = \delta_i - \gamma_i$ and $V_i = V(0, \pi_i)\forall i = H, L, U$

i) If $V_H' < V_H$ and $V_L' < V_L$ nobody buys insurance.

ii) If $V_H' < V_H' < V_L'$ and $p_H V_H' + p_L V_L' - c < V_U'$ only informed low risk buy insurance and the only contract offered is $\delta^*_L$. 

12
iii) If \( V_H < \pi_H \); \( V_L > \pi_L \) and \( p_H V_H + p_L V_L - c > \pi_U \), then uninformed become informed and the low risk buy the only contract offered \( \delta^*_L \).

iv) If \( V_H > \pi_H \) and \( V_L > \pi_L \) and \( p_H V_H + p_L V_L - c < \pi_U \), then uninformed remain uninformed and remain uninsured, both high risk and low risk reveal information and are offered the contracts \( (\delta^*_H, \delta^*_L) \).

v) If \( V_H > \pi_H \) and \( V_L > \pi_L \) and, \( p_H V_H + p_L V_L - c > \pi_U \) then uninformed become informed, both high risk and low risk reveal information and are offered the contracts \( (\delta^*_H, \delta^*_L) \).

*Proof:* Straightforward comparison of the utilities of each type of agent with or without insurance with the clause that insurance requires mandatory testing leads to above results.

4. Discussion

Nature of insurance contracts under each of the policy proposals has been summarized in the propositions. Under the first proposal where insurers are neither permitted to ask whether the consumer has undergone any test nor ask for test results, the contracts offered are - full insurance for the high risk and incomplete insurance for the uninformed and the low risk. These contracts rely on pure self-selection a la Rothschild and Stiglitz. If release of information is allowed by consent law and insurers cannot observe the information status (Proposition 2), then for the low aggregate disutilities due to testing and loss of privacy, the low risk consumers can also get complete insurance. Since the uninformed cannot be distinguished from the high risk, the insurer has to distort the contract for the uninformed so that the high risk does not prefer this contract to her own. For higher values of aggregate disutilities, the insurer is better off offering the same contract that it offers under the first proposal. Thus from a welfare perspective, Proposal 2 is better than Proposal 1. The results of Proposition 3 are relevant to Proposal 3, where the insurer always knows the information status of the consumers. The results suggest that the uninformed can also get full insurance if they choose to become informed. The key difference between
proposals 2 and 3 is that the insurers can identify the informed high risk from the uninformed. Since this information is conveyed to the insurers without any inherent loss to the consumers of any type, from a welfare perspective Proposal 3 dominates Proposal 2. This result should be kept in mind by policy makers when they choose from the different proposals on genetic testing. Mandatory testing and mandatory revelation of information does the worst in terms of overall social welfare since, it imposes additional costs on the high risk and uninformed agents without providing any benefits to the high risk and providing benefits to the uninformed only in some cases.

Based on the relative magnitudes of the loss in privacy and the cost of testing, there could be 1-4 equilibriums in a given region as is shown in Fig. 2 and Fig. 3. It is possible that some of these Nash equilibrium can be eliminated by the intuitive criterion. All the equilibrium calculated are separating equilibrium. Pooling and partial pooling equilibrium can be ruled out since they won’t exist in the competitive insurance setting described here. If all insurers offer a partial pooling or a pooling contract then there is an incentive for an insurer to deviate and offer separating contracts.

References

http://www.wired.com/news/medtech/0,1286,37271,00.html


**APPENDIX**

**Proof of Proposition 1.**

We verify that it is an equilibrium for the low risk consumers to reveal information and get the contract $\delta^*_L$, and then show that not revealing information can also be an equilibrium.

If insurers expect the low risk to reveal information, then they offer the contracts $(\delta^*_H, \delta^*_L)$. The value of revealing information to the low risk is
\[ u^* = [V(\delta^*_L, \pi_L) - \gamma] - V(\delta^*_H, \pi_L) \]

If \( \gamma < V(\delta^*_L, \pi_L) - V(\delta^*_H, \pi_L) \) then \( u^* > 0 \). Since there is value to revealing information all low risk consumers would reveal information. If \( \gamma > V(\delta^*_L, \pi_L) - V(\delta^*_H, \pi_L) \), then value to revealing information is negative \( (u^* < 0) \). All low risk consumers would choose the contract \( \delta^*_H \).

However Rothshild and Stiglitz (1976) have shown that a pooling contract cannot be equilibrium. So \((\delta^*_L, \delta^*_H)\) is an equilibrium if \( \gamma < V(\delta^*_L, \pi_L) - V(\delta^*_H, \pi_L) \).

If the insurers expect the low risk not to reveal information, then they offer the contracts \((\delta^*_L, \delta^*_L)\). The value of revealing information is

\[ u^* = [V(\delta^*_L, \pi_L) - \gamma] - V(\delta^*_L, \pi_L) = -\gamma \]

Since value of information is negative, low risk do not reveal information. The high risk types are indifferent between \( \delta^*_H \) and \( \delta^*_L \), so \((\delta^*_L, \delta^*_L)\) is also an equilibrium.

The insurers and the high risk types are indifferent between the contracts \((\delta^*_H, \delta^*_L)\) and \((\delta^*_H, \delta^*_L)\). The low risk types prefer \((\delta^*_L, \delta^*_L)\) if \( \gamma < V(\delta^*_L, \pi_L) - V(\delta^*_L, \pi_L) \). Thus the equilibrium \((\delta^*_H, \delta^*_L)\) Pareto dominates the equilibrium \((\delta^*_H, \delta^*_L)\) if \( \gamma < V(\delta^*_L, \pi_L) - V(\delta^*_L, \pi_L) \)

**Proof of Proposition 2.**

There are three players in the game – the uninformed, the informed low risk type and the insurer. The strategy space for the uninformed is to get informed or to remain uninformed\(^9\). The strategy space for the informed low risk is to reveal information or not to reveal information. The strategy space for the insurer is to design the set of contracts. We adopt the following approach to calculate the equilibriums - the insurer picks the best contracts given one of the four possible

---

\(^8\) We assume for simplicity throughout that the low risk prefer the contract meant for the high risk to remaining uninsured i.e. \( V(\delta^*_H, \pi_L) > V(0, \pi_L) \)

\(^9\) The informed high risk consumer always end up with the contract \( \delta^*_H \) and are not part of the game.
combinations of strategies of the other two players. If none of the other two players has an incentive to deviate given the contract offered by the insurer and the strategy of the other player, then the contract is an equilibrium. We calculate all the equilibriums in four steps; each step corresponds to the insurers’ belief on the four possible strategy combinations of the other two players.

**Step 1:** Assume that the insurers expect the informed low risk type to reveal information and the uninformed type to become informed, then it would offer the contracts \((\delta_H^*, \delta_L^*)\). The value of revealing information for the informed low risk types is

\[
u^* = [V(\delta_L^*, \pi_L) - \gamma] - V(\delta_H^*, \pi_L)
\]

If \(\gamma < V(\delta_L^*, \pi_L) - V(\delta_H^*, \pi_L)\), \(u^* > 0\) and so all informed low risk types would reveal information. The value of information for the uninformed is

\[
I^* = \{p_HV(\delta_H^*, \pi_H) \plus p_L[V'(\delta_L^*, \pi_L) - \gamma] - c\} - V(\delta_H^*, \pi_U)
\]

\[
I^* = p_L[V(\delta_L^*, \pi_L) - V(\delta_H^*, \pi_L) - \gamma] - c
\]

If \(\gamma < V(\delta_L^*, \pi_L) - V(\delta_H^*, \pi_L) - c'\) where \(c' = \frac{c}{p_L}\), then \(I^* > 0\), all uninformed would become informed. Thus given that the insurers expect the informed low risk to reveal information and the uninformed to become informed, the informed low risk have an incentive to reveal information and the uninformed have an incentive to get informed. So \((\delta_H^*, \delta_L^*)\) is an equilibrium. If \(\gamma > V(\delta_L^*, \pi_L) - V(\delta_H^*, \pi_L) - c'\), then all uninformed would remain uninformed and \((\delta_H^*, \delta_L^*)\) cannot be an equilibrium.

**Step 2:** Assume now that the insurers expect the informed low risk not to reveal information and the uninformed to get informed. In this case, the insurers offer the contracts \((\delta_H^*, \delta_L^*)\). The value of revealing information for the informed low risk is

\[
u^* = [V(\delta_L^*, \pi_L) - \gamma] - V(\delta_L^*, \pi_L) = -\gamma
\]
Since the value of revealing information is negative, the informed low risk will not reveal information. The value of getting informed for the uninformed is

\[ I^* = \{ p_H V(\delta_H^*, \pi_H) + p_L V(\delta_L^*, \pi_L) - c \} - V(\delta_H^*, \pi_U) \]

\[ I^* = p_L [V(\delta_L^*, \pi_L) - V(\delta_H^*, \pi_L)] - c \]

If \( c' < V(\delta_L^*, \pi_L) - V(\delta_H^*, \pi_L) \) then the value of information \( I^* \) is positive, uninformed will become informed and \((\delta_H^*, \delta_L^*)\) will be an equilibrium. \((\delta_H^*, \delta_L^*)\) cannot be an equilibrium if \( c' > V(\delta_L^*, \pi_L) - V(\delta_H^*, \pi_L) \).

**Step 3:** Assume now that the insurers expect the informed low risk type to reveal information and the uninformed to remain uninformed. In this case the, insurers offer the contracts \((\delta_H^*, \delta_U^*, \delta_L^*)\).

The value of revealing information for the informed low risk types is

\[ u^* = [V(\delta_L^*, \pi_L) - \gamma] - V(\delta_U^*, \pi_U) \]

If \( \gamma < V(\delta_L^*, \pi_L) - V(\delta_U^*, \pi_U) \), then \( u^* > 0 \), and all informed low risk reveal information. The value of getting informed for the uninformed is

\[ I^* = \{ p_H V(\delta_H^*, \pi_H) + p_L [V(\delta_L^*, \pi_L) - \gamma] - c \} - V(\delta_H^*, \pi_U) \]

\[ I^* = p_L [V(\delta_L^*, \pi_L) - V(\delta_U^*, \pi_L) - \gamma] - c \]

If \( \gamma < V(\delta_L^*, \pi_L) - V(\delta_U^*, \pi_L) - c' \), then the value of getting information is positive, they choose to become informed and so the contract \((\delta_H^*, \delta_U^*, \delta_L^*)\) cannot be an equilibrium. If \( \gamma > V(\delta_L^*, \pi_L) - V(\delta_U^*, \pi_L) - c' \), then uninformed continue to remain uninformed. Thus for the range \( V(\delta_L^*, \pi_L) - V(\delta_U^*, \pi_L) - c' < \gamma < V(\delta_L^*, \pi_L) - V(\delta_U^*, \pi_L) \), the set of contracts \((\delta_H^*, \delta_U^*, \delta_L^*)\) is also an equilibrium.
Step 4: Assume now that the insurers expect the informed low risk type not to reveal information and the uninformed to remain uninformed. In this case the insurers offer the contracts \((\delta_H^*, \delta_U^*, \delta_L^*)\). The value of revealing information for the informed low risk type is
\[
\begin{align*}
    u^* &= [V(\delta_L^*, \pi_L) - \gamma] - V(\delta_L^*, \pi_L) = -\gamma
\end{align*}
\]
Since the value of revealing information is negative, the informed low risk will not reveal information. The value of getting informed for the uninformed is
\[
\begin{align*}
    I^* &= p_H V(\delta_H^*, \pi_H) + p_L [V(\delta_L^*, \pi_L) - \gamma] - c
    \begin{align*}
    I^* &= p_L [V(\delta_L^*, \pi_L) - V(\delta_U^*, \pi_U) - \gamma] - c 
    \end{align*}
\]
If \(\gamma < V(\delta_L^*, \pi_L) - V(\delta_U^*, \pi_U) - c^*\), then the value of getting informed is positive and so uninformed will become informed. Thus \((\delta_H^*, \delta_U^*, \delta_L^*)\) cannot be an equilibrium. However if
\[
\begin{align*}
    \gamma > V(\delta_L^*, \pi_L) - V(\delta_U^*, \pi_U) - c^* \text{, then the value of information is negative and so uninformed will choose to remain uninformed and the contracts \((\delta_H^*, \delta_U^*, \delta_L^*)\) will be an equilibrium.}
\end{align*}
\]

**Proof of Proposition 3.**

As before there are three players in the game – the uninformed, the informed low risk type and the insurer. We calculate all the equilibriums in four steps; each step corresponds to the insurers’ best response based on beliefs on the four possible strategy combinations of the other two players.

**Step 1** and **Step 2** are identical to the corresponding steps in Proposition 2. \((\delta_H^*, \delta_L^*)\) is an equilibrium if \(\gamma < V(\delta_L^*, \pi_L) - V(\delta_H^*, \pi_L) - c^*\) and \((\delta_H^*, \delta_L^*)\) will be an equilibrium if \(c^* < V(\delta_L^*, \pi_L) - V(\delta_H^*, \pi_L)\).

**Step 3:** Assume now that the insurers expect the informed low risk type to reveal information and the uninformed to remain uninformed. Note that now since the insurers can identity the informed
from the uninformed, the, insurers offer the contracts \((\delta_H^*, \delta_U^*, \delta_L^*)\). The value of revealing information for the informed low risk types is

\[
 u^* = [V(\delta_L^*, \pi_L) - \gamma] - V(\delta_U^*, \pi_L)
\]

If \(\gamma < V(\delta_L^*, \pi_L) - V(\delta_U^*, \pi_L)\), then \(u^* > 0\), and all informed low risk reveal information. The value of getting informed for the uninformed is

\[
 I^* = \{p_H V(\delta_H^*, \pi_H) + p_L [V(\delta_L^*, \pi_L) - \gamma] - c\} - V(\delta_U^*, \pi_U)
\]

\[
 I^* = [p_L V(\delta_L^*, \pi_L) + p_H V(\delta_H^*, \pi_H) - V(\delta_U^*, \pi_L)] - (\gamma + c)
\]

The term in the square bracket is negative so \(I^* < 0\) and uninformed continue to remain uninformed. Thus \((\delta_H^*, \delta_U^*, \delta_L^*)\) is an equilibrium for \(\gamma < V(\delta_L^*, \pi_L) - V(\delta_U^*, \pi_L)\).

Step 4: Assume now that the insurers expect the informed low risk type not to reveal information and the uninformed to remain uninformed. In this case the insurers offer the contracts \((\delta_H^*, \delta_U^*, \delta_L^*)\). The value of revealing information for the informed low risk type is

\[
 u^* = [V(\delta_L^*, \pi_L) - \gamma] - V(\delta_U^*, \pi_L) = -\gamma
\]

Since the value of revealing information is negative, the informed low risk will not reveal information. The value of getting informed for the uninformed is

\[
 I^* = \{p_H V(\delta_H^*, \pi_H) + p_L [V(\delta_L^*, \pi_L) - \gamma] - c\} - V(\delta_U^*, \pi_U)
\]

\[
 I^* = [V(\delta_L^*, \pi_U) - V(\delta_U^*, \pi_L)] - (\gamma + c) < 0
\]

The value of information is negative, uninformed will choose to remain uninformed and the menu of contract \((\delta_H^*, \delta_U^*, \delta_L^*)\) will be an equilibrium.