

Welfare Effects of Banning Genetic Information in the  
Life Insurance Market:  
The Case of BRCA1/2 Genes

Michael Hoy\*  
and  
Julia Witt

Department of Economics  
University of Guelph  
Guelph, Ontario  
Canada N1G 2W1

April, 2005

**Abstract**

This paper is a contribution to the debate about whether regulations that ban insurance companies from access to individuals' genetic tests may lead in the near to medium term future to substantial adverse selection costs. We choose the specific possibility of widespread knowledge based on genetic testing for the so-called breast cancer (BRCA1/2) genes. We use a data set including economic, demographic, and relevant family background information to simulate the market for 10-year term life insurance targeted at women aged 35 to 39. Using standard welfare economic analysis for various information and regulatory scenarios concerning genetic test results, we find generally only modest adverse selection costs associated with such a regulatory ban. However, for family background groups which are at high risk for carrying one of the BRCA1/2 genes, the efficiency cost of adverse selection may be significant especially if a large fraction of women within such groups were to obtain genetic test results. These results, therefore, suggest some caution in developing regulations which protect individuals' genetic privacy.

\*e-mail: mhoy@uoguelph.ca

# 1. INTRODUCTION

The debate on whether insurance companies should be allowed to use results of genetic tests for the purpose of assessing differential prices for health, life, or disability insurance has been very lively and increasingly relevant as genetic technology and law making efforts are both progressing very rapidly.<sup>1</sup> The US senate in 2003 passed the so-called Genetic Information Nondiscrimination Act (GINA) by a unanimous vote and the bill is waiting attention by the House. Although GINA restricts the use of genetic test results only for health insurance and employment purposes, many other countries extend restrictions to life and long-term disability insurance.<sup>2</sup>

Arguments in favour of restricting the use of genetic test results for rate-making purposes are generally based on a concern with equity, while those who favour allowing insurance companies to use the information argue that not to do so would expose the market to serious adverse selection problems.<sup>3</sup> There is a limited empirical literature, mostly based on actuarial simulations that estimate possible price effects for insurance that would result from a ban on the use of genetic test results for rate-making purposes (e.g., Macdonald, 2003, and Macdonald and Pritchard, 2000, 2001). Although this is one important component of a full analysis of the associated adverse selection costs of such regulation, this work can be usefully supplemented by standard economic welfare analysis. In this paper we provide such an analysis for the specific case of information

---

<sup>1</sup> For example, the special issue of the North American Actuarial Journal (vol. 3, no. 1, Jan., 1999), which is devoted to papers on genetic technology and underwriting reflects such an interest.

<sup>2</sup>For a discussion on the variety of international regulatory responses concerning genetic information and insurance, see Lemmens, Trudo and P. Bahamin (1998), Lemmens (2000), and Knoppers, Godard, and Joly (2004). A recent editorial in *Nature Genetics* (May, 2004) discusses the US Genetic Information Nondiscrimination Act.

<sup>3</sup>See Kitcher (1996) for a good discussion of the equity arguments supporting prohibition of genetic information being used to risk-rate insurance. Also, see Hoy (1984), Hoy and Lambert (2000), and Bossert and Fleurbaey (2002). For a general discussion of efficiency concerns created by adverse selection in insurance see Crocker and Snow (1986). See Brockett and Tankersey (1997) for a discussion about regulation of genetic information in insurance markets.

relating to breast cancer; that is, the potential for genetic screening for the so-called breast cancer susceptibility genes (BRCA1-2 genes). The advantage of a standard welfare analysis is that it is based on a rational choice (utility-based) model that allows both for price effects from related demand analysis and efficiency measurement of alternative information scenarios and regulations.

We adopt a standard economic model of the life insurance market, based on Abel (1986), Brugiavini (1993), Villeneuve (1999, 2000), and Hoy and Polborn (2000). Insurance companies are assumed to offer nonexclusive contracts and to price these contracts linearly. This contrasts with the models involving price-quantity competition, pioneered by Rothschild and Stiglitz (1976) and Wilson (1977). Those models are well suited to property-liability insurance where the size of the loss is clearly identifiable and so the amount of insurance provided in the contract acts as an effective (self-) selection mechanism for different risk types. However, the economic cost associated with loss of life is not easily observable or objective. Moreover, individuals often change the amount of their life insurance holdings over time and through various heterogeneous products (e.g., whole life, group life, individual term insurance). Therefore, we develop a stylized model of the term life insurance market and simulate the effect of allowing consumers to hold private information (genetic test results) when deciding how much insurance to purchase.<sup>4</sup> We compare the implications of allowing consumers to hold this information as private to a *laissez faire* situation in which insurers do risk-rate premiums according to genetic test results. In the case of private information, higher risk individuals purchase more insurance than do lower risk individuals, hence driving the price higher than the population weighted actuarially fair price. From an efficiency perspective this means higher-risk types over-consume while lower-risk types under-consume, hence generating adverse selection costs. We measure the adverse selection costs as a standard deadweight loss based on the compensating variation measure.

---

<sup>4</sup> For empirical evidence on linearity of contracts and nonexclusivity of provision, see Cawley and Philipson (1999) and Pauly, et al. (2003). Cawley and Philipson (1999) even find concavity in the pricing of individual contracts (i.e., bulk discounting). This is explained by the existence of some contracting costs being fixed in relation to the amount of coverage for the policy. We abstract from such costs, although adopting a two-part tariff would not change our analysis qualitatively.

In the short-term future at least (say less than 5 years), it seems plausible that not many women will obtain genetic test results for the BRCA1/2 genes. One reason is cost. As Subramanian, et al. (1999) note, the cost of the test is very high (\$2,400) and since the test is patented the cost may not fall significantly for some time. Therefore, we consider the implications for adverse selection if only 5% of women become informed (i.e., take the test), if 20%, and if 100%; the last of these to cover the maximum adverse selection costs possible for this particular genetic test. We also break down our results by risk classes as defined by different family background for breast and ovarian cancer. Since higher risk classes as defined by family background imply different likelihoods for possessing the gene, this is an important consideration. In particular, women with more first-degree relatives having had breast or ovarian cancer, and at a younger age, will presumably feel more at risk and hence be more likely to obtain the genetic test.<sup>5</sup> Moreover, these classes also have a higher proportion of women who have the relevant cancer genes. Therefore, not surprisingly, we find that adverse selection costs can be substantially higher for classes represented by higher-risk family backgrounds.

We also use a variety of assumptions about risk preferences. We base our analysis of the demand for life insurance on the presumption that a family's objective is to maintain its per capita standard of living should an adult (income earning) member die.<sup>6</sup> In fact, we make explicit use of household equivalence scales to generate benchmark preferences for life insurance. However, we do assume that insurance purchasing is sensitive to price and so the above goal applies only if the individual considers the price to be actuarially fair. Within this framework, we allow for a variety of possible risk aversion levels. In this way we are able to provide extensive sensitivity analysis for our simulations. Our overall conclusion is that the size of adverse selection costs generated by a regulation prohibiting insurers from using genetic test results for the BRCA1/2 genes would probably be very modest in most circumstances. Thus, equity arguments that favour such regulation would not pale in comparison. However, for some higher risk family background types, if women in sufficient numbers obtain genetic tests,

---

<sup>5</sup> There is evidence that women at higher risk are more likely to obtain diagnostic testing (e.g., mammograms). See Picone, et al. (2004) and Witt (2004).

<sup>6</sup> See, for example, Bernheim, et al. (2003) and Gokhale and Kotlikoff (2002).

then adverse selection costs from such regulation could be substantial. This points to the possibility that as genetic information in society grows, there may come a point when genetic privacy may not be desirable. Thus, short to medium term moratoria on the use of genetic test results by insurance companies may be a more desirable policy framework than strict regulation through legislation that may be difficult to change in the future. In the long run, however, as information that is available to individuals about their own genetic predispositions grows and becomes more easily obtainable, regulations prohibiting insurers from access to genetic information may lead to substantial adverse selection costs. In this latter scenario, alternatives to banning the use of genetic test results by insurers for risk-rating premiums, such as limited public life insurance provision for high-risk types, may be more socially desirable.

The paper proceeds as follows. In section 2 we describe the basic underlying life insurance model. The application for this model is described in section 3, along with a brief discussion of the data sources and methods used to select parameters (see Appendix 1 for a fuller description). Results are discussed in section 4, with some extended results designed to provide sensitivity analysis in section 5. Summary and conclusions are given in the final section.

## **2. THEORETICAL BACKGROUND**

First we present the model we use to describe an individual's decision about how much life insurance to purchase.<sup>7</sup> We then show how market equilibrium is determined when there is a homogeneous group of consumers (i.e., individuals with the same probability of death), a scenario that applies under symmetric information conditions. This is followed by a discussion of equilibrium determination under conditions of asymmetric information (i.e., when the population is made up of individuals who have different probabilities of death and these are private information of consumers so that adverse selection occurs). In the following section we show how to apply these models to examine the welfare

---

<sup>7</sup> This model is based on Hoy and Polborn (2000).

implications of regulatory adverse selection in the case of genetic testing for BRCA1/2 genes.

An individual's probability of death,  $p$ , depends on family history and is assumed known to the individual. If the person has a genetic test for one of the BRCA1/2 genes, then the individual revises her assessment of this probability depending on the test result according to Bayes' law. The insurer may or may not know individuals' probabilities. Since this depends on legislation as well as inherent observability of "personal" information, we consider both cases of this information being used by firms and not being used by firms, the latter situation resulting in adverse selection.

The economic resources available to families are measured in net present value (NPV) terms, although for ease of discussion we generally simply refer to income. If the individual (woman) in question dies and no life insurance were purchased, the income earned by surviving family members is amount  $K$ , where  $K$  is the NPV of the income stream of the other adult household member(s) (generally the father or step-father). For single parent families  $K = 0$  (strictly speaking).<sup>8</sup> If the individual lives, income consists of the income of that individual,  $Y$ , as well as amount  $K$ . Life insurance can be purchased in amount  $L \geq 0$ , at price  $\lambda$ . If the individual purchases life insurance, then in the death state surviving family members have income  $K$  as well as income from the insurance claim, resulting in a net amount of  $K + (1 - \lambda)L$ . If the individual lives, income is  $Y + K - \lambda L$ ; that is, total income less the cost of life insurance purchased. Letting  $u(\cdot)$  and  $v(\cdot)$  represent von Neumann-Morgenstern elementary utility indices in the life and death states, respectively, with  $u'(\cdot), v'(\cdot) > 0$  and  $u''(\cdot), v''(\cdot) < 0$ , then the expected utility of the family as a function of insurance purchases is

$$EU(L) = (1 - p)u(K + Y - \lambda L) + pv(K + (1 - \lambda)L) \quad (1)$$

The family will maximize expected utility subject to  $L \geq 0$ , with first-order condition

$$-(1 - p)\lambda u'(K + Y - \lambda L) + p(1 - \lambda)v'(K + (1 - \lambda)L) \leq 0 \quad (2)$$

---

<sup>8</sup> Other sources of income could of course contribute to  $K$  and it is unlikely that the children of a single parent who dies would be left with no resources. We have no person-specific information about such resources and so in our simulations we set  $K$  equal to 5% of  $Y$  in these cases, an admittedly conservative assumption. We also added 30% of  $Y$  to each family's income (in the death state) as a robustness test and this did not change the qualitative nature of our results.

and second-order condition

$$(1-p)\lambda^2 u''(K+Y-\lambda L) + p(1-\lambda)^2 v''(K+(1-\lambda)L) < 0 \quad (3)$$

The second-order condition is always satisfied since we assume  $u''(\cdot) < 0$  and  $v''(\cdot) < 0$ .

For our simulations, we further assume families have constant relative risk aversion (CRRA) utility functions with degree of risk aversion  $\gamma$ :

$$u(X) = \frac{X^{1-\gamma}}{1-\gamma} \quad \gamma > 0, \gamma \neq 1 \quad (4)$$

$$u(X) = \ln X \quad \gamma = 1$$

Additionally, utility in the death state is a fraction of utility in the life state; i.e.

$$v(x) = a \cdot u(x) \quad a < 1$$

This is a simple way to parameterize the objective function in order to reflect the fact that when an adult family member dies the total income required to maintain per capita income levels for remaining members of the family falls. Note that this implies that, at equivalent total family income, marginal utility of income is less in the death state than in the life state.<sup>9</sup> Hence, expected utility is

$$EU(L) = (1-p) \left( \frac{(K+Y-\lambda L)^{1-\gamma}}{1-\gamma} \right) + pa \left( \frac{(K+(1-\lambda)L)^{1-\gamma}}{1-\gamma} \right) \quad (5)$$

with first-order condition

$$-(1-p)\lambda(K+Y-\lambda L)^{-\gamma} + pa(1-\lambda)(K+(1-\lambda)L)^{-\gamma} \leq 0 \quad (6)$$

and second-order condition

$$-(1-p)\lambda^2 \gamma (K+Y-\lambda L)^{-\gamma-1} - pa(1-\lambda)^2 \gamma (K+(1-\lambda)L)^{-\gamma-1} < 0 \quad (7)$$

From (6) it is straightforward to derive that the demand for insurance is the value of the following function:

$$L(a, p, K, Y, \lambda) = \frac{p^{1/\gamma} (1-\lambda)^{1/\gamma} a^{1/\gamma} (K+Y) - (1-p)^{1/\gamma} \lambda^{1/\gamma} K}{(1-p)^{1/\gamma} \lambda^{1/\gamma} (1-\lambda) + p^{1/\gamma} (1-\lambda)^{1/\gamma} a^{1/\gamma} \lambda} \quad (8)$$

---

<sup>9</sup> That is, with one less family member per capita income is higher in the death state for the same level of income that applies in the life state. Our method is consistent with the use of equivalence scales for standardizing income needs for families of different size, as described in the appendix.

Note that if  $a = 1$  [i.e.,  $v(\cdot) = u(\cdot)$ ] and  $Y > 0$ , then at zero insurance purchases ( $L = 0$ ) the marginal utility of income is higher in the death state than in the life state (since  $u''(\cdot) < 0$ ):

$$u'(K + Y) < v'(K) = u'(K) \tag{9}$$

If this were the case, then at actuarially fair pricing (i.e.,  $\lambda = p$ ) the family would purchase an amount of insurance so that income in the two states is the same. This is the standard result that applies for the case of state independent preferences and actuarially fair insurance pricing; i.e., that a risk averse individual chooses a quantity of insurance to equalize income across states of the world. However, if  $a < 1$  and insurance is priced at the actuarially fair rate, the optimal amount of insurance purchases which equalizes marginal utility across the two states implies less consumption in the death state than in the life state. This is a reflection of the presumption implicit in the choice of  $a < 1$  that economic needs for the family are lower when there is one less person in the family. Moreover, since  $u''(\cdot) < 0$ , the marginal utility in the life state is less the smaller is  $Y$ . Thus, if  $Y$  is sufficiently small relative to  $K$ , then given  $a < 1$  it is possible that marginal utility in the death state will be less than that in the life state even if no insurance is purchased. In this case, it is optimal for the individual not to purchase any insurance even if it is offered at the actuarially fair rate. We will see in our simulations that this is an important possibility.

We assume that insurers are risk neutral, face zero administrative costs, and are in a competitive price environment.<sup>10</sup> If insurers can observe family history and age of the individual, then they will also observe  $p$ , the individual's probability of death. Equilibrium in this situation is simply characterized by a price of  $\lambda = p$  for each consumer and demand determined by equation (8) above. In this case of symmetric information, we can allow for individuals to have different values of the various parameters ( $a$ ,  $K$ , and  $Y$ ) with no change to the determination of equilibrium price. If individuals vary according to probability  $p$ , but this is observed by insurers, then the same model applies except a different price is charged to each risk class.

---

<sup>10</sup> Also implicit in our price formation model is that insurers do not use exclusive contracts with consumers, which is the standard assumption in life insurance models (see footnote 4). Thus, linear pricing applies.

Now, suppose we allow for heterogeneity of consumers in risk type ( $p$ ) but assume that specific individual realizations of  $p$  are not observed by firms. We index risk type by  $r \in T$  where  $T = \{1, 2, \dots, t\}$ . Thus,  $p^r$  represents the probability of death for risk type  $r$  with  $p^1 \leq p^2 \leq \dots \leq p^t$ . Individuals also vary according to their inherent demand for life insurance, which depends on family composition and other considerations. We index demand type by  $d \in M$  where  $M = \{1, 2, \dots, m\}$ . Thus,  $a^d$  represents the taste parameter for demand type  $d$  with  $a^1 \leq a^2 \leq \dots \leq a^m$ . The equilibrium price when different risk types are pooled will be determined according to an interaction of these two dimensions of heterogeneity (i.e., risk type and demand type).

If insurers are privy to the same information about risk type as are consumers themselves, and are allowed legally to use this information, then the price charged will simply be the actuarially fair rate; that is,  $\lambda = p^r$ . Under these symmetric information conditions, the demand for insurance follows from equation (8) differentiated by risk and demand type, but with each individual belonging to a specific risk type being charged the same price. Thus, we can summarize this outcome in equation (10) below.

$$L^{rd} = L(a^d, p^r, K, Y, \lambda); \quad \lambda = p^r \quad (10)$$

Thus, the distribution of the parameter  $a^d$  within the population has no impact on the determination of price under symmetric information.

Now suppose insurers cannot charge separate prices to individuals of different risk types and must, instead, charge all individuals the same price  $\lambda$ . Let  $i = 1, \dots, N$  index all individuals, regardless of risk type or demand type. Each individual is distinguished by her risk type  $r$  and demand type  $d$ . Let  $p_i$  be the probability of death and  $a_i$  be the demand type parameter for individual  $i$ . Thus, for each  $i$  it follows that  $p_i = p^r$  for some  $r \in T$  and  $a_i = a^d$  for some  $d \in M$ , with risk type being private information, which in our application is due to regulation. We suppress the parameters  $K_i$ ,  $Y_i$ , and  $a_i$ , and denote the demand for life insurance by individual  $i$  as  $L^i(p_i, \lambda)$ . Thus, the total revenue from selling insurance to the entire group of individuals, where of course in some situations demand is zero, is  $TR(\lambda) = \sum_i \lambda L^i(p_i, \lambda)$  while expected total cost is

$ETC(\lambda) = \sum_i p_i L^i(p_i, \lambda)$ . Under conditions of perfect competition equilibrium price is determined by the zero expected profit condition,  $\Pi(\lambda) = TR(\lambda) - ETC(\lambda) = 0$ , and so

$$\lambda_e = \frac{\sum_i p_i L^i(p_i, \lambda_e)}{\sum_i L^i(p_i, \lambda_e)} \quad (11)$$

We are ensured that an equilibrium price  $\lambda_e$  exists provided at least one person of the highest risk type will demand a positive amount of insurance if charged the actuarially fair rate for that risk type (i.e., if  $\lambda = p^l$ ). The reason for this is that if  $\lambda = p^l$ , then expected profit,  $\Pi(\lambda)$ , will be negative since the firm just breaks even on sales to the lowest risk type individuals and all other contracts sold generate expected losses. If  $\lambda = p^l$  then the profit will be greater or equal to zero since the firm breaks even on sales to the individuals of the highest risk type and, if anyone of lower risk type purchases insurance, then the insurer will earn positive expected profits on those contracts. Since all component functions of  $\Pi(\lambda)$  are continuous in  $\lambda$ , then so is  $\Pi(\lambda)$  continuous. Thus, the function  $\Pi(\lambda)$  must equal zero for at least one value of  $\lambda$ . The lowest such value is then the equilibrium price  $\lambda_e$ .<sup>11</sup>

### 3. APPLICATION

Our application is a simulation of the market for 10-year term life insurance for Canadian women in the age group 35 to 39. We use data from the 1999 Canadian National Population Health Survey (NPHS), Health Component Data Set, that provides both detailed health information, including family background of breast cancer, as well as relevant economic and demographic information that affect demand for life insurance. The survey does not include information about life insurance purchases, and this is simulated using a standard expected utility model. Risk factors (family background) that determine the probability of a woman getting breast cancer as well as being positive for one of the BRCA1/2 cancer genes is taken from the health survey and used in the

---

<sup>11</sup> If there is more than one value of  $\lambda$  for which  $\Pi(\lambda) = 0$ , then the lowest such value is the equilibrium price since any higher price would be undercut by some firm(s). A formal proof would be essentially the same as the analysis in section 2 of Hoy and Polborn (2000).

program *CancerGene* to determine both the probabilities of getting cancer and the probability that if the woman were to obtain a genetic test for the BRCA1/2 genes that she would test positive. This allows us to explore how the insurance market would behave in the various scenarios discussed earlier in this paper; in particular, with and without a regulation which bans insurers from using genetic test results known to insureds.

Here we provide a sketch of how this information was utilized in generating simulations of the insurance market in both scenarios of symmetric and asymmetric information regarding genetic test results. A more detailed explanation of the computations and assumptions involved in the simulations is provided in the appendix.

The variables needed to describe the insurance market are  $\gamma, p_i, Y_i, K_i, \alpha_i, q_i, p_i^{neg}, p_i^{pos}$ , where person-specific information is indicated by subscript  $i$ . Although the risk preference parameter,  $\gamma$ , would realistically vary across individuals, we assume it is the same for all. Not only would it be a substantial complication to depart from this assumption, but we have no person-specific information about this risk parameter. We do, however, vary this parameter in our robustness tests and so can demonstrate how our results depend on this parameter.

First we establish a 10-year death probability from all causes (i.e., the probability that a woman in the age group 35 to 39 will die within the next ten years) as a benchmark by using the CDC Life Tables (Arias, 2000). We then deduct from this value the probability that a woman would die of breast cancer over this period (i.e., not conditioned on family background). This is our benchmark 10-year probability of death due to all causes except breast cancer. Then, using the *CancerGene* program and the family background information from the NPHS data set mentioned above, we create 13 risk categories according to family background (e.g., based on whether any first- or second-degree relatives of a woman had breast cancer).<sup>12</sup> The death probabilities are denoted simply by  $p_j, j=1,2,\dots,13$ ; that is, each person  $i$  belongs to one of these risk groups.

---

<sup>12</sup> The probability of getting breast cancer is taken from Ford, et al., 1998, and the 10 year survival rate for women getting breast cancer is taken from the SEER Cancer Registry (see Ries, et al., 2004), based on all women under the age of 50.

If individuals obtain genetic test results, then their probability of dying from breast cancer of course depends on both family background and whether the person tested positive or negative. Integrating these differential probabilities of death from breast cancer with the probability of death from all other causes leads to probabilities of death denoted  $p_j^{pos}$  and  $p_j^{neg}$ , respectively. Summary data describing the thirteen risk groups and the relevant probabilities for those who test positive and negative, should test results be obtained, is given below in Table 1. Also included is the probability that a woman from a given risk group will test positive ( $q$ ) for one of the BRCA1/2 genes and the number of individuals in each group. (NOTE: We suppress subscripts  $i$  or  $j$  whenever it will not lead to confusion.)

*Table 1: Risk Groups*

Group	$p$	$p^{neg}$	$p^{pos}$	$q$	# ind's per group
1	0.017340	0.017309	0.048413	0.001	491
2	0.017344	0.017313	0.048417	0.001	157
3	0.017476	0.017269	0.086038	0.003	4
4	0.017483	0.017208	0.086037	0.004	21
5	0.017494	0.017287	0.085975	0.003	2
6	0.017500	0.017294	0.085982	0.003	5
7	0.017584	0.017240	0.086009	0.005	1
8	0.017608	0.017195	0.086005	0.006	2
9	0.017769	0.017219	0.086018	0.008	8
10	0.017814	0.017194	0.086017	0.009	1
11	0.018088	0.017193	0.086000	0.013	6
12	0.020125	0.017165	0.086007	0.043	1
13	0.021480	0.016994	0.086010	0.065	1

As can be seen from Table 1, there is very little difference in the overall probability of death for family background types 1 through 10. However, there is substantial difference in the relative likelihoods of women from these groups testing positive for the breast cancer gene (from 0.001 in group 1 to 0.009 in group 10 and up to 0.065 in group 13) and it is this latter factor that is an important driver of the cost of adverse selection when insurers are not allowed to use genetic test results for rate-making.

Incomes of the household and the woman in particular are available from the data set, although in group form. This information is used to generate the important economic

variables  $Y_i$  and  $K_i$ . Details on how grouped information was handled are in the appendix. The functional form of the utility function is that of constant relative risk aversion, as mentioned earlier. We used three different values for the degree of relative risk aversion; namely 0.5, 1 and 3. We believe this reflects a reasonable range as suggested by the empirical and experimental literature on risk-taking under uncertainty (see Blake, 1996).

Finally, we use a benchmarking exercise to determine the parameter “ $a$ ” for the death state utility function,  $v(x) = a \cdot u(x)$ . The purpose of this exercise is to find a utility function so that the “average consumer” of insurance buys that amount of insurance, when offered at the actuarially fair rate, so that the household’s standard of living remains unchanged should she die and her income be lost to the family. If insurance is priced higher than the actuarially fair rate, which happens in the presence of adverse selection for individuals who test negative for the BRCA1/2 genes, then demand is less. Details of how these values are calculated for individuals from families with different size and composition are provided in the appendix.

Notice in Table 1 that for most risk categories there are very few observations. Thus, for each risk group we simulate the insurance market based on socio-economic information of all 700 observations available rather than use only the individuals in each group. As long as socio-economic information is not correlated with risk factors (i.e., family background of the woman with respect to breast cancer incidence), this process is not of concern. To check on this we did separately simulate the market for each of the risk categories 1, 2, and 4 using only the observations from their respective groups rather than the entire set of observations. We obtained qualitatively similar results.

Below we describe each of the five sets of simulations we do, which we refer to as modules 1 through 5. Each module refers to a different information scenario regarding whether individuals have taken genetic tests and whether insurers are banned from using genetic test results in rate-making. In each module we compute the efficiency gain from insurance by comparing the per capita compensating variation ( $CV$ ) that results from the opportunity to purchase insurance in the given information scenario. That is, we compute the value of  $CV$  that equates the utility an individual would receive with no insurance (and no genetic test information) to expected utility if insurance is available (conditional

on a given information scenario) with income in both states of the world reduced by amount  $CV$ . Note that insurance availability always leads to a higher expected utility level and so for all of our modules, or information scenarios, the per capita  $CV$  will be positive. It is the difference in the levels of the per capita  $CV$ , however, that we use to compare the relative efficiency of the market in the different modules. In particular, this allows us to determine a measure of inefficiency due to adverse selection for the case in which insurers are banned from using genetic test results for rate-making.

Another measure that is often reported as a measure of the impact of adverse selection is the amount by which the price of insurance rises, relative to the average probability of death (also the population weighted actuarially fair price), due to adverse selection. We also report this value for each risk group (i.e.,  $\lambda_e^j / p_j$  for each group  $j$  where  $\lambda_e^j$  is the equilibrium price under adverse selection).<sup>13</sup>

We now describe more formally the five modules or information scenarios. In all cases insurance companies are assumed to observe and be allowed to use family background information in their pricing behaviour. Thus, the only difference between modules involves genetic tests.

## MODULE 1

*Individuals have no genetic test information.*

In this case the price of insurance is actuarially fair according to family background; i.e.,  $\lambda = p_j$ ,  $j = 1, 2 \dots, 13$ . There are 13 risk groups, hence 13 different prices. This case provides an interesting benchmark with which to compare our other results. Moreover, it reflects well from an informational perspective the existing insurance market since few people now currently obtain genetic test results.

## MODULE 2

*All individuals obtain genetic tests and insurance companies are allowed to use this information in setting prices (i.e., in addition to family background information).*

---

<sup>13</sup> Note that  $\lambda_e/p_j = 1$  if there is no adverse selection.

Again, from individuals' perspectives the price of insurance is actuarially fair. However, within each of the family background groups,  $j = 1, 2, \dots, 13$ , there are now two prices. One price applies to those who test positive for having one of the BRCA1/2 genes, in which case  $\lambda = p_j^{pos}$ , while a second price applies to those who test negative (i.e.,  $\lambda = p_j^{neg}$ ). The fraction of individuals who test positive,  $q_j$ , depends on family background and so we write  $q_j$  to represent this fraction for individuals from risk (family background) group  $j$ ,  $j = 1, 2, \dots, 13$ .<sup>14</sup> Thus, this scenario reflects a situation in which each individual knows whether she has one of the breast cancer genes BRCA1/2 and insurers are also privy to this information.

### MODULE 3

*All individuals obtain genetic tests and insurance companies are **not** allowed to use this information in setting prices.*

Since we presume in this scenario that every individual obtains information about whether or not she has one of the BRCA1/2 genes and that insurers are not allowed to use this information for rate-making purposes, then this case represents the maximum possible effect of private information on adverse selection costs created by this genetic test and a ban on its use.

Pricing is not so straightforward as in the previous two modules as, *ceteris paribus*, a person who tests negative will want to reduce her purchases of life insurance while a person who tests positive will want to increase her purchases. Thus, the price that persists will be higher than the population weighted death probability (i.e.,  $\lambda_e > p$ ). The price  $\lambda_e$ , which is referred to as the average clientele risk, is determined according to a specific application of equation (11). In the context of the information set for this module, this means finding the minimum value of  $\lambda$  that satisfies equality of expected total revenue (ETR) and expected total cost (ETC) for each of the 13 risk groups ( $j = 1, 2, \dots, 13$ ). We will denote this value by  $\lambda_e^j$ . Thus, letting  $n_j$  represent the number of individuals in risk group  $j$  and indexing these individuals by index  $i$ , ( $i = 1, 2, \dots, n_j$ ) and

---

<sup>14</sup> Obviously the fraction that test negative is  $1 - q_j$ .

$q_j$  the probability that any individual in this risk group is determined by the genetic test to be carrying one of the BRCA1/2 genes, we have:<sup>15</sup>

$$ETR(\lambda_e^j) = \lambda_e^j \cdot \sum_{i=1}^{n_j} [q_j \cdot L^i(p_j^{pos}, \lambda_e^j) + (1 - q_j) \cdot L^i(p_j^{neg}, \lambda_e^j)] \quad (12)$$

$$ETC(\lambda_e^j) = \sum_{i=1}^{n_j} [q_j \cdot p_j^{pos} \cdot L^i(p_j^{pos}, \lambda_e^j) + (1 - q_j) \cdot p_j^{neg} \cdot L^i(p_j^{neg}, \lambda_e^j)] \quad (13)$$

As discussed in section 2 (Theoretical Model) we know some such value  $\lambda_e^j \in (p_j^{neg}, p_j^{pos}]$  exists. To compute this value we simply begin with the value  $\lambda = p_j^{neg}$  and increase it in “small increments” until  $ETR = ETC$ . We know that high risk types (i.e., those who test positive) will purchase more insurance than will low risk types (i.e., those who test negative), and so  $\lambda = p_j^{neg}$  will generate negative expected profits ( $ETR < ETC$ ). For  $\lambda = p_j^{pos}$  we will have  $ETR > ETC$  as long as at least one person who tests negative buys some positive amount of insurance, otherwise  $ETR = ETC$ . Since all functions that make up  $ETR$  and  $ETC$  are continuous in the price  $\lambda_e^j$ , a minimum price that establishes equality between  $ETR$  and  $ETC$  (i.e., zero expected profits) must exist and so be found (approximately) by our iterative method. Of course, it is possible that all those who test negative will be driven out of the market and we end up with  $\lambda_e^j = p_j^{pos}$ . If, however, at least some persons who test negative would purchase insurance in equilibrium, we will have  $p_j^{neg} < \lambda_e^j < p_j^{pos}$ .

#### MODULE 4

*In this module, information is again asymmetric, as individuals can obtain a genetic test and insurance companies cannot ask for the results. However, not all individuals obtain a genetic test. We consider two cases: in the first case we assume that only 5% of individuals take a genetic test while in the second case it is assumed that 20% obtain a genetic test.*

---

<sup>15</sup> As earlier, we suppress the variables  $a_i, Y_i, K_i$  in  $L^i(\cdot)$

The structure for price determination is essentially the same as that of Module 3 except that insurance companies, when they calculate  $\lambda$  for each group, now take into account that only 5% (20%) will know their true genetic status, hence only  $5\% \cdot q$  ( $20\% \cdot q$ ) will actually know that they are a high risk type. As mentioned earlier, these cases we believe are more realistic scenarios for the near-term future as the cost of obtaining a genetic test for the BRCA1/2 gene is very high and so it may well be the case that only a relatively small fraction of the population will take such a test in the near-term future. With fewer people holding private information in this module, relative to Module 3, we expect and find a smaller adverse selection effect.

## MODULE 5

*In this module, information is again asymmetric, as individuals can obtain a genetic test and insurance companies cannot ask for the results. We consider the specific case in which 20% of individuals take a genetic test. Although insurance companies cannot ask for test results, we assume that those who test negative may present their test result to the insurer in order to obtain an actuarially fair (lower) insurance premium.*

This module reflects the possibility, allowed by some regulatory regimes, for insurers to accept negative test results and price accordingly but not to allow insurers access to positive test results. Thus, those who test negative, and who make up  $20\% \cdot (1-q)$  of the population within a risk group, will receive price  $\lambda = p_j^{neg}$ . Those who test positive will be pooled with those who remain uninformed. Insurance companies offering insurance to those in this pooled group face a group composed of fraction  $20\% \cdot q$  who are informed that they have death probability  $p_j^{pos}$  while the uninformed (composed of 80% of the population within the overall group) will have death probability  $p_j$ . The extent of adverse selection for this pooled group is in fact worse than in Module 4 when 20% of people get tested because nobody who tests negative will be in this group (i.e., only uninformed individuals and those who test positive form this pool).

For each of the modules we compute the compensating variation (CV) of the insurance opportunity implicitly by equating expected utility with no insurance,  $EU_{no}$ , to

expected utility with insurance but with an amount of income taken away ( $CV$ ) in both states of the world. Thus, we have for all modules:<sup>16</sup>

$$EU_{no}^j = (1 - p_j)u(K + Y) + p_jv(K) \quad (14)$$

The precise method of determining  $CV$  varies across the modules since individuals hold different information sets. In Module 1 individuals within each risk group face the same probability of death and so  $CV$  is determined according to:

$$EU^j(CV^j) = (1 - p_j)u(K + Y - \lambda L - CV^j) + p_jv(K + (1 - \lambda)L - CV^j) = EU_{no}^j \quad (15)$$

where  $L$  is the optimal amount of insurance demanded. In Modules 2 and 3 we determine separate values for the  $CV$  for those who test positive and for those who test negative. In Module 4 we determine the  $CV$  for three types; namely those who test positive, those who test negative, and those who don't test at all. A similar process is applied in Module 5 noting that those who test negative are not pooled with the others and they receive price  $\lambda = p_j^{neg}$ . In each Module we compute the per capita  $CV$  value and use this as our relative measure of efficiency. Note that the per capita  $CV$  in each Module is based on the same benchmark value of expected utility with no insurance. This is important since if one were not to do this one would be led to rather odd conclusions. For example, if suddenly the probability of death were to become zero for all individuals, an admittedly unrealistic scenario, the willingness to pay for insurance would become zero. But this of course does not mean people become worse off.<sup>17</sup>

#### 4. RESULTS

In this section our results, which we present in detail, are based on the elementary utility function  $u(x) = \ln(x)$ . In section 5 we provide a summary table of the results obtained from using the other utility functions mentioned earlier.

---

<sup>16</sup> We suppress individual subscript  $i$  whenever doing so doesn't lead to confusion.

<sup>17</sup> We also computed the equivalent variation for the insurance market opportunity. Although we do not report these values here, we obtained qualitatively similar results.

The average amount of insurance purchased, and the participation rate, does not vary much across risk groups in Module 1. In each of the risk categories 70.3% of women purchase insurance and, among those who participate in the market, average demand for insurance varies between 390,400 and 390,900 across the risk groups, which implies an average replacement rate of 45% (i.e., women who purchase insurance on average buy an amount which would replace 45% of their income for their family's use should they die).

In Module 2 individuals continue to face actuarially fair prices, although of course those who test positive face a higher price reflecting their correspondingly higher probability of death. Since only 0.1% of this risk group would test positive, the probability of death changes very little for those who test negative relative to the pooled probability from Module 1. Not surprisingly, those who test negative have very similar demand behaviour as they would in Module 1. Those who test positive demonstrate a slight rise in the amount of insurance, varying from 394,200 to 398,900 across risk groups, for an average replacement rate of 45.5%. Thus, although a state contingent utility function allows for the amount of insurance to vary as the probability of death varies, even under the assumption of risk-type specific actuarially fair pricing, we see that demand behaviour is quite stable for the range of variations in  $p$  that is relevant to our scenarios. We also find that the introduction of information concerning whether an individual has or does not have one of the BRCA1/2 genes does not change insurance buying behaviour very much as long as insurers are allowed to use this information for rate-making purposes.

In Table 2 below we provide the per capita compensating variation value of the insurance market for each risk group for both Modules. Two patterns are of interest. First, the value of the insurance market opportunity is higher the higher is the probability of death within each of these modules, which is not surprising. Also, for a given risk group, the per capita  $CV$  is slightly higher in Module 1 than in Module 2. The reason is that, although under symmetric information there is no loss of efficiency due to adverse selection in Module 2, there is premium risk associated with the introduction (and use by

insurers) of genetic test results. Since  $q_j$  is very small, it is not too surprising that this premium risk reduces welfare as measured by the  $CV$  value only by a small amount.<sup>18</sup>

Table 2: Per Capita CV for each risk group for Modules 1 and 2

Group	Per Capita CV Module 1	Per Capita CV Module 2
1	4,017.3	4,016.8
2	4,018.3	4,017.6
3	4,048.1	4,047.5
4	4,050.4	4,046.8
5	4,053.4	4,052.8
6	4,054.4	4,052.6
7	4,073.1	4,072.8
8	4,078.3	4,077.3
9	4,114.7	4,109.6
10	4,124.9	4,122.6
11	4,186.9	4,184.5
12	4,646.1	4,633.1
13	4,950.1	4,932.4

### MODULE 3

Recall that Module 3 represents our scenario with the greatest potential for adverse selections costs. All individuals are presumed to obtain genetic test results and insurers are not allowed to use that information to risk-rate insurance policies. The result is that those who find that they are lower risk types (i.e., test negative) reduce their insurance purchases while those who find they are higher risk types (i.e., test positive) increase insurance purchases. Table 3 shows the pooled price for each risk type within each risk group (i.e., as defined by family background which is still used in risk-rating). In risk groups 1 and 2 the ratio  $p^{pos}/p^{neg}$  is substantially lower than for the other groups. Moreover, the fraction of individuals who would test positive is only 0.001. For these two categories ( $j=1,2$ ) the impact of adverse selection is to increase the price to a level

<sup>18</sup> Notice from Table 1 that, roughly speaking,  $q$  increases through the risk categories 1 to 13, as does the reduction in  $CV$  associated with moving to the scenario of Module 2. In other words, the efficiency loss due to premium risk is increasing in  $q$ .

only 1.5% above the population weighted actuarially fair price (or probability of death). The price increases more for groups 3 – 13, rising to a substantial price increase (relative to the average price in Modules 1 and 2) of almost a factor of 3.

Table 3: Price Effects of Adverse Selection from Module 3

Group	$p$	$p_{neg}$	$p_{pos}$	$\lambda$	$\lambda/p$
1	0.017340	0.017309	0.048413	0.017609	1.015513
2	0.017344	0.017313	0.048417	0.017613	1.015510
3	0.017476	0.017269	0.086038	0.021469	1.228484
4	0.017483	0.017208	0.086037	0.023208	1.327460
5	0.017494	0.017287	0.085975	0.021487	1.228249
6	0.017500	0.017294	0.085982	0.021494	1.228228
7	0.017584	0.017240	0.086009	0.025140	1.429708
8	0.017608	0.017195	0.086005	0.027095	1.538788
9	0.017769	0.017219	0.086018	0.030519	1.717541
10	0.017814	0.017194	0.086017	0.032194	1.807230
11	0.018088	0.017193	0.086000	0.038093	2.105987
12	0.020125	0.017165	0.086007	0.057565	2.860394
13	0.021480	0.016994	0.086010	0.063495	2.955983

Not surprisingly, looking at average insurance purchases by lower and higher risk types within these risk groups displays an expected pattern with average demand by lower risk types (i.e., those testing negative) falling drastically compared to Modules 1 and 2 for the higher risk groups.. The participation rate for those who test negative for the BRCA1/2 genes in risk group 13 falls from 70.3% in Modules 1 and 2 to 21.3% in the presence of adverse selection. Moreover, for those who test negative and continue to purchase insurance, their replacement rate falls from 45% to 12.4%. These results are summarized in Table 4 below. Note that, conversely, those who test positive have a revised probability of death which implies insurance is priced at less than the actuarially fair price. As a consequence, these higher risk types within each risk category purchase more insurance than they would have under risk-type specific actuarially fair pricing. Their replacement rates range from 0.92 to 7.2, implying substantial degrees of over-insurance in some cases. Of course, even those who test positive for life insurance are not

insensitive to price and so even though the ratio  $p^{pos}/p^{neg}$  is very large for risk groups 12 and 13, the fact that participation in the insurance market by those who test negative is very low leads to a relatively high price and a lower degree of over-insurance by those who test positive.

Table 4: Demand Effects of Adverse Selection in Module 3

Group	Test Group	Avg. Amt LD	Avg. Replacement Rate	% with LD > 0
1	Low (NEG)	394,277	0.450	66.6
2		394,281	0.450	66.6
3		310,818	0.343	50.1
4		275,499	0.305	46.4
5		310,969	0.344	50.1
6		311,027	0.344	50.1
7		252,096	0.281	42.3
8		265,009	0.298	34.0
9		221,125	0.255	32.4
10		209,723	0.245	30.7
11		172,809	0.212	26.9
12		109,843	0.145	21.3
13		94,405	0.124	21.3
1	High (POS)	2,003,352	4.120	100.0
2		2,002,950	4.119	100.0
3		3,346,135	7.198	100.0
4		3,036,453	6.490	100.0
5		3,339,541	7.183	100.0
6		3,338,551	7.181	100.0
7		2,741,281	5.814	100.0
8		2,486,171	5.230	100.0
9		2,118,326	4.388	100.0
10		1,966,416	4.040	100.0
11		1,536,431	3.055	100.0
12		749,960	1.236	98.4
13		622,562	0.920	95.3

In Table 5 we show the per capita  $CV$  computations for each risk group associated with the insurance market opportunity when adverse selection is present. Not surprisingly we find a reduction in per capita  $CV$  in comparison to that in Module 2 in

which insurers are allowed to risk-rate according to genetic test results. The loss in market efficiency is, roughly speaking, greater for those cases where adverse selection has a greater effect on price.

*Table 5: Per Capita CV for each risk group for Module 3*

Group	Per Capita CV
1	3,963.2
2	3,964.0
3	3,395.3
4	3,185.9
5	3,400.1
6	3,394.7
7	3,027.7
8	2,867.5
9	2,673.0
10	2,593.8
11	2,380.1
12	2,389.2
13	2,669.6

Notice that for risk groups 1 and 2, where the price increase due to adverse selection is about 1.5%, the loss of efficiency (relative to Module 2) is modest at approximately 1.3%. The loss of efficiency, however, rises to as much as 43% for risk group 11. This comparison across risk groups demonstrates that one needs to consider separately different risk groups according to family background when measuring efficiency effects arising from a ban on insurers using genetic test results for risk-rating purposes

#### MODULE 4

In this module we reconsider the effects of adverse selection when less than 100% of the population obtains genetic test results. In particular, we examine the cases for  $x = 5\%$  and  $x = 20\%$  of the population obtaining such information. This means that a substantial proportion of the population remains uninformed about their genetic type and so with fewer individuals knowing they are high risk types, the impact of adverse selection is not

as great as it was for Module 3. Below we compare the effect on price of adverse selection of these three experiments.

*Table 6: Price Effects of Adverse Selection from Module 4*

Group	$\lambda/p$ ( $x = 5\%$ )	$\lambda/p$ ( $x=20\%$ )	$\lambda/p$ ( $x=100\%$ ) <i>Module 3</i>
1	1.001096	1.002826	1.015513
2	1.001096	1.002825	1.015510
3	1.011044	1.039654	1.228484
4	1.012870	1.052908	1.327460
5	1.011032	1.039657	1.228249
6	1.011086	1.039657	1.228228
7	1.014559	1.065742	1.429708
8	1.021978	1.078771	1.538788
9	1.025325	1.104114	1.717541
10	1.026945	1.116762	1.807230
11	1.038976	1.177189	2.105987
12	1.101366	1.528695	2.860394
13	1.135661	1.666390	2.966983

Not surprisingly, the results in Table 6 demonstrate that reducing the fraction of privately informed individuals reduces the degree of adverse selection as indicated by the extent to which the equilibrium price exceeds the population weighted probability of death (i.e., the ratio  $\lambda/p$ ). Similarly, it is not surprising that the demand effects also are less when a smaller fraction of insureds are privately informed. The comparative results on demand effects from cases in Modules 3 and 4 are summarized in Table 7 below. Also, the loss of efficiency as measured by the per capita CV measure (Table 8) is less when a smaller fraction of the population holds private information about risk type.

Table 7: Ranges of Demand Effects for Modules 3 and 4

Module	Test Status	Average Replacement Rate		% with LD > 0	
		Group 1	Group 13	Group 1	Group 13
3	Negative	0.450	0.124	66.6%	21.3%
	Positive	4.120	0.920	100%	95.3%
4 (x=20%)	Uninformed	0.448	0.282	69.7%	32.6%
	Negative	0.445	0.223	69.7%	27.6%
	Positive	4.201	3.401	100%	100%
4 (x=5%)	Uninformed	0.450	0.346	69.7%	63.1%
	Negative	0.444	0.291	69.7%	42.3%
	Positive	4.212	6.061	100%	100%

Table 8: Per Capita CV for each risk group for Module 4

Group	Per Capita CV (x=5% tested)	Per Capita CV (x=20% tested)
1	4,016.0	4,009.4
2	4,016.6	4,010.8
3	4,007.8	3,909.4
4	4,004.5	3,865.5
5	4,011.4	3,913.1
6	4,013.1	3,913.1
7	4,022.2	3,846.3
8	3,996.2	3,810.7
9	4,022.4	3,768.4
10	4,028.3	3,743.6
11	4,045.7	3,632.5
12	4,274.5	3,310.9
13	4,442.3	3,343.5

## MODULE 5

In this module we reconsider the effect of a regulation that (i) allows individuals who have had a negative test result to present that result to an insurer and to receive a corresponding actuarially fair price but (ii) prohibits insurers from asking those with positive test results to disclose them. We consider one of the cases of Module 4 (i.e., 20% of the population obtains a test). Thus, only those who test negative will declare their test results and, as noted above, these people receive price  $\lambda = p_j^{neg}$ . Those who test positive keep this information private and are pooled with the uninformed consumers, thus creating an adverse selection scenario which will be more intensive than in the similar situation in Module 4, as described earlier. The comparison of the extent of adverse selection in this case and with that of Module 4 is illustrated in Table 9 below.

*Table 9 Price Effects of Adverse Selection from Modules 4 and 5*

Group	$\lambda/p$ ( $x = 20\%$ ) Module 5	$\lambda/p$ ( $x=20\%$ ) Module 4
1	1.003403	1.002826
2	1.003402	1.002825
3	1.051099	1.039654
4	1.070068	1.052908
5	1.051046	1.039657
6	1.051086	1.039657
7	1.082802	1.065742
8	1.101488	1.078771
9	1.137880	1.104114
10	1.156057	1.116762
11	1.232474	1.177189
12	1.647951	1.528695
13	1.782779	1.666390

Thus, adverse selection conditions are in a sense worsened by allowing people who test negative to have their test results used to reduce their insurance price and this is relatively more important for higher risk groups. Although we don't present the results here, not

surprisingly people in these adverse selection scenarios purchase less insurance coverage. However, those who do test negative and receive price  $\lambda = p_j^{neg}$  avoid adverse selection altogether and are clearly made better off by the opportunity provided by a regulation that allows people the freedom to present negative test results to the insurer.

We see from the above computations from the various modules that the smaller the fraction of the population that takes genetic tests, and hence the smaller the fraction of the population that holds private information about their risk type, the less the impact of adverse selection. This holds regardless of whether one looks at demand effects (i.e., price effects and quantity demanded effects) or our efficiency measure based on the *CV* value. Although this is hardly surprising, we believe the exercise provides at least some rough sense of how important the size of this particular parameter is from a regulatory perspective. This is especially true if one were to consider the potential availability of genetic tests for a wide variety of diseases which have some significant impact on longevity. For this reason, a moratorium on insurers using genetic test results to risk-rate life insurance policies may make sense for the short run given the counter-arguments concerning use of this information. Given the greater difficulty to change legislation than to review terms of a fixed-term (renegotiable) moratorium, we believe that legislation banning insurers from using genetic test results is too inflexible a response to the justifiable concerns about the potential use of this information because of the possibility that adverse selection costs could eventually become quite important.

## **5. SENSITIVITY ANALYSIS**

A more detailed description of results from various changes in our assumptions is available in a technical appendix. However, due to its length it is not included here. There are many assumptions which deserve attention from the point of view of sensitivity analysis. Perhaps chief among these is the accuracy of the probabilities used to determine likelihoods of getting cancer for women of different family backgrounds and whether they have one of the BRCA1/2 genes. Also important is the determination of the 10-year survival rate for women diagnosed with cancer and the fraction of women,

conditional on family background, who are assumed to have one of the BRCA1/2 genes. However, simply by the fact that we have a wide variety of such probabilities across our family backgrounds provides a good sense of how sensitive operation of our life insurance model is to these factors. We see that the level of insurance demanded in our model is not very sensitive to the probabilities of death (and hence, indirectly of getting cancer) when symmetric information applies.

However, as  $q$ , the fraction of individuals in the risk group (as determined by family background) to have one of the BRCA1/2 genes varies across these groups, we can see significant sensitivity as to how the impact of adverse selection is realized in the insurance market. This is also true for the relative size of the probabilities of death for those with and without one of the BRCA1/2 genes (i.e., the ratio  $p^{pos}/p^{neg}$ ). The insurance market behaviour under adverse selection is also very sensitive to the proportion of individuals in a risk group who obtain information from genetic testing. In fact, identifying these sensitivities is, we believe, the main contribution of this paper.

For our presentation of results in Section 4, we assumed the particular elementary utility function representing constant relative risk aversion (CRRA) preferences with degree of risk aversion of one (i.e., the logarithmic utility function). We also used constant relative risk aversion utility functions with degrees  $\frac{1}{2}$  (i.e., less risk averse) and 3 (i.e., more risk averse). In terms of insurance market behaviour under symmetric information a lower degree of risk aversion leads to a lower demand for insurance and a lower value, in terms of  $CV$ , for the insurance market opportunity. By comparing the results of Module 1 for these three assumptions on risk preferences, we do indeed find this to be the case. Most significant, perhaps, is the fact that the value of insurance is significantly affected by our assumption about the degree of risk aversion.

*Table 10: Demand Sensitivity to Choice of Risk Preferences (Module 1)*

Degree of CRRA	% participation	Ave. Demand	Per Capita CV
CRRA = $\frac{1}{2}$	44.9	252,500 to 253,100	625 to 768
CRRA = 1	70.3	390,400 to 390,500	4,017 to 4,950
CRRA = 3	82.7	581,900 to 582,200	37,250 to 41,700

When an individual faces an actuarially fair price for life insurance the “rule” that determines optimal insurance demand is to purchase that amount of insurance that equates marginal utility of income across states of the world (i.e., the life and death states).<sup>19</sup> Another important implication of the degree of risk aversion is that when price varies from the actuarially fair rate, the higher is the individual’s degree of risk aversion the less an individual wishes to deviate from the “rule” that determines demand when she is faced with the actuarially fair price. Thus, one might expect that when information is asymmetric and adverse selection occurs, the ultimate equilibrium price will be higher (i.e., the greater will be the impact of adverse selection on price) the lower is the degree of risk aversion. The reasoning can be explained by considering what happens in the thought experiment when we increase the price of insurance from the population weighted probability of death in order to cover the cost of adverse selection (i.e., to arrive at a price so that firms earn zero expected profit rather than negative expected profit). As the price rises, the low risk types who have tested negative for the gene tend to reduce their purchases of insurance, thus creating a wedge between the state-contingent consumption levels that would equate marginal utility of income across states. But the more risk averse they are the less willing are they to do this. So they don’t decrease demand for insurance as much as they would if they were less risk averse. For those who test positive and so recognize they are higher risk types, the converse applies. They purchase more insurance than under symmetric information because the price is favourable relative to their higher perceived probability of death. Thus, they over-insure and their final consumption vector across states of the world represents a wedge relative to that consumption vector that would equate marginal utility of income across states. The extent to which they are willing to over-insure is weakened by being more risk averse and so they respond less in terms of increased demand for insurance than if they were less risk averse. Both of these effects will reduce the impact of adverse selection on the resulting price of insurance (i.e., the fact that greater risk aversion leads to less flight from insurance by low risk types and less over-insurance by high risk types). The results in Table 10 demonstrate this effect.

---

<sup>19</sup> This is a straightforward implication of the fundamental theorem of risk bearing .

Table 10: Impact on Adverse Selection from Choice of Risk Preferences (Module 3)

Risk Group	$\lambda/p$	$\lambda/p$	$\lambda/p$
	CRRA = $\frac{1}{2}$	CRRA = 1	CRRA = 3
1	1.090484	1.015513	1.003979
2	1.090463	1.015510	1.003978
3	4.925040	1.228484	1.022488
4	4.925298	1.327460	1.035749
5	4.915285	1.228249	1.022465
6	4.914000	1.228228	1.022514
7	4.893141	1.429708	1.042994
8	4.889595	1.538788	1.050375
9	4.841015	1.717541	1.064719
10	4.832996	1.807230	1.071853
11	4.759725	2.105987	1.099790
12	4.276567	2.860394	1.285217
13	4.008142	2.955983	1.396368

## 6. SUMMARY AND CONCLUSIONS

In this paper we have developed simulations of the market for 10-year term life insurance targeted at women aged 35 to 39 under various information and regulatory scenarios concerning genetic testing based on the BRCA1/2 genes. We first generate benchmark results (Module 1) based on a model of rate-making in which genetic test results are not available to either insured or insurers but insurers have access to and use relevant family background to establish risk groups. In particular, we compute the compensating variation ( $CV$ ) measure of the opportunity to insure in this environment for all women in the model. We then introduce information, available to both insureds and insurers, about genetic test results for the BRCA1/2 genes for all individuals (Module 2). Under the assumption that insurers are allowed to use this information to risk-rate

insurance policies, this results in separate premiums for those who test positive and those who test negative within each risk group, as determined by family background. The per capita  $CV$  is computed for this scenario as well and turns out to be slightly lower than that for Module 1. This result reflects the fact that in Module 2 individuals face premium risk in comparison to the single premium charged in Module 1.

In Module 3 we simulate the insurance market when 100% of consumers have genetic tests for the BRCA1/2 genes and that insurers are banned by regulation (or mutual agreement) not to use this information to risk-rate premiums. The resulting impact of adverse selection is captured in a number of ways. For each family background risk group we consider (1) the effect on demand for life insurance according to whether individuals test positive or negative for one of the genes, (2) the overall effect on the price of insurance in relation to the population weighted actuarially fair price of insurance, and (3) the effect on the per capita  $CV$  value of the insurance market opportunity. We find that for each of these criteria, the impact of adverse selection varies substantially across family background risk groups. This is because, roughly speaking, women with a stronger history of family members having had breast or ovarian cancer are more likely to test positive for one of the BRCA1/2 genes and so this enhances the impact of adverse selection on the market. Our analysis demonstrates the importance of considering separately the various risk groups as defined by family background when estimating or simulating the impact of a regulation banning insurers from using genetic test results.

In Module 2, where symmetric information prevails, 70% of women buy life insurance, and the average amount is very similar across risk types (from approximately 395,000 to 400,000).<sup>20</sup> In Module 3 we saw that the size of the effects of adverse selection can be quite substantial. Adverse selection leads to price increases ranging (across risk groups) from 1.5% to almost a tripling in the price for insurance. The result is that those women who discover they are lower risk types as a result of testing negative for the BRCA1/2 genes end up participating less in the life insurance market (from 67%

---

<sup>20</sup> The participation rate and average amount of insurance purchased is very similar in Modules 1 and 2, suggesting that the effect of genetic testing when information is symmetric will not change insurance buying behaviour very much, although there are significant price differences according to genetic test results.

in Groups 1 and 2 to only 21% in Groups 12 and 13). For those who do purchase life insurance, the average amount purchased drops slightly to 394,300 (Group 1) to a substantial drop to \$94,400 (Group 13). For those who test positive, all individuals purchase life insurance, except in groups 12 and 13, where 98% and 95% do. The average amounts purchased are much higher when individuals know that they are high risk, ranging from \$622,500 to over \$3 million.

The loss of efficiency due to adverse selection, as measured by the reduction in per capita  $CV$  in moving from modules 2 to 3, tends to rise as  $q$  rises (the proportion of individuals in the population that will test positive) and also as the ratio  $p^{pos}/p^{neg}$  rises (i.e., the difference in the probability of death between those who test positive and those who test negative). This is because the price distortion caused by higher risks buying more than lower risks rises in these circumstances. The range in the reduction of per capita  $CV$  caused by adverse selection is from less than 2% (for risk groups 1 and 2) to over 40% (for higher risk groups).

In Module 3 it is assumed that 100% of women know their genotype in terms of the BRCA1/2 genes. This is most likely too large a fraction even for the near-term future as the cost of this test is very high. Therefore, in Module 4 we assume two cases of only a fraction of women obtaining this information, 5% and 20%. In these cases the impact of adverse selection is indeed substantially smaller. We also considered (in Module 5) the effects of a regulation that allows those who obtain genetic tests and test negative to provide their results to the insurance company in order to obtain a “discount” while those who test positive are allowed to keep this information private and so get pooled with uninformed individuals. The result of this scenario was an overall efficiency effect that, although not very large, did imply less overall efficiency than if insurers were not allowed to use any test results, whether positive or negative, for risk-rating purposes. Of course, the option in such a regulation did improve welfare of those who test negative.

We do not think that our study should on its own be considered sufficiently comprehensive as to provide an answer to the question of whether life insurers should have access to insurance buyer’s genetic test results. There are many aspects of genetic privacy besides the efficiency effects resulting from adverse selection. However, considering the efficiency effect of adverse selection should be a component of a general

policy discussion concerning genetic privacy. Our simulation exercises suggest that at least for some family background types, if a sufficiently large fraction of women were to become informed about whether they have one of the BRCA1/2 genes, the efficiency effects may be quite substantial. If one were to consider the possibility of insurance buyers being informed about a wide range of genetic test results which have significant impact on mortality, then this concern would be reinforced. On the other hand, if the fraction of informed individuals is quite small our results suggest that the efficiency effects may well be quite low. Thus, the establishment of reasonably short moratoria (e.g., 5 years) on the use of genetic test results by insurers, with a review of the situation at the end of said moratoria before renewal is considered, seems to us a reasonable response to this controversy. In this way if the possibility of significant adverse selection costs from genetic privacy should arise over time, then such a stance could be reconsidered and alternative means of dealing with the equity issues revolving around the so-called genetic underclass could be considered.

### *Acknowledgements*

This research was supported by the CIHR (Canadian Institutes of Health Research) in conjunction with Genome Canada, Grant number FFO 62452 (Program - Facing Our Future: Genetics, Ethics, Law and Society). Opinions expressed are those of the authors and are not meant to necessarily reflect those of the CIHR or Genome Canada. The first author also thanks SSHRC for funding and the CESifo Institute for their generous financial support and hospitality while he worked on this project.

## REFERENCES

Abel, A. (1986), "Capital Accumulation and Uncertain Lifetimes with Adverse Selection," *Econometrica*, vol. 54, pp. 10797-97.

Arias, Elizabeth, Centres for Disease Control (2000), "National Vital Statistics Reports: United States Life Tables 2000". Volume 51, No. 3

Bernheim B.D., L. Forni, J. Gokhale, and L.J. Kotlikoff, (2003) "The Mismatch Between Life Insurance Holdings and Financial Vulnerabilities: Evidence from the Health and Retirement Study," *American Economic Review*, vol. 93, pp.354-365.

Blake, D. (1996), "Efficiency, Risk Aversion and Portfolio Insurance: An Analysis of Financial Asset Portfolios Held by Investors in the United Kingdom," *The Economic Journal*, vol. 106, pp. 1175-1192.

Bossert, W. and M. Fleurbaey (2002), "Equitable Insurance Premium Schemes," *Social Choice and Welfare*, vol. 19, pp. 113-125.

Brockett, Patrick L. and E. Susan Tankersley (1997), "The Genetics Revolution, Economics, Ethics and Insurance," *Journal of Business Ethics*, vol. 16, pp. 1661-1676.

Brugiavini, A. (1993), "Uncertainty Resolution and the Timing of Annuity Purchases," *Journal of Public Economics*, vol. 50, pp. 31-62.

Cawley, J. and T. Philipson (1999), "An Empirical Examination of Information Barriers to Trade in Insurance," *American Economic Review*, vol. 89, no. 4, pp. 827-846.

Crocker, K. J., and Snow, A. (1986), "The Efficiency Effects of Categorical Discrimination in the Insurance Industry," *The Journal of Political Economy*, vol. 94, pp. 321-44.

Ford D, Easton DF, Stratton M, Narod S, Goldgar D, Devilee P, Bishop DT, Weber B, Lenoir G, Chang-Claude J, Sobol H, Teare MD, Struewing J, Arason A, Scherneck S, Peto J, Rebbeck TR, Tonin P, Neuhausen S, Barkardottir R, Eyfjord J, Lynch H, Ponder BA, Gayther SA, Zelada-Hedman M, et al. "Genetic heterogeneity and penetrance analysis of the BRCA1 and BRCA2 genes in breast cancer families. The Breast Cancer Linkage Consortium." *American Journal of Human Genetics*, 1998; Mar 62(3):676-89.

Gokhale, J. and L.J. Kotlikoff (2002), "The Adequacy of Life Insurance," unpublished mimeo.

Hoy, M. (1984), "The Impact of Imperfectly Categorizing Risk on Income Inequality and Social Welfare," *Canadian Journal of Economics*, vol. 17, no. 3, pp. 557-568.

- Hoy, M. and Polborn, M., (2000), "The Value of Genetic Information in the Life Insurance Market," *Journal of Public Economics*, vol. 78, pp. 235-252.
- Hoy, M., and Lambert, P. J., (2000), "Genetic Screening and Price Discrimination in Insurance Markets," *The Geneva Papers on Risk and Insurance Theory*, vol. 25, pp. 103-130.
- Kitcher, P., (1996), *The Lives to Come: The Genetic Revolution and Human Possibilities*, New York: Simon and Schuster.
- Knoppers, B., B. Godard, and Y. Joly (2004), "Life Insurance and Genetics: A Comparative, International Overview," in *Genetics and Life Insurance: Medical Underwriting and Social Policy*, ed. Rothstein, M., MIT Press.
- Lemmens, Trudo and P. Bahamin (1998), "Genetics in Life, Disability and Additional Health Insurance in Canada: A Comparative Legal and Ethical Analysis," in : B. Knoppers (ed.), *Socio-Ethical Issues in Human Genetics*, Cowansville: Yvon Blais.
- Lemmens, T. (2000), "Selective Justice, Genetic Discrimination, and Insurance: Should We Single Out Genes in Our Laws?," *McGill Law Journal*, vol. 45, no. 2, pp. 347-412.
- Macdonald, A. (2003), "Genetics and Insurance: What Have We Learned So Far?," *Scandinavian Actuarial Journal*, forthcoming.
- Macdonald, A. and D. Pritchard (2000), "A Mathematical Model of Alzheimer's Disease and The APOE Gene," *Astin Bulletin*, Vol. 30, no. 1;, pp. 69-110.
- Macdonald, A. and D. Pritchard (2001), "Genetics, Alzheimer's Disease, and Long-Term Care Insurance," *North American Actuarial Journal*, vol. 5, no. 2, pp. 54-78
- North American Actuarial Journal*, (1999), Feature Articles on Genetic Technology and Underwriting: The 1998 Bowles Symposium, vol. 3, no. 1.
- Parmigiani, G., Berry, D.A., Aguilar, O. (1998), "Determining Carrier Probabilities for Breast Cancer-Susceptibility Genes BRCA1 and BRCA2", *American Journal of Human Genetics*, 62: pp. 145 – 158.
- Pauly, M. V., et al. (2003), "Price Elasticity of Demand for Term Life Insurance and Adverse Selection," NBER working paper #9925.
- Picone, G., Sloan, F., Taylor, Jr. D. (2004), "Effects of Risk and Time Preference and Expected Longevity on Demand for Medical Tests," *The Journal of Risk and Uncertainty* vol. 28, no. 1, pp. 39 - 53.

Ries LAG, Eisner MP, Kosary CL, Hankey BF, Miller BA, Clegg L, Mariotto A, Feuer EJ, Edwards BK (eds.), (2001), *SEER Cancer Statistics Review, 1975-2001*, National Cancer Institute.

Rothschild, M. and Stiglitz, M. (1976), "Equilibrium in Competitive Insurance Markets: An Essay on the Economics of Imperfect Information," *Quarterly Journal of Economics*, vol. 90, no. 4, pp. 630-49.

Subramanian, et al., (1999), "Estimating Adverse Selection Costs from Genetic Testing for Breast and Ovarian Cancer: The Case of Life Insurance," *The Journal of Risk and Insurance*, vol. 66, no. 4, pp. 531-550.

Villeneuve, B. (1999), "Mandatory Insurance and the Intensity of Adverse Selection," CREST/INSEE working paper, unpublished mimeo.

Villeneuve, B. (2000), "Life Insurance," in *Handbook of Insurance*, ed., Georges Dionne, Kluwer Academic Publishers.

Wilson, C. (1977), "A Model of Insurance with Incomplete Information," *Journal of Economic Theory*, vol. 16, pp. 167-207.

Witt, J. (2004), "The Effect of Information in the Utilization of Preventive Health Care Strategies: An Application to Breast Cancer," unpublished mimeo.

**Technical Appendix to  
Welfare Effects of Banning Genetic Information in the Life Insurance Market:  
The Case of BRCA1/2 Genes**

**Appendix A: Data Set and Construction of Variables**

THE DATA SET

We use the National Population Health Survey (NPHS) 1999 Health Component Data Set, with  $n=703$  females aged 35-39 to construct our variables. Included in the data set is detailed health information, including family background of breast and ovarian cancer. As described below some data is provided in group format and we constructed individual data as best as possible. For example, we chose as the actual individual income and household income the mean of the grouped income category to which each individual observation was assigned. For income in the highest groups (\$80,000 and up for household income, \$60,000 and up for individual income), we referred to the distribution in the NPHS Master File, where actual income is declared, and set those categories to their medians (\$100,000 for household income, \$75,000 for individual income). (NOTE: Unfortunately we could not access much of the family background data from the master file and so the public use version of the data had to be used.)

CALCULATION OF 10 YEAR CUMULATIVE DEATH PROBABILITIES

Since the object of the exercise is to simulate the market for 10 year term life insurance for women aged 37 at time of purchase, we compute (estimate) 10-year cumulative death probabilities from all sources including death from breast cancer.

To calculate  $p$  (the probability of death) the NPHS family history information concerning (i) whether the woman's birth mother ever had breast and/or ovarian cancer; (ii), if so, was it the cause of her death; and (iii) whether the woman had a sister who had breast and/or ovarian cancer. To calculate  $p$  according to these possible family histories we used the *CancerGene* program, and within that, used the "BRCAPRO"<sup>21</sup> results to determine 10-year probabilities of getting breast cancer, ovarian cancer, and the probability of being BRCA1/2-positive conditional on family history. Accordingly we developed 13 risk groups, corresponding to the different family backgrounds of individuals in the data set. To calculate the risk of death, the excess mortality from being at risk for breast cancer was calculated and added to the 10-year average mortality rate from all causes other than breast cancer. The *CancerGene* program uses information

---

<sup>21</sup> BRCAPRO is a program that uses BAYES' rules of determination of the probability of a mutation and the probability of risk of breast cancer stratified by age, given family history. An estimate of the mutation frequencies in the normal and Ashkenazi Jewish populations provides the probability of the mutation in the proband prior to the ascertainment of family history. For a detailed discussion on the methodology see Parmigiani ,et. al. (1998).

about family members to calculate an individual's risk of developing breast and ovarian cancer, and the probability of being BRCA1/2 positive. It requires some input on the physiology of the breast cancer, which was not available from the NPHS data set. We assumed that any breast cancer in the family was unilateral (as opposed to bilateral); this will quite significantly underestimate the 10-year risk of developing breast cancer if the affected family member really had bilateral breast cancer (this could lead to an up to 4-fold increase in risk for these individuals). The data set did not specify the number of sisters that an individual has; hence we assumed one sister, if the individual had "sisters". The sister's age was assumed to be 39, the mother's age was assumed to be 66, if alive. Only mother and sister (if present) were entered into the program since no breast and/or ovarian cancer history could have been obtained for male family members (e.g. father's sister, father's mother, and so on). If breast cancer was present in the family, the program also required to know the presence of other cancers. Since we had no information about this, we assumed that there weren't any. To find the same probabilities for individuals that are BRCA1/2 positive (the "high risk groups"), all family members that were entered into the program (mother and possibly sister) were assumed to have been tested for the BRCA gene, and found to be positive (on 182 delAG). Finally, no one was assumed to be of Ashkenazi Jewish background (since this information was not obtainable from the NPHS public use file), and that would underestimate the frequency of BRCA1/2 positive individuals for this group.

There was also insufficient detail in the NPHS data set about ages of affected family members, if any, for precise use of the program. Thus, the following additional assumptions were made. If the mother had experienced breast cancer, it was assumed she did at age 58 (if after age 50) or at age 40 (if before age 50). Likewise, if the mother had ovarian cancer, it was assumed she did so at age 66 (if after age 50) or at age 36 (if before age 50). If the sister had breast cancer, it was assumed she had it at age 37; if the sister had ovarian cancer it was assumed she had it at age 35. This is the family history information that was entered into the *CancerGene* program with 13 sets of probabilities being computed. The description of these risk groups, defined by alternative family background is described in the following table:

Table A.1 Family History Information

Group #	Mother had BC after 50 (58)	Mother had BC before 50 (40)	Mother had OC after 50 (66)	Mother had OC before 50 (36)	Cause of Death BC?	Has a Sister ?	Sister BC (age 37)	Sister OC (age 35)
1					n	y		
2					n	n		
3				y	n	y		
4	y				n	y		
5				y	n	n		
6	y				n	n		
7			y		n	y		
8			y		n	n		
9		y			n	y		
10		y			n	n		
11					n	y	y	
12				y	n	y		y
13	y			y	n	y		

The following probabilities were obtained for each group, including the probability of a woman with her specific family background having one of the BRCA1/2 mutations.

Table A.2 Cancer Probabilities for Various Family Background Groups

Group #	Prob. Of BC (next 10 years)	Prob of BRCA1/2 Mutation present
1	0.012682	0.001000
2	0.012700	0.001000
3	0.013242	0.003000
4	0.013271	0.004000
5	0.013315	0.003000
6	0.013339	0.004000
7	0.013686	0.005000
8	0.013786	0.006000
9	0.014449	0.008000
10	0.014633	0.009000
11	0.015761	0.013000
12	0.024145	0.043000
13	0.029722	0.065000

To get the high risk probabilities (i.e., for those with one of the BRCA1-2 mutations), the exact same family history information was entered for each group, but this time, the *CancerGene* program was told that the person had a genetic test, and that it was positive. Also, it was assumed that the mother had a positive genetic test (since BRCA mutations are identified through existing mutations in a family member), and that the sister (if applicable) had also tested positive. From these probabilities ( $p^{pos}$ ), and the “gross” probabilities ( $p$ ; i.e., the probabilities of getting breast cancer unconditional of mutation status), and the  $q$ ’s (i.e.; the probability that an individual would test positive for a mutation), we can use the identity

$$p = p^{pos} \cdot q + p^{neg} \cdot (1 - q)$$

to determine the probability  $p^{neg}$  (i.e.; the probability of getting breast cancer over the next 10 years conditional on a person obtaining a genetic test and testing negative). The results of these computations are provided in the following table:

*Table A.3 Probabilities of Breast Cancer According to BRCA1/2 Status*

<b>Group #</b>	<b>Prob of BC given BRCA1/2 positive</b>	<b>Prob of BC given BRCA1/2 negative</b>
1	0.140765	0.012554
2	0.140766	0.012572
3	0.295257	0.012393
4	0.295265	0.012138
5	0.295259	0.012467
6	0.295267	0.012207
7	0.295264	0.012271
8	0.295266	0.012087
9	0.295262	0.012184
10	0.295264	0.012084
11	0.295259	0.012080
12	0.295262	0.011963
13	0.295282	0.011261

These were used to calculate the overall morbidity of the individuals in each group.

The probability of death from all causes over the next 10 years for a 37 year old female was calculated as follows. Using the CDC Life Tables (Arias, 2000) for Females, the probability of survival was calculated for each year period starting at age 37 until age 47. The life tables give the probability of dying ( $d_t$ ) between periods  $x$  and  $x + 1$ , hence the survival probability was calculated as follows:

$$s_t = 1 - d_t$$

where  $s_t$  is the one year survival probability from age  $t$  to age  $t + 1$ . The 10-year survival probability ( $S$ ) was then calculated as the product of the annual survival rates:

$$S = \prod_{t=37}^{47} s_t = 0.983311952$$

The 10-year probability of dying ( $D$ ) was then calculated as the complement of the 10-year survival rate, hence

$$D = 1 - S = 0.016688047 \approx 0.016688$$

This is the probability of death from all causes that was used in the calculations.

The 10 year survival rate with breast cancer for someone affected by invasive breast cancer before age 50 in 1991 is 75.7% (Ries et. al., 2004). That is, out of all individuals in the SEER Cancer Registry who were diagnosed with invasive breast cancer in 1991 and were under age 50 at that time, 75.7% were still alive 10 years later. We considered this to be the best survival probability to use since the SEER database is very reliable and since 10 year probabilities were difficult to find. The mortality rate is one minus the survival rate:

$$M = 1 - 0.757 = 0.243$$

This is the 10 year probability of death from breast cancer given an invasive breast cancer diagnosis before age 50 that was used as the 10 year probability of death from breast cancer. The 10 year probability of getting breast cancer for the general population that we used is 0.01 for age group 40 – 49 (Ford et. al., 1998). Therefore, the probability of death from breast cancer in the general population is:

$$0.243 \cdot 0.01 = 0.00243$$

Subtracting this from the general population's probability of death:

$$0.016688 - 0.00243 = 0.014258$$

The mortality from breast cancer is then calculated for each group, e.g. for Group 1 it is:

$$0.243 \cdot 0.0127 = 0.0030861$$

and this is added to the general population's probability of death from all causes except breast cancer:

$$0.014258 + 0.0030861 = 0.0173441$$

These probabilities are referred to as  $p_j$  values (i.e., probability of death for an individual in family background group  $j$  given no knowledge about BRCA1/2 mutation status. This exercise was repeated for all base probabilities conditioned on individuals taking a genetic test with the probability being  $p_j^{neg}$  for a negative test result and  $p_j^{pos}$  for a positive genetic test result. Note that the “re-computed” probability of death even for the lowest risk class is slightly higher than that for the population as a whole (i.e., 0.0173441 for Group 1 compared to 0.016688). This occurs because *CancerGene* undoubtedly assigns different probabilities of cancer than would implicitly be the case of the overall probability from the Ford, et al. (1998) study. Moreover, this study is for a somewhat different age group and only provides a single significant digit. However, the absolute sizes of the probabilities are very close and it is the relative probabilities that matter to our analysis and we believe these are sound.

## ASSIGNMENT OF HOUSEHOLD SIZE

Household size was not directly given in the data set, so we calculated it as follows: a variable called “income adequacy” places each household into one of five categories (lowest income quintile to highest income quintile) as shown in the following table.

*Table A.4 Income Adequacy Definitions*

Income Adequacy	Description	Income	Household Size
1	Lowest Income	Less than \$10,000	1 to 4 persons
		Less than \$15,000	5 or more persons
2	Lower Middle Income	\$10,000 to \$14,999	1 or 2 persons
		\$10,000 to \$19,999	3 or 4 persons
		\$15,000 to \$29,999	5 or more persons
3	Middle Income	\$15,000 to \$29,999	1 or 2 persons
		\$20,000 to \$39,999	3 or 4 persons
		\$30,000 to \$59,999	5 or more persons
4	Upper Middle Income	\$30,000 to \$59,999	1 or 2 persons
		\$40,000 to \$79,999	3 or 4 persons
		\$60,000 to \$79,999	5 or more persons
5	Highest Income	\$60,000 or more	1 or 2 persons
		\$80,000 or more	3 persons or more
9	Unknown	Not stated	--

The variable “household arrangement” indicates the number of adults and whether any children are in the family. Thus, for example, if the household arrangement variable (HHARR) is coded a 3 (HHARR = 3) this implies a family made up only of a couple (2 adults). HHARR = 4 implies a couple with children but the number of children is not directly provided. However, if HHARR=4 and income adequacy is coded at 1 (lowest quintile) and family income is between 10,000 and 14,999 then their income quintile can only be the lowest (coded 1) if they have at least 3 children (5 family members), hence we assign 3 children to this family. But if this family had income below \$9,999 they may have one or more children. In this latter case we would assign the lowest possible number of children. We used the lower bound in all such cases. The following table (Household Size Assignment) summarizes the results of this exercise. The number in brackets following each imputed household size number is the number of observations allocated to a particular cell.

Table A.5 Household Size Assignments

Household Arrangement	Income Adequacy Group	Household Income (\$ ,000)											n
		0	< 5	5 - 10	10 - 15	15 - 20	20 - 30	30 - 40	40 - 50	50 - 60	60 - 80	> 80	
Couple	1	2 (0)	2 (0)	2 (1)	-	-	-	-	-	-	-	-	1
	2	-	-	-	2 (3)	-	-	-	-	-	-	-	3
	3	-	-	-	-	2 (2)	2 (4)	-	-	-	-	-	6
	4	-	-	-	-	-	-	2 (4)	2 (5)	2 (9)	-	-	18
	5	-	-	-	-	-	-	-	-	-	2 (21)	2 (21)	42
	9	-	-	-	-	-	-	-	-	-	-	-	0
Couple with Kid(s)	1	3 (2)	3 (1)	3 (3)	5 (2)	-	-	-	-	-	-	-	8
	2	-	-	-	3 (4)	4 (11)	5 (20)	-	-	-	-	-	35
	3	-	-	-	-	-	3 (24)	4 (58)	5 (33)	5 (19)	-	-	134
	4	-	-	-	-	-	-	-	3 (58)	3 (49)	4 (111)	-	218
	5	-	-	-	-	-	-	-	-	-	-	4 (98)	98
	9	3 (1)	3 (2)	3 (1)	-	-	3 (1)	-	-	-	-	5 (1)	-
Single Parent with Kid(s)	1	3 (0)	3 (2)	3 (8)	5 (1)	-	-	-	-	-	-	-	11
	2	-	-	-	2 (29)	3 (18)	5 (3)	-	-	-	-	-	50
	3	-	-	-	-	2 (7)	3 (25)	3 (10)	5 (1)	5 (0)	-	-	43
	4	-	-	-	-	-	-	2 (12)	3 (7)	3 (5)	4 (2)	-	26
	5	-	-	-	-	-	-	-	-	-	2 (0)	3 (3)	3
	9	-	-	-	2 (1)	-	-	-	-	-	-	-	1

## BENCHMARKING OF DEMAND PARAMETER ( $a$ )

The elementary utility function for the death state is  $v(x) = au(x)$  and so an increase in  $a$  leads to a higher marginal utility of income in the death state and, therefore, an increase in the demand for life insurance (see also demand function in equation (8)). We benchmark our value of this demand parameter ( $a$ ) on the basis that, if insurance is priced at the actuarially fair rate, then the optimal amount of insurance for an individual to purchase is that amount that would lead to maintaining per capita living standards of the family upon that person's death.<sup>22</sup> We adopt a standard household equivalence scale to guide us in our choice of  $a$ . Thus, if insurance is priced at the actuarially fair rate ( $\lambda=p$ ), then we want the value of " $a$ " to be such that per capita equivalent income in the life and death states is equal (i.e., standard of living of remaining household members is the same as it would be in the life state).

Let  $L^*$  be this value and  $L$  be the actual demand for life insurance as determined from utility maximization. So, let  $S_0$  be the number of person equivalents in the life state and  $S_D$  the number of person equivalents in the death state. These are not simply the number of household members in each state of the world due to economies of scale in consumption from living together. Using a standard (OECD) equivalence scale,<sup>23</sup> these numbers are determined according to the following rule:

$S_i$  takes on the value of 1 for the first adult, plus 0.7 for each additional adult, and then an additional 0.5 for each child. So, for example, for a couple with no children we have  $S_0 = 1.7$  and  $S_D = 1.0$ . For a couple with one child, we have  $S_0 = 2.2$  and  $S_D = 1.5$ . There is one exception to this rule. In the case of a one parent family,  $S_0 = 1.5$  and  $S_D = 0.8$  (i.e., loss of the adult reduces  $S_0$  by 0.7 rather than 1).

So, to equalize equivalent per capita income across states of the world requires an amount of life insurance,  $L^*$ , such that:

$$\frac{Y + K - \lambda L^*}{S_0} = \frac{K + (1 - \lambda)L^*}{S_D}$$

Once  $L^*$  is found, we then determine what must be the value of " $a$ " to give a level of insurance demand  $L(a, p, K, Y, \lambda) = L^*$  when insurance is priced at the actuarially fair rate (i.e.,  $\lambda=p$ ) using median values within each family type for the other parameters. The end result is given in the following table.

<sup>22</sup> For a discussion of this principle see Gokhale and Kotlikoff (2002) and Bernheim et al (2003).

<sup>23</sup> In particular, we use the formula  $S(N_1, N_2) = (1 + 0.7N_1 + 0.5N_2)$  to determine the equivalence weight (number of person-equivalents) for a family with  $N_1$  additional adults and  $N_2$  children. See Figini (1998) for a detailed discussion.

*Table A.6 Demand Parameter Values for Various Family Demographic Types*

Household Arrangement	Household Size	$\alpha$
Couple Alone	2	0.58787
Couple with Child(ren)	3	0.68160
	4	0.74031
	5	0.78074
Single Parent with Child(ren)	2	0.53031
	3	0.64704
	4	0.71730
	5	0.76423

### **Appendix B: Detailed Demand Analysis**

Except for the case of Module 3 only summary information was provided about the demand for life insurance of various groups. As noted in the main text of the paper, there was almost no variation in demand behaviour across risk groups and information sets (i.e., whether a person tested positive or negative for one of the BRCA1/2 genes). In this appendix we provide more complete information for these remaining cases (i.e., Modules 4 and 5).

Table A.7 Demand Effects of Adverse Selection in Module 4 ( $x=5\%$  testing)

Group	Test Group	Avg. Amt LD	Avg. Replacement Rate	% with LD > 0
1	Uninformed	392,489	0.450	69.7
2		392,490	0.446	69.7
3		382,370	0.437	69.7
4		398,614	0.455	66.6
5		382,383	0.437	69.7
6		382,331	0.437	69.7
7		396,987	0.453	66.6
8		389,853	0.444	66.6
9		386,685	0.440	66.6
10		385,155	0.438	66.6
11		382,318	0.434	65.1
12		332,462	0.373	64.3
13		310,934	0.346	63.1
1	Low (NEG)	390,645	0.448	69.7
2		390,646	0.448	69.7
3		388,687	0.443	66.6
4		388,107	0.442	65.7
5		388,712	0.443	66.6
6		388,722	0.443	66.6
7		386,102	0.439	65.1
8		375,319	0.426	65.1
9		366,541	0.415	64.9
10		361,297	0.408	64.9
11		337,033	0.379	64.7
12		297,899	0.329	47.7
13		261,484	0.291	42.3
1	High (POS)	2,043,744	4.193	100.0
2		2,043,329	4.192	100.0
3		4,233,870	9.229	100.0
4		4,222,767	9.203	100.0
5		4,224,982	9.208	100.0
6		4,223,422	9.205	100.0
7		4,184,091	9.115	100.0
8		4,141,108	9.017	100.0
9		4,081,371	8.880	100.0
10		4,061,384	8.834	100.0
11		3,931,870	8.538	100.0
12		3,214,910	6.898	100.0
13		2,849,313	6.061	100.0

Table A.8 Demand Effects of Adverse Selection in Module 4 (x=20% testing)

Group	Test Group	Avg. Amt LD	Avg. Replacement Rate	% with LD > 0
1	Uninformed	390,712	0.448	69.7
2		390,714	0.448	69.7
3		381,634	0.433	65.1
4		371,201	0.420	64.9
5		381,674	0.433	65.1
6		381,634	0.433	65.1
7		359,768	0.406	64.9
8		348,429	0.393	64.9
9		330,103	0.370	64.3
10		325,762	0.364	63.1
11		307,244	0.340	57.7
12		259,593	0.291	36.9
13		247,589	0.282	32.6
1	Low (NEG)	388,871	0.445	69.7
2		388,872	0.445	69.7
3		371,898	0.421	64.9
4		356,237	0.402	64.9
5		371,948	0.421	64.9
6		371,968	0.421	64.9
7		342,136	0.385	64.7
8		329,615	0.369	64.3
9		307,982	0.343	63.1
10		302,989	0.336	61.4
11		313,654	0.347	50.1
12		223,296	0.258	31.4
13		184,927	0.223	27.6
1	High (POS)	2,038,836	4.201	100.0
2		2,038,423	4.199	100.0
3		4,095,921	8.913	100.0
4		4,032,560	8.768	100.0
5		4,087,404	8.894	100.0
6		4,085,949	8.890	100.0
7		3,945,712	8.570	100.0
8		3,882,038	8.424	100.0
9		3,734,426	8.086	100.0
10		3,671,944	7.944	100.0
11		3,378,371	7.272	100.0
12		2,094,609	4.334	100.0
13		1,687,448	3.401	100.0

Table A.9 Module 5: Average LD and Replacement Rate

Group	Risk Group	Avg. Amt LD	Avg. Replacement Rate	% with LD > 0
1	Uninformed	390,122	0.447	69.7
2		390,123	0.447	69.7
3		371,202	0.420	65.1
4		355,966	0.402	64.9
5		371,251	0.420	65.1
6		371,215	0.420	65.1
7		344,973	0.388	64.9
8		332,250	0.373	64.3
9		309,148	0.344	63.1
10		303,725	0.337	61.4
11		317,271	0.351	50.1
12		247,745	0.282	33.3
13		225,200	0.260	31.4
1	Low	390,419	0.448	70.3
2		390,419	0.448	70.3
3		390,414	0.448	70.3
4		390,406	0.448	70.3
5		390,416	0.448	70.3
6		390,417	0.449	70.3
7		390,410	0.448	70.3
8		390,405	0.448	70.3
9		390,408	0.448	70.3
10		390,405	0.448	70.3
11		390,405	0.448	70.3
12		390,401	0.448	70.3
13		390,380	0.448	70.3
1	High	2,037,203	4.193	100.0
2		2,036,787	4.192	100.0
3		4,042,841	8.792	100.0
4		3,955,389	8.592	100.0
5		4,034,467	8.773	100.0
6		4,033,044	8.769	100.0
7		3,871,242	8.399	100.0
8		3,785,865	8.204	100.0
9		3,600,394	7.780	100.0
10		3,520,520	7.597	100.0
11		3,191,609	6.845	100.0
12		1,885,008	3.854	100.0
13		1,524,386	3.027	100.0

### **Appendix 3: Further Sensitivity Results**

In the base case described in the results of the main text of the paper, women considered as their insurable loss the net present value of all future (hypothetical) income that they would earn up to age 65. This amounted to almost 30 years of lost income. In this appendix we show that the qualitative results are not affected even if we reduce the time horizon of lost income to 10 years. Less insurance is purchased by all concerned, not surprisingly, and the impact on price of adverse selection does not change significantly (i.e., according to the  $\lambda/p$  ratio). The results for Modules 1 through 4 for this case are presented below.

*Table A.10 Module 1 Results for Revised Insurance Demand*

Group	Avg. Amt LD	Avg. Replacement Rate	% with LD > 0
1	142,272	0.453	65.7
2	142,273	0.453	65.7
3	142,279	0.453	65.7
4	142,279	0.453	65.7
5	142,280	0.453	65.7
6	142,280	0.453	65.7
7	142,284	0.453	65.7
8	142,285	0.453	65.7
9	142,292	0.453	65.7
10	142,294	0.453	65.7
11	142,306	0.453	65.7
12	142,396	0.454	65.7
13	142,456	0.454	65.7

*Table A.11 Module 2 Results for Revised Insurance Demand*

Group	Risk Type	Avg. Amt LD	Avg. Replacement Rate
1	Low	142,272	0.453
2		142,272	0.453
3		142,270	0.453
4		142,267	0.453
5		142,271	0.453
6		142,271	0.453
7		142,268	0.453
8		142,267	0.453
9		142,267	0.453
10		142,267	0.453
11		142,266	0.453
12		142,265	0.453
13		142,258	0.453
1	High	143,656	0.458
2		143,656	0.458
3		145,369	0.463
4		145,368	0.463
5		145,366	0.463
6		145,366	0.463
7		145,367	0.463
8		145,367	0.463
9		145,368	0.463
10		145,368	0.463
11		145,367	0.463
12		145,367	0.463
13		145,367	0.463

Table A.12 Module 3 Results for Revised Insurance Demand

Pooled Price Group	$p$	$p_{neg}$	$p_{pos}$	$\lambda$	$\lambda/p$
1	0.017340	0.017309	0.048413	0.017609	1.015513
2	0.017344	0.017313	0.048417	0.017613	1.015510
3	0.017476	0.017269	0.086038	0.022069	1.262817
4	0.017483	0.017208	0.086037	0.024208	1.384659
5	0.017494	0.017287	0.085975	0.022087	1.262547
6	0.017500	0.017294	0.085982	0.022094	1.262514
7	0.017584	0.017240	0.086009	0.026140	1.486578
8	0.017608	0.017195	0.086005	0.028095	1.595580
9	0.017769	0.017219	0.086018	0.031819	1.790702
10	0.017814	0.017194	0.086017	0.033694	1.891435
11	0.018088	0.017193	0.086000	0.039393	2.177859
12	0.020125	0.017165	0.086007	0.057765	2.870332
13	0.021480	0.016994	0.086010	0.063495	2.955983

Table A.12 Module 3 Results for Revised Insurance Demand

Group	Risk Group	Avg. Amt LD	Avg. Replacement Rate	% with LD > 0
1	Low	137,761	0.437	64.9
2		137,763	0.437	64.9
3		106,236	0.330	46.0
4		93,669	0.293	41.4
5		106,296	0.330	46.0
6		106,320	0.330	46.0
7		97,808	0.309	33.9
8		88,231	0.283	32.4
9		73,686	0.244	30.4
10		71,506	0.242	27.9
11		58,370	0.206	25.7
12		39,043	0.144	21.3
13		33,716	0.124	21.3
1	High	748,826	4.40	100.0
2		748,673	4.40	100.0
3		1,217,584	7.47	100.0
4		1,082,269	6.58	100.0
5		1,215,177	7.45	100.0
6		1,214,823	7.45	100.0
7		978,567	5.91	100.0
8		888,412	5.32	100.0
9		747,377	4.40	100.0
10		687,971	4.01	100.0
11		541,619	3.05	100.0
12		269,171	1.25	98.4
13		223,699	0.93	94.7

Table A.13 Module 4 ( $\alpha=5\%$  testing) Results for Revised Insurance Demand

Pooled Price Group	$p$	$p_{neg}$	$p_{pos}$	$\lambda$	$\lambda/p$
1	0.017340	0.017309	0.048413	0.017409	1.001096
2	0.017344	0.017313	0.048417	0.017413	1.001096
3	0.017476	0.017269	0.086038	0.017669	1.011044
4	0.017483	0.017208	0.086037	0.017808	1.018589
5	0.017494	0.017287	0.085975	0.017687	1.011032
6	0.017500	0.017294	0.085982	0.017694	1.011086
7	0.017584	0.017240	0.086009	0.017940	1.020246
8	0.017608	0.017195	0.086005	0.017995	1.021978
9	0.017769	0.017219	0.086018	0.018319	1.030953
10	0.017814	0.017194	0.086017	0.018394	1.032559
11	0.018088	0.017193	0.086000	0.018893	1.044505
12	0.020125	0.017165	0.086007	0.022465	1.116273
13	0.021480	0.016994	0.086010	0.024794	1.154282

Table A.14 Module 4 ( $\alpha=5\%$  testing) Results for Revised Insurance Demand

Group	Risk Group	Avg. Amt LD	Avg. Replacement Rate	% with LD > 0
1	Uninformed	141,865	0.452	65.7
2		141,866	0.451	65.7
3		139,437	0.443	65.1
4		137,315	0.436	64.9
5		139,442	0.443	65.1
6		139,423	0.443	65.1
7		136,723	0.434	64.9
8		136,103	0.432	64.9
9		132,927	0.421	64.9
10		132,365	0.419	64.9
11		128,233	0.404	64.9
12		112,027	0.348	61.3
13		124,941	0.387	50.3
1	Low	141,198	0.450	65.7
2		141,199	0.450	65.7
3		135,668	0.430	64.9
4		131,537	0.416	64.9
5		135,677	0.430	64.9
6		135,681	0.430	64.9
7		129,549	0.409	64.9
8		127,516	0.402	64.9
9		124,238	0.390	63.6
10		122,304	0.384	63.6
11		113,901	0.355	63.1
12		99,977	0.311	46.0
13		96,980	0.305	36.9
1	High	764,208	4.50	100.0
2		764,049	4.50	100.0
3		1,598,371	9.95	100.0
4		1,583,453	9.85	100.0
5		1,594,986	9.93	100.0
6		1,594,392	9.92	100.0
7		1,568,886	9.76	100.0
8		1,563,050	9.72	100.0
9		1,530,204	9.51	100.0
10		1,522,675	9.46	100.0
11		1,473,887	9.14	100.0
12		1,190,042	7.28	100.0
13		1,048,815	6.36	100.0

Table A.15 Module 4 ( $x=20\%$  testing) Results for Revised Insurance Demand

Pooled Price Group	$p$	$p_{neg}$	$p_{pos}$	$\lambda$	$\lambda/p$
1	0.017340	0.017309	0.048413	0.017409	1.002826
2	0.017344	0.017313	0.048417	0.017413	1.002825
3	0.017476	0.017269	0.086038	0.018269	1.045376
4	0.017483	0.017208	0.086037	0.018508	1.058628
5	0.017494	0.017287	0.085975	0.018287	1.045330
6	0.017500	0.017294	0.085982	0.018294	1.045371
7	0.017584	0.017240	0.086009	0.018840	1.071429
8	0.017608	0.017195	0.086005	0.019195	1.090129
9	0.017769	0.017219	0.086018	0.019919	1.120997
10	0.017814	0.017194	0.086017	0.020294	1.139216
11	0.018088	0.017193	0.086000	0.021693	1.199303
12	0.020125	0.017165	0.086007	0.031865	1.583353
13	0.021480	0.016994	0.086010	0.036994	1.722257

Table A.16 Module 4 (x=20% testing) Results for Revised Insurance Demand

Group	Risk Group	Avg. Amt LD	Avg. Replacement Rate	% with LD > 0
1	Average	141,223	0.449	65.7
2		141,224	0.449	65.7
3		127,916	0.404	64.9
4		125,983	0.396	63.6
5		127,932	0.404	64.9
6		127,918	0.404	64.9
7		121,780	0.382	63.6
8		116,631	0.364	63.1
9		110,587	0.343	61.3
10		125,627	0.389	51.6
11		114,167	0.353	50.1
12		93,269	0.298	32.6
13		81,956	0.267	31.4
1	Low	140,557	0.447	65.7
2		140,558	0.447	65.7
3		126,217	0.397	63.6
4		121,261	0.380	63.1
5		126,238	0.397	63.6
6		126,246	0.397	63.6
7		115,787	0.361	63.1
8		111,989	0.348	61.3
9		124,244	0.384	50.3
10		118,738	0.367	50.1
11		107,671	0.334	46.9
12		72,999	0.242	30.4
13		63,361	0.220	26.0
1	High	762,339	4.49	100.0
2		762,180	4.49	100.0
3		1,535,687	9.54	100.0
4		1,511,824	9.40	100.0
5		1,532,472	9.52	100.0
6		1,531,922	9.52	100.0
7		1,479,101	9.17	100.0
8		1,445,910	8.96	100.0
9		1,382,302	8.54	100.0
10		1,350,981	8.34	100.0
11		1,243,398	7.64	100.0
12		745,695	4.39	100.0
13		597,834	3.42	100.0

### References (for Appendix)

Arias, Elizabeth, Centres for Disease Control, “National Vital Statistics Reports: United States Life Tables 2000”. Volume 51, No. 3, December 19<sup>th</sup>, 2000.

Figini, P. (1998), “Inequality Measures, Equivalence Scales and Adjustment for Household Size and Composition,” LIS Working Paper No. 185.

Ford D, Easton DF, Stratton M, Narod S, Goldgar D, Devilee P, Bishop DT, Weber B, Lenoir G, Chang-Claude J, Sobol H, Teare MD, Struewing J, Arason A, Scherneck S, Peto J, Rebbeck TR, Tonin P, Neuhausen S, Barkardottir R, Eyfjord J, Lynch H, Ponder BA, Gayther SA, Zelada-Hedman M, et al. “Genetic heterogeneity and penetrance analysis of the BRCA1 and BRCA2 genes in breast cancer families. The Breast Cancer Linkage Consortium.” *American Journal of Human Genetics*, 1998; Mar 62(3):676-89.

Parmigiani, G., Berry, D.A., Aguilar, O., “Determining Carrier Probabilities for Breast Cancer-Susceptibility Genes BRCA1 and BRCA2”, *American Journal of Human Genetics*, 62: 145 – 158, 1998.

Ries LAG, Eisner MP, Kosary CL, Hankey BF, Miller BA, Clegg L, Mariotto A, Feuer EJ, Edwards BK (eds). *SEER Cancer Statistics Review, 1975-2001*, National Cancer Institute. Bethesda, MD, [http://seer.cancer.gov/csr/1975\\_2001/](http://seer.cancer.gov/csr/1975_2001/), 2004.