

# Exiting the Health Insurance Market as a Rational Choice: Demand for Health Insurance in a Learning Model\*

Rahul Jain<sup>12</sup>

---

\*Ph.D. candidate in the Economics Department at State University of New York at Buffalo, NY

<sup>1</sup>I thank Professor Isaac Ehrlich, Professor Jose M. Plehn-Dujowich, Professor Yong Yin for invaluable comments. I am solely responsible for all the remaining errors.

<sup>2</sup>407 Fronczak Hall, Department of Economics, State University of New York at Buffalo, Amherst, NY 14226. Telephone: (716) 645 2121, Fax Number: (716) 645 2127, Email: rhswj@buffalo.edu

## Abstract

It is widely believed that individuals exit the insurance market due to adverse shock to their income, insurance premium or both. This paper studies the role of *imperfect information* and subsequent *learning* about health endowment on individuals' decisions to continue health insurance coverage. We show that rational individuals may exit the insurance market *even in the absence* of adverse shocks to income and/or insurance premium. We assume that unknown health endowment and age are the only determinants of losses due to expenditures on medical care.

Individuals receive noisy signals about their endowment by observing these losses as they advance in age. Following Jovanovic (1982), we develop a model in which individuals incorporate this new information in the decision making process by using Bayesian Learning to update their beliefs about their health. Favorable new information diminishes the valuation of continuing insurance coverage; similarly, unfavorable new information increases its valuation. As individuals grow older, they accumulate additional information, which increases the precision of their beliefs. Beliefs that are more precise react imperceptibly to new information. Therefore, new information influences the beliefs of young people more than of old. Moreover, for a given belief, increased precision induces a mean preserving decrease in risk, reducing the demand for health insurance. Since precision increases as individuals age, we call it the *learning effect* of age. The other effect of growing older is the biological depreciation of health, which increases the size of loss. This increases the demand for health insurance and we call it the *aging effect* of age. Consequently, there is a continuous trade off between the gain in precision and increase in loss as individuals grow older. Bayesian Learning implies that the learning effect weakens with age. However, biological depreciation strengthens with age. Hence, we expect that the aging effect will override the learning effect at a unique point over the life span of an individual. Thus, unlike previous studies, we predict that controlling for changes in income, premium and new information, middle-aged individuals are least likely to renew insurance coverage.

In the second part of the paper, we use a panel component of MEPS data for 1997-2000 to test our model. We select individuals who are single, non-elderly, ineligible for public insurance and privately insured in the reference period and estimate their probability of insuring in the subsequent period. We construct a measure for new information and find that the likelihood of continuing coverage increases for adverse surprises. This relationship between new information and continuing coverage is stronger for young people than for old. We find that for the young, the probability of continuing coverage increases by 5.33 % points (95%CI 0.39, 13.60 % points) and for the old by 0.37 %points (CI -1.62, 1.20 % points), when new information is changed from 5<sup>th</sup> percentile to 95<sup>th</sup> percentile. We also find that the impact of age on the likelihood of continuing insurance coverage is *non-monotonic*. It decreases until 37 years of age (95%CI 34, 39 years of age), and then increases, as predicted by the theoretical model. Thus, we find some empirical evidence that unanticipated new information affects the demand for insurance and the interaction between learning and biological depreciation as individuals age.

# 1 Introduction

Health Insurance is of an overriding concern in the discussion of the health economy. Not having insurance promotes postponement in seeking care, non compliance of the treatment regime and results in an overall poor health outcome (Hadely 2002).

Mostly the individuals do not pay for the expenditures on health care directly. To a certain extent, a private insurance company or other program indirectly pays for much of the care, with the consumers paying only a part of it<sup>3</sup>, if at all. The Health Insurance coverage is provided through the payment of premiums in privately financed systems or by taxes in the publicly financed systems. Even if the coverage is purchased through a privately financed system, the consumer end up paying only a part of the premium if the insurance is purchased through the consumer's or its family member's participation in the labor force. This greatly reduces the out-of-pocket premium for the individual. However with the decline in the economy, many employers, who were earlier subsidizing the premium, either reduce their share of contribution or terminate the offer of insurance. This increases the cost of purchasing insurance for formerly insured consumers. These individuals may now choose to exit the insurance market all together (see footnote 4). This dynamic relationship is reflected in the health insurance coverage patterns seen in the past - better economy is characterized with high coverage rates<sup>4</sup>.

The standard economic theory predicts that a risk-averse agent would pay to avoid severe financial consequences of the unfortunate state of the world. This willingness-to-pay to avoid risk leads to the existence of insurance markets. In health insurance context, the unfortunate state of the world is usually an event of illness, which would require an individual or family to bear the whole cost of medical care out of income or wealth. Risk averse consumers, facing actuarially fair prices will fully insure. However, in real world, there are loading costs and individuals prefer to buy only an incomplete insurance contract. Given loading costs, a more risk averse individual

---

<sup>3</sup>Through co-payment, co-insurance and/or deductibles

<sup>4</sup>The federal government estimated that 43.6 million Americans lacked health insurance in 2002. The new Census Bureau figures show that 15.2 percent of Americans lacked coverage for all of 2002, an increase of 2.4 million people from 2001, when 14.6 percent were uninsured. The 5.8 percent rise in the uninsured resulted from a decline in the percentage of people covered by employer-based insurance - 61.3 percent in 2002, down from 62.6 percent the year before. That deterioration reflected increases in unemployment and the rise in health care costs, which saw some employers withdraw health insurance coverage. At the same time, the Census Bureau reported that the percentage of people covered by government programs (primarily Medicaid, the state-federal program for the poor and disabled, and the State Children's Health Insurance Program) rose to 25.7 percent in 2002 from 25.3 percent in 2001 (see Schaefer and Mcginley, 2003).

would buy higher degree of insurance coverage. The degree of risk aversion can be reasonably be assumed to depend on wealth, income, education, family status, access to other private<sup>5</sup> or public health insurance, health status, gender<sup>6</sup>, perceived risk and expected loss and the spread of this loss.

In view of this, most of the researchers have focused on the relationship between insurance premium, wealth and income, and access to free health care for the uninsured with the willingness-to-pay. Researchers have mostly ignored the relationship between perceived risk and expected loss, and the willingness-to-pay, with the exception of Bundorf, Herring, and Pauly (2005). They defined ‘health risk’ as the expected expenditure on medical care next period based on the health outcome in the current period. They find that those who expect future losses to be large are more likely to purchase insurance, a finding consistent with Ehrlich and Becker (1972). However, to our knowledge, no one has examined the mechanism by which the individuals’ perceptions about their expected loss are formed, nor has the question of how these perceptions change and what effect, if any, would these changes have on the demand for health insurance been addressed.

## 1.1 General Trends in the Number of Uninsured

In 1998, approximately 18 percent of the population was without insurance coverage and a bit more than 16 percent were uninsured in 2000 and 2001 (Fronstin 2002). Children and elderly are less likely to be uninsured due to availability of comprehensive public program like State Children’s Health Insurance Program and Medicare. The uninsured population is disproportional distributed in the population sub groups comprised by young adults, working poor and Hispanic populations. Young adults have an insurance rates of about 30 percent (Glied and Stabile 2001) while Hispanics have an insurance rate of about 32 percent (Mills 2002).

Overall un insurance rate for non-elderly, adult population was about 18% at the time of 1998 census, with Hispanics and low income groups bearing a disproportionate share of the uninsured. Garrett and Hudman 2002 find that women who leave welfare have 40 % un insurance rate. Cunningham, Schaefer and Hogan 1999 find that cost of insurance was the commonly cited barrier to enrollment in employer sponsored insurance coverage.

The situation after 1998 changed with the percentage of uninsured rising steadily. This increase uninsurance rate was attributed to health care cost rising faster than

---

<sup>5</sup>Through family member’s participation in labor force

<sup>6</sup>Since men and women have different health use profiles

the personal income and decline in the employment based health coverage due to weakening economy. Of all the uninsured, about 20% of the population are those who decline employer sponsored coverage.

## 1.2 Related Literature

Most of the work on health insurance demand is empirical and in the static framework. The research focuses on identifying causes for coverage or lack of it, the characteristics of the uninsured, and consequences of being without coverage.

There is some empirical work, which studies the problem in the dynamic setting. Swartz (1994), Swartz et. al. (1993a and 1993b) have found that employment status is an important determinant for continuing insurance coverage. Other determinants for continuing coverage were income, industry of employment, education, demographics, and region. Recent work has focused on the dynamics of insurance and un insurance on particular subgroups of population. Czajka (1999 and 2000) examine the dynamics of children. They found that ‘trigger events’ such as changes in parents employment status (employed to unemployed as well as full time to part time) are the determinants of transition from insured to uninsured status for children. Jensen (1992) and Sloan and Conover (1998) study this dynamics for near elderly. They had similar findings that change in employment status and income were main reasons for discontinuing insurance coverage.

On the theoretical side, there are relatively few research papers analyzing the insurance demand in the dynamic environment. Notable exceptions are Tesfatsion (1981) and Venezia & Levy (1983). In these models, however, the loss distribution is known and the losses are stochastically independent over time. Losses due to expenditure on medical care are, however, stochastically dependent over time. Those who have realized large expenditures in the past are likely to face large losses in future as well.

Thus, the general belief seems to be that individuals exit the insurance market whenever they face an adverse shock either to their income or to insurance premium or both.

We show that *imperfect information* and subsequent *learning* about health endowment may lead individuals to exit the insurance market *even in the absence* of adverse shocks to their income and insurance premiums. The uncertainty about the endowment affects individuals’ decisions through its impact on their beliefs and the precision of those beliefs.

In our model, insurance demand is modeled as a recursive decision problem and the learning-process provides the link between periods. In this model, therefore, losses are *stochastically dependent* over time.

Individuals receive *noisy signals* about their endowment by observing their realized expenses on medical care. We assume that these expenses depend only on underlying health endowment and age. By observing the loss, individuals update their beliefs by incorporating this new information. Thus, individuals accumulate information as they advance in age and the Bayesian learning process provides a structure to incorporate new information.

If the new information reveals low health endowment, individuals update their beliefs accordingly and expect the future loss to be large. This increases the value of insuring. However, if the new information discloses good endowment, beliefs adjust accordingly and individuals therefore have a low valuation of buying insurance. Thus, new information affects beliefs about the endowment, which in turn affects the value of insuring.

As individuals continue to accumulate additional information (i.e. advance in age), the *precision of the beliefs increases*. This increased precision has two effects. First, if two individuals reach the *same* beliefs, one who uses more information has more precise beliefs, i.e., for a given belief, more information induces a *mean preserving compression* on the posterior distribution. This suggests that the demand for insurance is reduced for all risk averse individuals when the loss distribution undergoes a decrease in risk. Second, beliefs that are more precise react imperceptibly to new information, i.e., the effect of new information on the beliefs decreases as individuals accumulate more information. Since individuals are learning *passively* and accumulate information as they advance in age, new information will affect the beliefs of younger age groups *more than* those of the older generation. This implies that the learning effect weakens as individuals age.

The other effect of the aging process is the *depreciation of health*. This increases the size of the loss as individuals grow older. Hence, the subjective value of being insured increases as individuals advance in age. The incentive to buy insurance is even stronger if the insurance premium does not adjust to include the higher expected loss.

Therefore, we show that new information affects the demand for insurance, and aging has two competing effects on the probability of continuing coverage. *Unanticipated bad news* increases the probability of staying insured. Similarly, *unanticipated good news* increases the chance of exiting the insurance market. Further, the effect of the new information is stronger at a younger age than at an older age. Aging has

two competing effects on the decision to continue the coverage. For a given belief, increased precision increases the possibility for exiting the market and increased size of the loss due to biological depreciation of health increases the possibility for continuing coverage. Thus, unlike previous studies, we predict that controlling for changes in income, premium and new information, middle-aged individuals are least likely to renew insurance coverage.

This paper is organized as follows. Section 2 details the learning process, elaborates on the analytical framework used to analyze the demand for insurance with learning and solves the model under different assumptions and draws out the behavioral implications on the demand for health insurance. All the proofs of the propositions are relegated to Appendix. Section 3 elaborates on the testable hypotheses as implied by the economic model. We discuss the data source, measures for new information and changes in income and insurance premium in Section 4 and the empirical results in Section 5. We conclude in Section 6.

## 2 Economic Model

Individuals are risk averse and they maximize their expected utility. Every period each individual receives income,  $w$ , and faces an uncertain financial loss due to expenditure on medical care. Let  $M(x_n, n)$  represent the loss due to medical care, which depends on random realization of ill-health index,  $x_n$ , and age  $n$ , independently. Ill-health index,  $x_n$ , is a positive, strictly increasing, continuous, and differentiable transformation of the ill health parameter. Since we model health as ill-health, medical expenditure is increasing in  $x_n$  and age. Individuals *do not know* their endowed ill health. They receive noisy signals about their ill health in the form of realized expenditures on medical care and can infer the signal about their endowment by observing the size of their losses and their age. They incorporate this new information in the decision making process by updating their beliefs about their ill health and their insurance purchase decisions are based on the *updated distribution* of future losses.

We are interested in isolating the dynamic implications of information accumulation between periods; hence we assume that a) individuals cannot postpone these losses to future periods and b) there is no *ex-post moral hazard* i.e. medical care expenditure is independent of the insurance status. Individuals have a choice to purchase medical care insurance from the market, at offered premium, *before they realize their losses*.

We do not model the insurer's problem and assume that individuals take the

premium as given. We assume that individuals consume all their income and spend money on consumption,  $c$ , and either on premium or medical care expenses or both.

## 2.1 Learning Process

### 2.1.1 Ill Health Parameter

Let  $\theta$  denote endowed the ill-health parameter which is normally distributed over the population, with *known* mean,  $\bar{\theta}$ , and *known* variance  $\sigma_\theta^2$ . Assume that individuals differ by their endowed ill health parameter and each individual knows that he is a random draw from this distribution, but does not know his *true* ill health parameter. Let  $\eta_n$  be the observed signal at age  $n$ .  $\eta_n$  is the sum of the *unknown* true ill-health parameter,  $\theta$ , and an *idiosyncratic random shock* or noise,  $\epsilon$ , which is assumed to be identically and independently distributed normal, with mean zero and variance  $\sigma_\epsilon^2$ .  $\epsilon$ 's are independent across individuals and over time. Specifically:

$$\eta_n = \theta + \epsilon_n \quad \text{with} \quad \begin{cases} \epsilon_n \sim i.i.d. N(0, \sigma_\epsilon^2) \\ Cov(\theta, \epsilon_n) = 0 \forall \theta, n \\ Cov(\epsilon_m, \epsilon_n) = 0 \forall m \neq n \end{cases} \quad (1)$$

$\eta_n$  is the sum of two normally distributed random variables; therefore its distribution is also normal with *unknown* mean,  $\theta$ , and *known* variance.

As age progresses, individuals gather information about their ill health parameter and integrate this new information by updating their prior beliefs about  $\theta$ . Following Jovanovic (1982), we assume that individuals use the Bayesian updating rule. At birth, all individuals have the same prior beliefs about their ill health parameter, which they set equal to the prior distribution that is, at birth, *prior mean* is  $\bar{\theta}$ , and *prior variance* is  $\sigma_\theta^2$ . Since  $\theta$  and  $\epsilon$  have a normal distribution, the *posterior distribution* of  $\theta$  is also normal, conditional on the past observations of  $\eta$ . Let  $\eta^n \equiv \{\eta_1, \dots, \eta_n\}$  be the information available at age  $n$ , then the posterior distribution of  $\theta$  has posterior mean and variance as below (see DeGroot (1970); also see Appendix A).

$$E[\theta|\eta^n] = \lambda(n)\bar{\eta}_n + (1 - \lambda(n))\bar{\theta} \quad (2)$$

$$Var[\theta|\eta^n] = \frac{\sigma_\theta^2 \sigma_\epsilon^2}{n\sigma_\theta^2 + \sigma_\epsilon^2} = \sigma_n^2 \quad (3)$$

where  $\lambda(n) = \sigma_\theta^2 / \left( \sigma_\theta^2 + \frac{\sigma_\epsilon^2}{n} \right)$ ,  $\bar{\eta}_n = \frac{1}{n} \sum_1^n \eta_i$  is the *sample* mean and  $\bar{\theta}$  is the *prior* mean. As is evident from (2) and (3) and the fact that normality is assumed,  $\bar{\eta}_n$  and  $n$  are *sufficient statistics* for posterior distribution of  $\theta$ .

The random shock,  $\epsilon$ 's, precludes the individual from ever learning his true  $\theta$  and ensures that learning is gradual. He does not know whether the large realization of his ill health parameter is due to the large value of the true ill health parameter or because of the large realization of the shock. His accumulation of more observations on his ill health parameter leads him to better estimate his true  $\theta$ . A large value of  $\sigma_\epsilon^2$  implies that the shock is extremely dispersed and therefore signals are less informative. Thus, the individual learns about his  $\theta$  slowly. By assumption,  $\sigma_\epsilon^2$  is the same for all individuals and, as a result, the speed of learning is the same for all the individuals<sup>7</sup>.

Reorganizing equations (2) and (3), posterior mean and variance may be written in terms of *sample precision* and *prior precision*<sup>8</sup>.

$$\begin{aligned} E[\theta | \eta^n] &= \frac{h_\epsilon}{h_\theta + h_\epsilon} \bar{\eta}_n + \frac{h_\theta}{h_\theta + h_\epsilon} \bar{\theta} \\ \text{Var}[\theta | \eta^n] &= \frac{1}{h_\theta + h_\epsilon} = (h_\theta + h_\epsilon)^{-1} \end{aligned}$$

This shows that the posterior mean, (2), is a *weighted average* of the sample mean,  $\bar{\eta}_n$ , and prior mean,  $\bar{\theta}$ , with weights that are proportional to sample precision,  $h_\epsilon$ , and prior precision,  $h_\theta$ . Additionally, posterior variance, (3), is reciprocal of the sum of sample precision and prior precision. Since the sample precision increases with sample size, this implies that a) the weight on the sample mean is increasing and b) posterior precision is increasing in  $n$ . In other words, as  $n$  increases, the posterior distribution of  $\theta$  becomes concentrated around the sample mean. Therefore, as sample size increases, sample mean becomes a more efficient predictor of the true ill health parameter. The rate of gain in efficiency is, however, decreasing in the sample size.

---

<sup>7</sup>Within a cohort, individuals may have different rates of learning if we assume either that there is heterogeneity in prior variance, i.e., different  $\sigma_\theta^2$ , or there is heterogeneity in the information contained in signals i.e. different  $\sigma_\epsilon^2$ . The first assumption implies that individuals are born with information about their health that is more precise, and therefore new information is less valuable. Hence, their prior beliefs react imperceptibly to new information at all ages relative to individuals who are born with less precise priors (see Blomberg and Harrington (2000)). The second assumption implies that the signals contain more information for some individuals and less for others, i.e. the idiosyncratic errors are heterogeneously dispersed, therefore individuals learn at different rates. We do not consider a situation where individuals pay to learn about their ill health. In other words, there is *no active learning* in this model.

<sup>8</sup>*Precision* of a random variable is the reciprocal of the variance. Here prior precision is  $h_\theta = (\sigma_\theta^2)^{-1} = 1/\sigma_\theta^2$  and sample precision is  $h_\epsilon = (\sigma_\epsilon^2/n)^{-1} = n/\sigma_\epsilon^2$  which is increasing in  $n$ .

Twice differentiating (3) with respect to  $n$ ,

$$\begin{aligned}\frac{\partial \sigma_n^2}{\partial n} &= -\frac{\sigma_\theta^2 \sigma_\epsilon^2}{(n\sigma_\theta^2 + \sigma_\epsilon^2)^2} \sigma_\theta^2 < 0 \\ \frac{\partial^2 \sigma_n^2}{\partial n^2} &= \frac{\sigma_\theta^2 \sigma_\epsilon^2}{(n\sigma_\theta^2 + \sigma_\epsilon^2)^3} 2\sigma_\theta^4 > 0\end{aligned}$$

we see  $\partial \sigma_n^2 / \partial n < 0$  and  $\partial^2 \sigma_n^2 / \partial n^2 > 0$  that is posterior variance decreases at a decreasing rate as  $n$  increases.

Writing (2) and (3) recursively, we get:

$$E[\theta|\eta^n] = \frac{\sigma_\epsilon^2}{\sigma_{n-1}^2 + \sigma_\epsilon^2} E[\theta|\eta^{n-1}] + \frac{\sigma_{n-1}^2}{\sigma_{n-1}^2 + \sigma_\epsilon^2} \eta_n \quad (4)$$

$$Var[\theta|\eta^n] = \frac{\sigma_\epsilon^2 \sigma_{n-1}^2}{\sigma_{n-1}^2 + \sigma_\epsilon^2} = \sigma_n^2 \quad (5)$$

where  $E[\theta|\eta^n]$  is the *updated belief*  $E[\theta|\eta^{n-1}]$  is the *prior belief*, and  $\eta_n$  is the *new information*. Collecting the terms together, we get:

$$E[\theta|\eta^n] - E[\theta|\eta^{n-1}] = (\eta_n - E[\theta|\eta^{n-1}]) \frac{\sigma_n^2}{\sigma_\epsilon^2} \quad (6)$$

where left hand side of (6) is the *size of revision* of beliefs and on the right hand side of (6), the difference between new information and prior belief is the *unanticipated new information* in the signal, or surprise, weighted with *posterior variance*,  $\sigma_n^2$ .

It is immediate from (6) that the change in belief is proportional to the deviation of new realization from prior belief. If  $\eta_n > E[\theta|\eta^{n-1}]$ , then belief adjusts to a larger value; otherwise it amends to a lower value. Unanticipated *bad* news induces the individual to believe that his  $\theta$  is large and it is likely that he will observe large realizations in the future as well. Since  $\sigma_n^2$  is decreasing in age,  $n$ , beliefs react strongly to unanticipated *new* information at a younger age than in older age. This implies that the importance of surprises decreases as one accumulates more information with advancing age. It also shows that the learning effect weakens as an individual advances in age.

### 2.1.2 Ill Health Index

Expenditure on medical care depends on the individual's age,  $n$ , and *ill-health index*,  $x_n$ , which is a positive, strictly increasing, continuous, and differentiable transformation of the ill health parameter. Accordingly, for individuals with realized ill health parameter  $\eta_n$ , ill health index  $x_n$  is  $x_n = \xi(\eta_n)$  with,  $\lim_{\eta_n \rightarrow -\infty} \xi(\eta_n) = 0$  and  $\lim_{\eta_n \rightarrow \infty} \xi(\eta_n) = \infty$ .

$\xi(\eta_n) = \text{constant} \leq \infty$ . Let  $\tilde{x}_{n+1}$  be the posterior expectation of  $x_{n+1}$ , conditional on information available up to age  $n$ ,

$$E[x_{n+1}|\eta^n] = \tilde{x}_{n+1} = \int \xi(\eta_{n+1})dF_{n+1}(\eta_{n+1}) \quad (7)$$

where,  $F_{n+1}(\eta_{n+1})$  is the normal posterior distribution  $\theta$  after  $n$  realizations of  $\eta$ . Since  $\xi(\cdot)$  is strictly increasing in  $\eta$ ,  $\tilde{x}_n$  is strictly increasing in  $\bar{\eta}_n$  for each  $n$ , and the pair  $(\tilde{x}_n, n)$  are also *sufficient statistics* for the posterior distribution of the ill health index,  $x_n$ .

We assume that  $\xi(\cdot)$  is an exponential transformation of  $\eta$ . Therefore,  $x$  is log normal and individuals have log normal priors for their ill health index with posterior mean as  $E[x_{n+1}|\eta^n] = \tilde{x}_{n+1}$  and posterior variance as  $\text{Var}[x_{n+1}|\eta^n] = \tilde{x}_{n+1}^2 [\exp(\sigma_n^2) - 1]$  where,  $\tilde{x}_{n+1}$  is as in (7) and  $\sigma_n^2 = \frac{\sigma_\epsilon^2 \sigma_\theta^2}{n\sigma_\theta^2 + \sigma_\epsilon^2}$  (see (3)). It is easy to show that  $\tilde{x}_n$  is a strictly increasing and convex function of  $\bar{\eta}_n$  (Appendix A), hence  $\tilde{x}_n$ , and  $n$  are *sufficient statistics* for the posterior distribution of  $x_n$ . Also, since  $\sigma_n^2$  is a decreasing function of  $n$ , as  $n$  approaches infinity, the subjective distribution collapses to the *true* value of ill-health index.

Let  $\tilde{x}_n$  be the *updated belief*,  $\tilde{x}_{n-1}$  be the *prior belief*,  $x_n$  be the *new information*, and  $0 < \gamma(n) = \sigma_n^2/\sigma_\epsilon^2 < 1$ , with  $\partial\gamma(n)/\partial n < 0$ , then the *size of the revision of beliefs* is:

$$\frac{\tilde{x}_n - \tilde{x}_{n-1}}{\tilde{x}_{n-1}} = \left[ \left( \frac{x_n}{\tilde{x}_{n-1}} \right)^{\gamma(n)} - 1 \right] \quad (8)$$

It is clear from (8) that the size of revision of an individuals belief depends on the new information. If  $x_n > \tilde{x}_{n-1}$ , then  $\left[ \left( \frac{x_n}{\tilde{x}_{n-1}} \right)^{\gamma(n)} - 1 \right]$  is positive and individuals adjust their beliefs upwards, otherwise they revise their beliefs downwards. Since  $\partial\gamma(n)/\partial n$  is negative, the beliefs react strongly to unanticipated information at a younger age than at an advanced one.

## 2.2 Medical Care Expenditure

We model the financial losses due to expenditures on medical care as a function of an individual's realized ill health index and age, independently, where ill health represents the individual's specific endowment and age captures the biological depreciation. This approach gives us *heterogeneity across* cohorts, as well as *within* cohorts. Researchers have usually modeled the biological depreciation as exogenous and we follow this convention. We assume that age enters the loss function deterministically and is independent of ill health. Thus, older individuals have larger medical

expenditures and individuals with larger ill health realize larger expenditures. The above assumptions imply that  $M(x_n, n)$  is increasing in age ( $M_n > 0$ ) and increasing in the ill health index ( $M_{x_n} > 0$ )<sup>9</sup>. For simplicity, we assume the following functional form:

$$M(x_n, n) = m\delta^n x_n \quad (9)$$

where  $x_n$  is the *realized* ill health index at age  $n$ , and  $m$  and  $\delta$  are constants. We assume that individuals know these constants and  $\delta (> 1)$  represents the rate of biological depreciation. The assumption that age enters the loss function exponentially implies that the aging effect of age gets stronger as individuals age. Since individuals know the depreciation rate,  $\delta$ , and  $m$ , they can infer their realized ill health index by observing their realized expenditure.

## 2.3 Individual's Problem

We model the insurance demand under two assumptions. Initially, individuals face a binary choice of either full insurance or no insurance. We relax this assumption later and allow individuals to buy partial insurance. In the former case, individuals face predetermined insurance premium,  $p$ , which they take as given. If they choose to be insured, they pay the premium  $p$ , and the insurer bears all financial loss due to medical expenditure. If they choose to remain uninsured, they are responsible for all the expenditure in medical care. As mentioned earlier, these expenses are independent of insurance status and may not be postponed to future period(s). In the partial insurance case, if premium adjusts to changes in beliefs,  $\tilde{x}$ , we show that the demand for insurance has ambiguous sign with respect to changes in the beliefs about their ill-health.

### 2.3.1 Full Insurance Case

The representative individual maximizes his expected utility by choosing consumption,  $c$ , and insurance status,  $I_n$ . His budget constraint is

$$w = c + pI_n + (1 - I_n)M(x_n, n) \quad (10)$$

where  $w$  is income,  $c$  is consumption, and its price is normalized to one,  $M(x_n, n)$  is the random expenditures on medical care when realized ill health is  $x_n$  at age  $n$ ,  $p$  is the insurance premium, and  $I_n$  is the indicator function of insurance status. If the

---

<sup>9</sup>Where,  $M_{x_n} = \partial M(x_n, n)/\partial x_n$  and  $M_n = \partial M(x_n, n)/\partial n$

individual chooses to buy insurance,  $I_n = 1$ , else  $I_n = 0$ . Thus, individuals maximize expected utility, subject to (10).

$$\text{Max}_{\{c, I_n\}} E(u(c)|\eta^n) \quad \text{subject to} \quad w = c + pI_n + (1 - I_n)M(x_n, n)$$

where  $u(c)$ , is a concave utility function, i.e.,  $u'(\cdot) > 0$  and  $u''(\cdot) < 0$ .

The individual has a pair of *sufficient statistics*,  $(\tilde{x}_n, n)$ , which completely characterizes his posterior distribution about future realization of ill health,  $x_n$ . Consequently, the maximization problem reduces to  $\text{Max}_{\{c, I_n\}} E(u(c)|\tilde{x}_n, n)$ , subject to (10). Substituting this budget constraint in the objective function, an individual chooses  $I_n$  according to  $\text{Max} \{E(u_0|\tilde{x}_n, n), E(u_1|\tilde{x}_n, n)\}$ , where  $E(u_0|\tilde{x}_n, n) = \int_0^\infty u(w - M(x_n, n)) f(x_n|\tilde{x}_n, n) dx_n$  and  $E(u_1|\tilde{x}_n, n) = u(w - p)$ . Therefore, the individual buys insurance when he expects that the utility of being insured is greater than being uninsured. Formally,

$$I_n^* = \begin{cases} 0 & \text{if } u(w - p) < E[u_0|\tilde{x}_n, n] \\ 1 & \text{if } u(w - p) \geq E[u_0|\tilde{x}_n, n] \end{cases} \quad (11)$$

Individuals who believe that their ill health index is large also expect their future loss to be large and have lower expected utility of being uninsured. Conversely, individuals who have gained evidence indicating lower endowed ill health expect their future losses to be lower and as a result, have higher expected utility of being uninsured than those whose experience has suggested otherwise. Therefore, for a given  $n$ , the expected utility of being uninsured decreases in  $\tilde{x}_n$ . This result is stated in Proposition (1) and the proof is given in Appendix (B.1)

**Proposition 1**  $E[u_0|\tilde{x}_n, n]$  is monotonically decreasing in  $\tilde{x}_n$ .

Utility with full insurance is independent of  $\tilde{x}$  and  $n$ , since the insurer bears all the financial loss. As a result, for a given  $n$ , there is a unique  $\tilde{x}$ , say  $\tilde{x}^*$ , such that  $u(w - p) = E[u_0|\tilde{x}^*, n]$  and therefore  $\forall \tilde{x} < \tilde{x}^*$ ,  $u(w - p) < E[u_0|\tilde{x}, n]$  and  $\forall \tilde{x} \geq \tilde{x}^*$ ,  $u(w - p) \geq E[u_0|\tilde{x}, n]$ . Hence, for a given  $n$ , there is a unique *reservation belief*,  $\tilde{x}^*$ , such that individuals buy insurance if their belief about their ill health is greater than  $\tilde{x}^*$  i.e.

$$I_n^* = \begin{cases} 0 & \text{if } \tilde{x}_n < \tilde{x}_n^* \\ 1 & \text{if } \tilde{x}_n \geq \tilde{x}_n^* \end{cases} \quad (12)$$

As individuals grow older, their beliefs change stochastically according to new information. Every period, individuals compare their updated beliefs against the

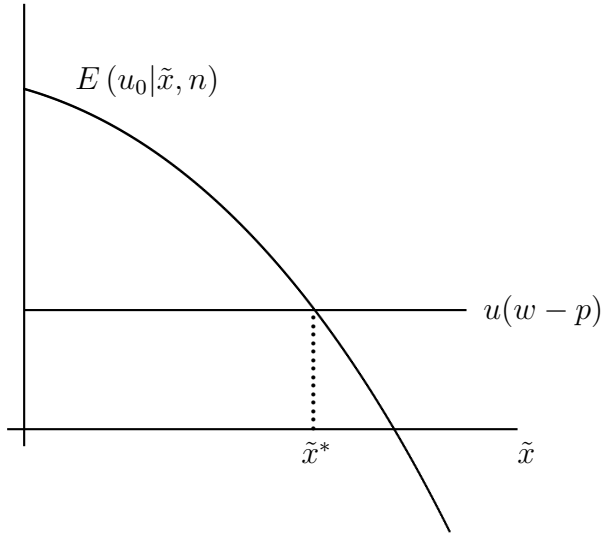


Figure 1: Reservation Belief for a Given  $n$ .

reservation beliefs and make decisions according to equation (12). If the new information suggests lower endowed ill health, individuals adjust their beliefs accordingly. These individuals will therefore, have a larger likelihood of exiting the insurance market. However, since the beliefs of the young adjust more than those of the old, the effect of the new favorable information would be larger for the young than for the old.

Since the expected utility of being uninsured is age dependent<sup>10</sup>, the reservation belief also depends on age. If the expected utility of being uninsured increases, the reservation belief,  $\tilde{x}^*$ , increases. Similarly, if the expected utility of being uninsured decreases, the reservation belief,  $\tilde{x}^*$ , decreases.

We show that growing older has competing effects on the expected utility of being uninsured. As individuals advance in age, their uncertainty about their endowed ill health diminishes, which increases the expected utility of being uninsured. At the same time, the depreciation of health leads to increase in loss and hence decreases the expected utility. We call the former effect of age the *learning effect* of age and the latter, the *aging effect*. The learning effect weakens with age, while the aging effect gets stronger. Therefore, we show that the reservation belief first increases and then decreases.

We analyze each effect in turn. To examine in isolation, the effect of aging as implied by the learning process, we consider the special case in which the rate of biological depreciation is set to one i.e.  $\delta = 1$ . Substituting this in the expected

<sup>10</sup> $E(u_0|\tilde{x}_n, n) = \int_0^\infty u(w - M(x_n, n)) f(x_n|\tilde{x}_n, n) dx_n$

utility of being uninsured, we get:

$$E[u_0|\tilde{x}_n, n] \Big|_{\delta=1} = \int_0^\infty u(w - mx_n) f(x_n|\tilde{x}_n, n) dx_n$$

Aging is the source of information, which reduces uncertainty individuals have about their endowed ill health. Thus, the posterior distribution of ill health depends on the sample size; hence, the reservation beliefs also depend on sample size. As sample size increases, the learning process gradually dispels uncertainty about endowed ill-health index. Nevertheless, random shocks rule out individuals from ever learning their true endowment. There is always an element of uncertainty because individuals do not know whether high realization is due to high true ill health or due to large realization of shock. The increased precision about the ill-health index reduces the risk in decision making process. Though the expected utility without insurance is random i.e. depends on new information, for a given  $\tilde{x}$ , it is increasing in age. If there is no change in premium, the reservation belief,  $\tilde{x}_n^*$  increases. This result is stated in Proposition 2 and proved in Appendix B.2.

**Proposition 2** *For given  $\tilde{x}$ ,  $E(u_0|\tilde{x}, n) \Big|_{\delta=1}$  is monotonically increasing in  $n$ .*

This shows that under imperfect information, due to uncertainty, some individuals buy insurance that would choose to exit the insurance market as they gather more information about their endowed ill-health.

When we include biological depreciation, for a given belief, the future losses are increasing, implying that expected utility of being uninsured is decreasing. This suggests that the reservation ill-health is decreasing as individuals advance in age. However, aging is the source of information as well which reduces the uncertainty about endowed ill health, hence the learning effect due to aging increases the expected utility of being uninsured (see Proposition (2)). In the end, whether the reservation ill health increases or decreases depends on the relative strength of the two effects. As seen above, the learning effect weakens as individuals advance in age and the aging effect gets stronger as individuals grow older. Thus, if learning effect were ever to dominate the aging effect, it has to be at a younger age. If the learning effect is weak<sup>11</sup> or the rate of depreciation very strong, it is possible that aging effect dominates the learning effect for the whole lifetime of an individual.

---

<sup>11</sup>This may happen when the  $\sigma_\epsilon^2$  is small.

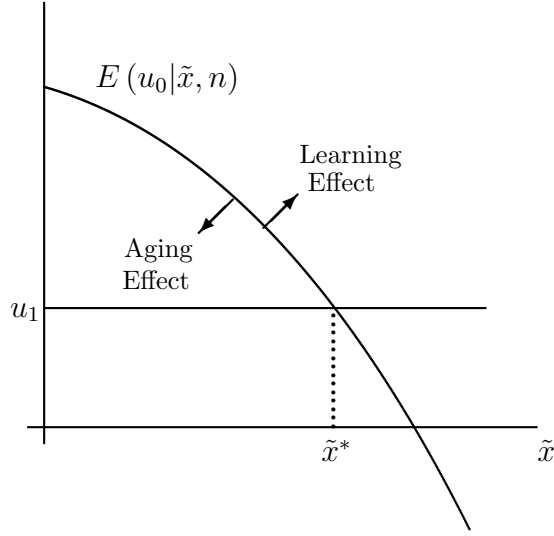


Figure 2: Competing Effects of Age on expected utility of being uninsured

Formally, as individuals advance in age, the aging effect shifts the expected utility of being uninsured to the left and the learning effect shifts it to the right (Figure (2)). Thus  $\tilde{x}_n^* - \tilde{x}_{n+1}^*$  depends on the relative strengths of the aging and learning effects. Since the learning effect is strongest at the young age and weakens as individuals grow older, we expect the reservation ill health to increase and then decrease as individuals grow older (Figure (3)). This result is formally stated in Proposition (3) below.

**Proposition 3** For given  $\tilde{x}$ ,  $E(u_0|\tilde{x}, n) \Big|_{\delta > 1}$  increases and then decreases as  $n$  increases.

**Proof** For given  $\tilde{x}$ ,

$$\Delta_n E[u_0] = E[u_0(w - \delta^{n+1}mx)|\tilde{x}, n+1] - E[u_0(w - \delta^n mx)|\tilde{x}, n]$$

Adding and subtracting  $E[u_0(w - \delta^{n+1}mx)|\tilde{x}, n]$ , we get

$$\begin{aligned} \Delta_n E[u_0] &= [E[u_0(w - \delta^{n+1}mx)|\tilde{x}, n+1] - E[u_0(w - \delta^{n+1}mx)|\tilde{x}, n]] \\ &\quad + [E[u_0(w - \delta^{n+1}mx)|\tilde{x}, n] - E[u_0(w - \delta^n mx)|\tilde{x}, n]] \\ &= \Delta_n^L E[u_0] + \Delta_n^D E[u_0] \end{aligned} \quad (13)$$

where,  $\Delta_n^L E[u_0]$  is the change in expected utility due to change in the posterior distribution, and  $\Delta_n^D E[u_0]$  is the change in expected utility due to increase in the size of loss. Thus,  $\Delta_n^L E[u_0]$  is the *pure learning effect* of age and  $\Delta_n^D E[u_0]$  is the *pure aging effect* of age.

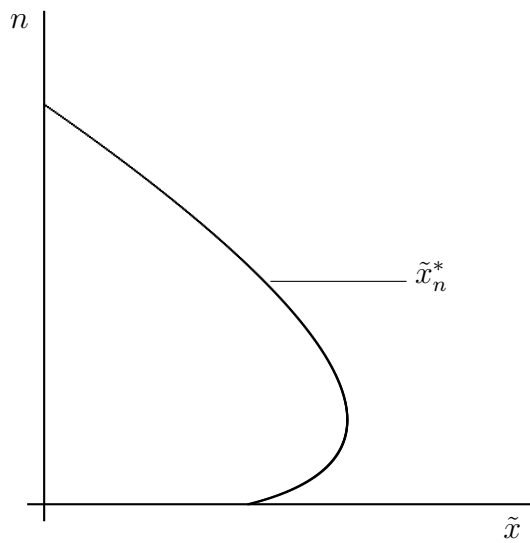


Figure 3: Reservation Ill Health - The Competing Effects of Age.

From proposition 2,  $\Delta_n^L E[u_0]$  is positive and from concavity of the utility function,  $\Delta_n^D E[u_0]$  is negative. The sign of the (13) depends on the relative size of  $\Delta_n^L E[u_0]$  and  $\Delta_n^D E[u_0]$ . We know that the learning effect wanes as individuals grow older and aging effect reinforces as individuals advance in age, thus learning effect dominates, if at all, at younger age and depreciation effect dominates later in life.

### 2.3.2 Partial Insurance Case

Thus far individuals could either buy full insurance or remain uninsured. Now we relax this assumption and show that similar results are obtained when individuals are allowed to purchase partial insurance.

We are interested in the change in demand for insurance coverage due to changes in *beliefs* about endowed health,  $\tilde{x}$  and age,  $n$ . As discussed in the previous chapter, both enter the posterior distribution of ill-health. The intuition of the results obtained below remains the same; however, the proofs are more involved and are relegated to the appendix. The difficulty of the proofs stems from the fact that under the partial insurance case, the first order stochastic dominant and second order stochastic dominant shifts are not sufficient for intuitive results. We elaborate on this in the next section.

Individuals maximize their expected utility by choosing their consumption,  $c$ , and partial insurance (coinsurance contract),  $\alpha \in [0, 1]$ , which pays an indemnity of  $\alpha M(x_n, n)$ . An individual who buys  $\alpha$  amount of insurance faces a premium  $P = \alpha \lambda m \delta^n p$ , where  $\lambda (> 1)$  is unity plus the loading factor, and  $m$  and  $\delta$  are as before.

We make alternative assumptions about  $p$ . If insurers have asymmetric information,  $p$  does not depend on an individual's belief. Hence  $P$  is age rated, but not experience rated. If on the other hand, insurers have symmetric information ( $p = \tilde{x}$ ), the premium reacts to changes in beliefs<sup>12</sup> and  $P$  is both age rated and experience rated. The budget constraint of the individual is

$$w = c + (1 - \alpha)M(x_n, n) + \alpha\lambda m\delta^n p \quad (14)$$

where  $w$  is income,  $c$  is consumption whose price is normalized to one,  $M(x_n, n)$  is the (stochastic) expenditures on medical care when realized ill health is  $x_n$  at age  $n$ ,  $\lambda m\delta^n p$  is the insurance premium where  $\lambda (> 1)$  is unity plus the loading factor,  $p$  is a constant in the case of asymmetric information and is equal to  $\tilde{x}_n$  in the case of symmetric information and  $\alpha$  is the amount of insurance bought. For the asymmetric information case, we assume that individuals may at most buy full coverage (optimal  $\alpha \in [0,1]$ ); we rule out the possibility of gambling. In the symmetric information case, the value of  $\lambda$  is such that the optimal  $\alpha \in (0, 1)$  (interior solution) for all  $\tilde{x}$  and  $n$ .

Given this representation, the individual's problem is

$$\text{Max}_\alpha E[u|\tilde{x}, n] = \int_0^\infty \left[ u(w - m\delta^n[(1 - \alpha)x + \alpha\lambda p]) \right] f(x|\tilde{x}_n, n) dx$$

and the optimal  $\alpha$ , say  $\alpha^*$ , solves

$$\int_0^\infty \left[ u'(w - m\delta^n[(1 - \alpha^*)x + \alpha^*\lambda p]) \right] \left[ m\delta^n(x - \lambda p) \right] f(x|\tilde{x}_n, n) dx = 0 \quad (15)$$

As individuals age, that is,  $n$  increases, they receive new information which they incorporate in the decision making process by adjusting their beliefs about their endowed ill health according to (8). Increased age also increases the size of loss due to the depreciation of health, and increases the precision individuals have about their beliefs.

### Shifts in the posterior distribution due to changes in $\tilde{x}$ and $n$

In the previous chapter we have shown that, for a given  $n$ , an increase in  $\tilde{x}$  induces a first order stochastic dominant shift in the posterior distribution. Similarly,

---

<sup>12</sup>The cost of monitoring individuals' medical care expenditures deters insurance firms from experience rating them. In reality, individuals who buy insurance from non-group settings face age dependent insurance premiums and in some extreme case, experience rating. However, the premium for individuals who buy insurance from group settings, is dependent on the characteristics of the group, rather than of specific individuals'. Therefore the premium faced by individuals is somewhere between these two extreme assumptions of exogenously determined premium and individual specific insurance premium (see Bundorf et al (2005)).

for a given  $\tilde{x}$ , an increase in  $n$  induces a mean preserving compression in the posterior distribution. However, neither the first order stochastic dominant shift nor the mean preserving compression of the posterior distribution necessarily imply intuitively consistent changes in the demand of insurance (see Rothchild and Stiglitz (1971), Hadar and Seo (1990) and Gollier (2001), pp 60). The foundation of the problem lies in the fact that  $\left[ u'(w - m\delta^n[(1 - \alpha^*)x + \alpha^*\lambda p]) \right] \left[ m\delta^n(x - \lambda p) \right] = g(x)$ , is not monotonic and concave in  $x$ .

Hadar and Seo (1990) have shown that assuming  $g(x)$  to be monotonic (for FOSD shift) is equivalent to assuming that the relative risk aversion is less than one and assuming  $g(x)$  is concave (for MPC shift) is equivalent to assuming that relative risk aversion is less than unity and increasing, and absolute risk aversion is decreasing (Hadar and Seo (1990), pp 724 and 726). These assumptions are contrary to empirical evidence at hand. For example, Kessler and Wolff (1991) have shown that households with low wealth excessively choose to put their wealth in risk free assets<sup>13</sup>. This suggests that relative risk aversion decreases with wealth. Other researchers have found additional restrictions on the utility function for example, see Dreze and Modigliani (1972), Diamond and Stiglitz (1974), Dionne and Eeckhoudt (1987), Dionne, Eeckhoudt and Briys (1989). Some researchers have imposed restrictions on the distribution function (Meyer and Ormiston (1983 and 1985), Black and Buckley (1989) Eeckhoudt and Hansen (1980, 1983)) to achieve an intuitively consistent result. Yet others have put restrictions on both the utility functions and the distribution functions (Sandmo (1970 and 1971), Meyer and Ormiston (1989)).

For our purposes, we need to find restrictions on the utility function, taking into account the restrictions implied by the learning process on the posterior distribution functions. As it turns out, we need a simple restriction on the utility function to get the desired results on changes in the demand for insurance due to shifts in the posterior distributions as a result of changes in  $\tilde{x}$  and  $n$ .

We assume that the utility function has decreasing absolute risk aversion (DARA). This assumption implies that individuals behave in a less risk averse fashion the larger their wealth; gambles accepted at a given level of wealth will be accepted at all higher levels of wealth (Dybvig and Lippman (1983), Yaari (1969)). Mathematically, decreasing absolute risk aversion means that the percentage decrease in marginal utility is itself decreasing<sup>14</sup>. This assumption is behavioral and can be directly interpreted in

---

<sup>13</sup>Over 80% of the lowest quintile of individuals by wealth, had portfolio wealth in liquid assets, while the highest quintile had less than 15% of wealth in liquid assets (Kessler and Wolff (1991)).

<sup>14</sup>In other words,  $-u'$  is a more concave than  $u$ .

terms of choice behavior<sup>15</sup>.

### Changes in demand for insurance due to changes in $\tilde{x}$

As individuals grow older, their beliefs adjust according to new information and aging affects the posterior distribution as well as the size of the loss.

Intuitively, if everything else stays the same, individuals who have large  $\tilde{x}$ , believe that their future losses would be large and demand more insurance. For the asymmetric information case, where premiums are independent of  $\tilde{x}$  (no experience rating), the demand increases unambiguously. However, in the symmetric information case, where premiums adjust according to the beliefs (i.e., with experience rating), larger  $\tilde{x}$  also implies a larger per unit price for insurance. This increase in the price of insurance, induces the usual substitution and income effect (Jang and Hadar (1995)) and the direction of change in the demand for insurance is ambiguous.

Let  $\alpha_a^*(\tilde{x}, n)$  be the optimum coinsurance under asymmetric information; then we show that  $\alpha_a^*(\tilde{x}, n)$  increases with  $\tilde{x}$ . This implies that if there is no change in premium, individuals with larger  $\tilde{x}$  would demand more insurance. Also, let  $\alpha_s^*(\tilde{x}, n)$  be the optimal insurance under the symmetric information case; then  $\alpha_s^*(\tilde{x}, n)$  has an ambiguous sign with respect to  $\tilde{x}$ . For the proof of this proposition, the additional assumption of decreasing absolute risk aversion utility function is required. The results are formally stated in Proposition 4 and the proofs are in Appendix (C).

**Proposition 4** *a) Under asymmetric information, where the insurance premium is independent of  $\tilde{x}$ ,  $\alpha_a^*(\tilde{x}, n)$  monotonically increases in  $\tilde{x}$ . b) Under symmetric information, where the insurance premium increases in  $\tilde{x}$ ,  $\alpha_s^*(\tilde{x}, n)$  has an ambiguous sign with respect to  $\tilde{x}$ .*

### Changes in demand for insurance due to changes in $n$

#### *Aging Effect of Age*

As before, age,  $n$ , enters the individual's objective function at two places. First, it enters the loss function directly (aging effect) and second, it enters the posterior distribution of the ill-health index (learning effect). We continue to study them one at a time. We start with the aging effect of growing older and keep the posterior distribution constant. Because of the depreciation of health, the size of the loss increases for all realizations of ill-health index  $x$ . This increased loss increases the demand for insurance. Let  $\alpha^*(\tilde{x}, n)$  be the optimum coinsurance contract at age  $n$ ;

---

<sup>15</sup>For example, decreasing absolute risk aversion implies that the risk premium is a decreasing function of wealth.

we show that for a given  $\tilde{x}$  and fixed posterior distribution,  $\alpha^*(\tilde{x}, n+1) > \alpha^*(\tilde{x}, n)$ . The result is stated in Proposition 5 and the proof is given in Appendix D.

**Proposition 5** *For a given  $\tilde{x}$  and fixed posterior distribution,  $\alpha^*(\tilde{x}, n)$  monotonically increases in  $n$ .*

#### *Learning Effect of Age*

As individuals grow older, they accumulate more information. This additional information, for a given  $\tilde{x}$ , induces a mean preserving compression in the posterior distribution of ill health index,  $x$ . The intuition is that if two individuals hold the *same beliefs*, the one who uses more information will have a mean preserving decrease in risk. This decrease in risk lessens the likelihood of extreme values, compared to the individual with less information and the same belief. Therefore, an older individual would expose a larger proportion of wealth to risk and hence would demand less insurance.

This decrease in demand for insurance is due to change in the posterior distribution induced by additional information about the unknown ill health index,  $x$ . Therefore, to get the pure learning effect of age, we keep the size of loss constant.

We show that the learning effect induces a relative strong decrease in risk (Black and Bulkley (1989) pp 120, Alarie, Dionne and Eeckhoudt (1991)). Therefore individuals decrease their demand for insurance. Proposition 6 formally states this result; the proof is in Appendix (E).

**Proposition 6** *For a given  $\tilde{x}$  and size of loss,  $\alpha^*(\tilde{x}, n)$  monotonically decreases in  $n$ .*

#### *Competing Effects of Age*

Combining the results of Propositions 5 and 6, we conclude that age has two competing effects on the demand for insurance. More information leads to a reduction in uncertainty and therefore a decline in the demand for insurance. At the same time, biological depreciation of health increases the size of a future loss, which increases the demand for insurance coverage. Since the learning effect wanes and the aging effect grows as individuals age, the learning effect will be dominant, if at all, for the young and the aging effect will be dominant for the old<sup>16</sup>.

---

<sup>16</sup>If the signals are less noisy ( $\sigma_\epsilon^2$  is small), the learning effect may be small. In that case, the aging effect is dominant throughout the individual's life.

### 3 Hypotheses

In the rest of the paper, we develop some testable hypotheses as implied by the economic model described above. Specifically, we examine the role of *new information* and *age* on the likelihood of an individual continuing insurance coverage.

#### 3.1 New Information

We have shown that individuals' demand for health insurance depends on their belief about their unknown endowed health. These beliefs are updated as individuals receive new information. If they receive adverse surprise, they adjust their beliefs accordingly and expect larger future loss due to higher medical care costs. Therefore, our model predicts, that they have lower likelihood of exiting the health insurance market. Conversely, according to our model, favorable surprise would increase the likelihood of exiting the insurance market. Proposition (1) and (4) are the basis of our first hypothesis.

**Hypothesis 1** *Adverse (favorable) surprise increases (decreases) the likelihood of continuing insurance coverage.*

However, as shown in equation (8), individuals beliefs react strongly to unanticipated new information at a younger age than at advanced age. As discussed above, this result is due to the fact that older individuals use more information than younger individuals to form their beliefs. Therefore, older individuals have more precise beliefs than younger individuals. This implies that older individuals beliefs react imperceptibly to new information. Therefore, we expect that the magnitude of the effect of new information on younger individuals decision to continuing coverage would be greater than that on the older individuals decision of staying insured. This leads us to our second hypothesis.

**Hypothesis 2** *The effect of new information is stronger for younger individuals than for older individuals.*

In US, most of the people obtain their coverage through their employer i.e. they obtain the insurance from group settings and the employee contribution tends to depend on the characteristics of the group rather than on the characteristics of individuals (Pauly and Herring 1999). Therefore, it is unlikely that new information that individuals receive is going to affect the insurance premium. Thus, we assume that insurance premiums may change due to factors beyond the control of individuals, for example, loss in the offer from the place of work will increase the insurance premium.

## 3.2 Age

As shown in Propositions (2) and (6), increased precision in beliefs would lead to decrease in demand for insurance by all risk averse individuals - *learning effect of age*. Similarly, due to biological depreciation of health, the size of future loss would increase as individuals advance in age, increasing the demand for insurance (Proposition (5)) - *aging effect of age*. Therefore, as long as the learning effect of age dominates over the aging effect, the demand for insurance will fall as individuals advance in age. When aging effect dominates over the learning effect, individuals would demand more insurance as they advance in age. This implies that the likelihood of remaining insured would decline when learning effect is dominating and it would increase when the aging effect is dominating. We have shown above that learning effect continuously declines as individuals accumulate more information. Therefore learning effect is strongest at younger age. This implies that learning effect would dominate, if at all, at younger age and aging effect would dominate as individuals approach death. This analysis generates our third hypothesis.

**Hypothesis 3** *Age has competing effects on the likelihood of remaining insured. We expect the likelihood of remaining insured would first decrease and then increase as individuals advance in age, because the learning effect would only dominate at young age.*

## 4 Methods

We are interested in testing the link between unanticipated new information and associated age effects and an individuals decision to continue health insurance coverage. We first study this relationship for all the individuals and then examine differences in the effects of new information and age for young and old individuals.

### 4.1 Data Source

We use the data provided Agency for Healthcare Research and Quality in Medical Expenditure Panel Survey (MEPS). The survey uses an overlapping panel design. New households are included each year and they are followed for two years. Each household is interviewed in five rounds over a thirty months period to collect information regarding their health care expenditures over two years. The Household Component is a nationally representative survey of United States civilian, non-institutionalized population and collects information about medical care expenditures, medical care use,

health care conditions, health insurance status for respondents along with information on demographic and socio-economic characteristics. This survey design provides us with the opportunity to observe the changes in insurance status, employment status and other factors that affect income, insurance premium and administrative load. This structure also allows us to create a proxy for new information that individuals receive in the form of realized expenditures on medical care, about which we elaborate below. We use data from three reference periods covering 1997-2000.

To test our model, we confine our study to single individuals, between the ages of 18 and 64, who are ineligible for any type of public health insurance and are privately insured in the base year (year 1). We study the insurance purchase decision of these individuals in the subsequent year (year 2).

Focusing only on individuals with family of size one restricts the empirical analysis, but it is necessary to eliminate any possible intra-family considerations that may enter the decision making process of a family. In addition, if the unit is the household, then it is not practical to estimate the age effects, as there is no implied proxy for the ‘precision’ of beliefs and ‘age’ of the household. Furthermore, our model in chapter two and three allows the analysis of the decision problem of currently uninsured individuals, however we do not model the demand for insurance by these individuals for the following reasons. Firstly, we have assumed that the medical expenditure can not be postponed. However uninsured may choose to postpone some of the expenditure to future periods, hoping to buy insurance in future and therefore subsidize the cost. Also, by assumption, individuals who are not insured are assumed to

## **4.2 Measure for unanticipated new information**

To obtain a measure of unanticipated new information that individuals receive about their health endowment, we need a proxy for the beliefs individuals held about their endowed health in the base period. Once we have this proxy, we may quantify new information by taking the deviation of actual realization of health from perceived endowed health.

We assume that regression predictions of an individuals medical care expenses in the base period based on individuals health status and their socioeconomic and demographic information is a good proxy for perceived losses due to medical expenses. The difference between realized expenses and predicted expenses is, therefore, new information about the loss. However, since we are interested in the perception about

the health and not the losses per se, we normalize these differences with the regression predictions of health care expenses conditional only on age and sex.

This ratio measures the extent to which the individuals actual expenditures deviate from expected losses, solely due to difference in the endowed health. Hence, this ratio captures the new information individuals receive about their endowed health.

We develop a standard two-stage estimation method. Jones (2000) has argued that if the aim is to predict the expected expenditures, then two-stage model is better than the sample selection model, or Heckit. In the first stage, we use the probit model to predict the likelihood of any positive annual medical expenditure. In the second stage, we regressed an ordinary least square regression of log transformation of positive expenditures. We rejected this approach since the residuals from the OLS regression were non-normal and heteroscedastic and therefore, the coefficients were biased (non-normal errors) and inefficient (heteroscedastic errors). Hence for the second stage, we chose the Generalized Linear Models. These models have the advantage over log-transformed models since they estimate  $\log(E(y|x))$ . To estimate these models, we have to identify the ‘link function’ and the ‘family of distribution’ of the error terms.

So, the following probit model is first estimated for all the individuals in our data.

$$p_j = \phi(X_j\beta_1) \tag{16}$$

where,  $p_j$  is a binary variable indicating whether individual  $j$  had non-zero medical expense and  $X$  includes  $j$ 's characteristics such as, gender, race, age, physical and cognitive limitations, health status, work conditions education level and wage.

For the second stage of this model, we estimate Generalized Linear Model for the expenses of those individuals who have positive expenditures. There are three constituents in the GLM. The first is the linear part:  $\eta_j = X_j\beta_2$ , where  $X$  includes  $j$ 's characteristics such as, gender, race, age, physical and cognitive limitations, health status, work conditions education level and wage and  $\beta_2$  is the column vector of unknown coefficients. The second constituent is a monotonic, differentiable link function, which relates the expected expenditures to the linear predictor.

$$g(\mu_j) = X_j\beta_2 \tag{17}$$

where,  $g(\cdot)$  is the link function and  $\mu_j = E(y_j|x)$ . We, as others have, assume that the link function is  $\log(\cdot)$ . The expenditures are i.i.d. drawn from an exponential distribution. This implies that the variance depends on the mean (see Blough et. al. 1999 for more details). We have to appropriately choose the family of distribution as

implied by the data. We performed specification tests for the link function and the family of distribution about which we elaborate below.

The link function specifies the relationship between the explanatory variables and the mean. It characterizes how the mean in the untransformed or raw scale is related to the predictors. We assume, as others have, that the link function is log. Then we use Pergibons Link Test to determine the feasibility of this assumption. After the regression, we predict  $\widehat{X\beta_2}$  and regress medical expenditures on  $\widehat{X\beta_2}$  and  $\widehat{X\beta_2}^2$  with log link. If the coefficient of  $\widehat{X\beta_2}^2$  is significant then the link function is rejected. The coefficient of  $\widehat{X\beta_2}^2$  was  $-0.0605$  with the p-value  $0.2281$ . This indicates that the log link is well specified.

The family of distribution specifies the mean variance relationship. Table (1) below identifies the distribution as reflected by their mean-variance relationships. We use Modified Park Test to identify the family of distribution. We regress log of the residuals (in raw scale) on log of predicted medical expenses and a constant term. Then test for the coefficient of the  $\log(\hat{y})$ . We found the estimate to be  $1.82026$  with t-value  $21.39$ . This indicates that Gamma specification is the best specification for the family of the distribution.

We used Modified Hosmer Lemshaw Test to determine if there was a systematic pattern of bias in the model fit on raw scale. For this test we estimate the residuals in the raw scale and regress them on ten groups dummies, sorted by  $\widehat{X\beta_2}$ , and without the constant term. We find that all the coefficients are individually insignificant (see Table 2) and the F-statistics is  $0.6864$  with p-value =  $0.7380$ .

Finally, with the predicted probability from part one and the predicted expense from part two, the expected expense for each individual  $j$  is defined as

$$\hat{y}_j = \hat{p}_j \times \hat{\mu}_j$$

where,  $\hat{p}_j$  is the prediction from equation (16), and  $\hat{\mu}_j$  is the prediction from equation (17) using  $\log(\cdot)$  link function and Gamma family of distribution.

Results from the first stage probit model are presented in Table (3). It uses the data of 1550 individuals in the MEPS data. In Table (4) we present the GLM regression results using 1358 individuals who had positive expense in the base year. We find in our regression results that white race, female gender, older age, higher education poor self reported health, higher wage, physical limitation are each significant predictor of the probability of positive medical expense. We also found that female gender, older age, physical limitation, poor self reported health, and high education are each significant determinant of expenditures.

Since the aim is to construct a measure of new information, we re-estimate the medical expense equation, only this time we use race, gender, and age as the explanatory variables. We again use the two part model and perform the specification and robustness tests (see Tables 5 and 6 for the estimates). We found that Pergibons Link test suggests log link function (coefficient of  $\widehat{X\beta_2}^2$  was  $-0.1419$  with p-value  $0.2496$ ). Using Modified Park test we find that gamma distribution is the correct specification for the family of distribution (coefficient of  $\log(\hat{y})$  was  $1.93117$  and p-value  $< 0.0001$ ). We also test for systematic pattern and performed Modified Hosmer Lemshow Test. We found the F-statistics to be  $1.4670$  with p-value  $0.1459$  and hence we reject the hypothesis that there is any systematic pattern of bias.

### 4.3 Measuring changes in insurance premium and income and wage

To test the hypotheses regarding new information and the competing age effects, we need to control for adverse shocks to both out-of-pocket insurance premium and income/wage. MEPS data provides detailed information about wage and salary income, hourly wage rate, employment status - currently employed, has a job to return to or not employed, whether or not self employed, number of hours per week worked, type of the establishment of current main job (CMJ) - private for profit, nonprofit or government, information about industry and occupation types for an individual's current main job. The data set also has information about whether the individual was offered health insurance from CMJ, whether the individual held insurance from CMJ, whether the individual disavowed the offer of health insurance from CMJ.

Using these variables, we are able to identify individuals, who, in the second period - 1) lost employment, 2) reduced the number of hours worked every week (part time), 3) lost the offer of insurance from current main job and 4) changed the type of establishment of the current main job. We create dummy variables indicating these changes. We also calculate the change in income to use in the regression below. There is a large variation in the administrative loading across industries. The general approach (see Pauly and Herring 1999) is to estimate this loading factor across industries, using estimates for group insurance loading by firm size as in Phelps (1997). These estimates then give each workers average administrative load based on the industry of employment. Since we are interested in the change in the administrative load, we take a different approach. If an industry has larger loading factor, the proportion of individuals who refuse the offer of insurance from the place of work would

be higher. As MEPS provides information about the offer of insurance from CMJ as well as whether the individuals refused the offer, we calculate the proportion of individuals who refused the offer of insurance from CMJ, for each industry. To calculate this fraction, we use all the individuals in the data set for each year. The measure for change in administrative load is, therefore, the difference between the proportions in two years.

Another variable of interest is the price of insurance that individuals face. However, it is not easy measure the premium faced by individuals. The difficulty arises because the insurance products heterogeneous and therefore individuals face multitude of prices. We could use the out of pocket premiums. However, it is well accepted in the literature that these out of pocket premiums may be endogenous; determined by the job selection behavior of the individuals. Therefore we resort to using exogenous proxies, for example, marginal tax rates which solves the potential endogeneity problem of using a price that cannot be affected by workers behavior<sup>17</sup>. Here, the implicit assumption is that workers choose jobs that will yield similar incomes and they believe that wages are reduced to finance his employers' contribution on his behalf. Hence, we also include marginal tax rates faced by individuals in the second period.

#### 4.4 Empirical Model and Estimation

We estimate the probability of renewing private health insurance coverage as a function of new information and age, controlling for other characteristics that affect the willingness and ability to renew the coverage. The dependent variable is one when individuals buy insurance in the second period (group or non-group coverage). The control variables included are race, sex, year, change in income, and change in the administrative load for each industry. Also included are dummy variables for loss of employment, loss of offer of health insurance from place of employment, transition from full time to part time worker. Since tax subsidy is a reasonable approximation of a price index for insurance when the decision to purchase vs. not to purchase is at hand, we also include marginal tax rates in the regressions.

We include in the study a sample of 1550 individuals who participated in the panel component of MEPS from 1997-2000. They were single in both base year and subsequent year, between 18 to 64 years of age in the base year, ineligible for any type of public health insurance in both the years and were privately insured in the base year. Table (7) contains the descriptive statistics for the study sample.

---

<sup>17</sup>Workers can of course choose among jobs with earnings sufficiently different and their marginal tax-rates would be changed and hence would be endogenous

We estimate the model using maximum likelihood logistic regression. As discussed above, our initial aim is to show that adverse surprises increased the likelihood of renewing insurance (H1), but our other aim is to show that the effect of new information (through its effect on the individuals beliefs about his health endowment) on the demand for insurance is greater for the young than for the old (H2). To achieve these goals, we first estimate the model on the entire sample. We then create dummy variables for young and old and regress the model using the entire sample but this time interacting the measure of new information with the dummy variables for young and old. Finally, we split the data set into young and old, and estimate separate models to differentiate between the effect of new information and age on the young and the old as implied by the economic model.

## 5 Results

In Figure (6), we examine how our measure of surprise or new information relates to age. If this measure is in fact capturing the unanticipated new information, then it should not show any systematic pattern with age. We also regress the square of this measure against age and, age and square of age. We find that in both the regressions the estimates of the coefficients were insignificant and the p-value of F - test of joint significance of coefficients was 0.87 and 0.47, respectively (see Table (10)). Thus, we find that there is no systematic relationship between age and the measure of unanticipated new information.

Table (11, equation 1) shows the results of multivariate regressions for the purchase of private insurance in the subsequent period for single individuals who had private insurance in the reference period. We include variables to control for changes in income and insurance premium. Our economic model predicts a non-monotonic (U-shaped) effect of age on the likelihood of continuing insurance coverage. Thus, we include age and square of age as regressors<sup>18</sup>. We find that the adverse surprise increases the likelihood of renewing insurance (estimate = 0.242; p-value = 0.06). We also find that age (estimate =  $-0.221$ ; p-value  $< 0.01$ ) and square of age (estimate = 0.003; p-value  $< 0.01$ ) are both significant and of opposite signs. This shows that the likelihood of insuring initially falls with age and then increases. The turn around age is 37 years. Individuals who leave the private sector for government sector job are more likely to remain insured and those who are now part time employees are

---

<sup>18</sup>We use likelihood ratio test (LR test) to test whether the inclusion of the square of age term significantly increases the fit. The test statistic is 9.27 with one degree of freedom (p-value  $< 0.01$ ).

less likely to continue coverage. Marginal tax rates are significant predictors for the renewal of coverage. Higher tax rates imply higher tax subsidy on the purchase of insurance.

Table (11, equation 2) shows the results for effects of new information on young and old separately. Since there is no well-accepted age to classify individuals as young and old, we take cue from the results of our first regression and define individuals below 37 years of age as young and above 37 years of age as old. Thus, we create dummy variables for young and old and interact these dummy variables with the measure of new information. As before, measures of decrease in income and increase in premium reduce the likelihood for continuing insurance. However, we find that new information has significant positive effect on the likelihood of young (estimate = 0.571; p-value < 0.01) to continue the insurance coverage and has no significant effect for the old<sup>19</sup>.

Table (12) shows the results for the new information and age for different sub sample split by age. We find that for young, there was no statistically significant relationship between age and likelihood of renewal of insurance coverage. However, adverse surprises have significant and positive (estimate = 0.759; p-value < 0.01) relationship with the likelihood of continuing insurance. In contrast, we find that there was no significant relationship between adverse surprises and the likelihood of remaining insured for old. On the other hand, for this sub sample, age increases the likelihood of continuing insurance coverage (estimate = 0.117; p-value < 0.01). In summary, for young, we find that adverse surprises are significant and positively related with the probability of continuing coverage, but aging has no significant effect<sup>20</sup>. Nonetheless, for old, there is no significant effect of new information, but positive and significant effect of age.

## 5.1 Magnitude of the effect of new information

As discussed in the previous chapters, new information influences the decision to continue insurance coverage through its impact on individuals' beliefs about their endowed health. We now discuss the magnitude of this effect on the demand for insurance. To express the magnitude of this effect, we keep all the variables constant at their mean values and calculate the change in probability when the new information

---

<sup>19</sup>We test for joint significant of adverse surprise for young and old. Test statistic is 7.149 and p-value is 0.02. We also test whether the effects are statistically same. The test statistic is 2.10 and p-value 0.14. Thus we conclude that the adverse surprise for young and old are jointly significant and the effect of adverse surprise for young is greater than for old.

<sup>20</sup>Our economic model predicts a negative age effect for the young.

variable is moved from the fifth percentile to ninety fifth percentile (see Table 13). When the sample is pooled across all ages, the movement of new information from 5<sup>th</sup> to 95<sup>th</sup> percentile increases the probability of continuing insurance by 1.57 percentage points (95% CI is -0.13 to 3.01 percentage points). However for the young the effect is much more stronger, relative to the effect on the old. For the young, the increase is 5.33 percentage points (95% CI is 0.39 to 13.60 percentage points) and that of old is 0.37 percentage points (95% CI is -1.62 to 1.20 percentage points). We get similar results when we use the split sample regressions. For the young the probability increases by 5.80 percentage points (95% CI is 1.47 to 18.29 percentage points) and that of old by 0.04 percentage points (95% CI is -0.31 to 0.18 percentage points)

## 6 Conclusion

We show imperfect information and subsequent learning about endowed health may lead individuals to exit the insurance market even in absence of adverse shocks to income and premium. We adapt the learning process of Jovanovic (1982). We show that new information affects individuals perception about their endowment. As more information accumulates, individuals beliefs about their endowment become more precise and therefore the consequence of new information decreases. Since individuals learn as they advance in age this implies that, the influence of the new information decreases as individuals advance in age. This result is important, as it gives a rationale to a seemingly curious finding by Sloan and Conover (1998) that elderly who I have sufficiently large deterioration of health is not more likely to renew their coverage than who does not'.

Aging is associated with biological depreciation of health that leads to increase in the size of future losses. Therefore, controlling for changes in income and premium, the demand for insurance should increase in age. We show that age need not have this monotonically increasing effect on the demand for insurance. This is due to the reduction in uncertainty that as implied by the learning process. Since utility is concave, and for given beliefs learning process implies a mean preserving compression in the posterior distribution, we show that the learning effect of age decreases the demand for insurance. Therefore, there are two opposing effects of age: learning effect decreasing the demand and the aging effect increasing it. Since learning effect wanes as individuals grow older and aging effect gets stronger as individuals approach death, we show that the likelihood of continuing insurance will first decrease and then increase with age.

In the second part of the paper, we use panel component of MEPS data for 1997-2000 to test this model. We select individuals who are single, non-elderly, ineligible for public insurance, and privately insured in the reference period and estimate their probability of insuring in the subsequent period. We find strong evidence of the effect of new information in the demand for insurance. Individuals who receive adverse surprise are more likely to continue coverage. We also find that the adverse surprise has a stronger impact on the decision of the young than on the decision of the old. In the pooled data regression, we see the non-monotonic pattern of age on the likelihood of continuing coverage. When data was pooled by age, we find that the effect of adverse surprise was stronger for the young and insignificant for the old, as predicted by the economic model. However, we did not find a strong negative relationship between age and continuation of coverage for the young as predicted by the economic model. What we did find was that age was not a significant predictor of continuing coverage.

## References

- ALARIE, Y., G. DIONNE, AND L. EECKHOUDT (1991): "Increases in Risk and the Demand for Insurance," in *Contributions to Insurance Economics*, ed. by G. Dionne. Kluwer Academic, Boston.
- APPLEBY, J. (2002): "More high-income Americans go without health insurance," *USA Today*, November.
- ATHEY, S. (2002): "Monotone Comparative Statics Under Uncertainty," *Quarterly Journal of Economics*, 117, 187–223.
- BAWA, V. S. (1975): "Optimal rules for ordering uncertain prospects," *Journal of Financial Economy*, 2(1), 95–121.
- BLACK, J. M., AND G. BULKLEY (1989): "A Ratio Criterion for Signing the Effects of an Increase in Uncertainty," *International Economic Review*, 30, 119–130.
- BLOMBERG, S. B., AND J. HARRINGTON (2000): "Rigid Extremists and Flexible Moderates With Applications to the U.S. Congress," *American Economic Review*, pp. 605–620.
- BLOUGH, D. K., C. W. MADDEN, AND M. C. HORNBROOK (1999): "Modeling risk using generalized linear models," *Journal of Health Economics*, 18, 153–171.

- BUNDORF, M. K., B. HERRING, AND M. PAULY (2005): “Health Risk, Income, and the Purchase of Private Health Insurance,” *NBER working paper series*, w11677.
- BUNDORF, M. K., AND M. PAULY (2002): “Is health insurance affordable for the uninsured?,” *NBER working paper series*, w9281.
- CUNNINGHAM, P., E. SCHAEFER, AND C. HOGAN (1999): “Who Declines Employer-Sponsored Health Insurance and is Uninsured?,” *Center for Studying Health System Change, Washington D.C.*, Issue Brief Number 22.
- CZAJKA, J. L. (1999): “Analysis of Children’s Health Insurance Patterns: Findings from the SIPP,” *Mathematica Policy Research, Inc.*, Report PR99-64.
- CZAJKA, J. L., AND C. OLSEN (2000): “The Effects of Trigger Events on Changes in Children’s Health Insurance Coverage,” *Mathematica Policy Research, Inc.*
- DEGROOT, M. (1970): *Optimal Statistical Decision*. McGraw Hill, New York, NY.
- DIAMOND, P., AND J. STIGLITZ (1974): “Increases in Risk and in Risk Aversion,” *Journal of Economic Theory*, 8, 337–60.
- DIONNE, G., AND L. EECKHOUDT (1987): “Proportional Risk Aversion, Taxation and Labor Supply Under Uncertainty,” *Journal of Economics*, 47, 353–366.
- DIONNE, G., L. EECKHOUDT, AND E. BRIYS (1989): “More on Insurance as a Giffen Good,” *Journal of Risk and Uncertainty*, 2, 415–420.
- DRÈZE, J., AND F. MODIGLIANI (1972): “Consumption decisions under uncertainty,” *Journal of Economic Theory*, 5(3), 308–335.
- DUAN, N., W. MANNING, AND ET. AL. (1983): “A Comparison of Alternative Models for the Demand for Medical Care,” *Journal of Business and Economics Statistics*, 1, 115–126.
- DYBVIK, P. H., AND S. LIPPMAN (1983): “An alternative characterization of decreasing absolute risk aversion,” *Econometrica*, 5(1), 223–24.
- EECKHOUDT, L., AND P. HANSEN (1980): “Minimum and Maximum Prices, Uncertainty and the Theory of the Competitive Firm,” *American Economic Review*, 70, 1064–1068.

- (1983): “Micro-economic Applications of Marginal Changes in Risk,” *European Economic Review*, 22, 167–176.
- EHRlich, I., AND G. S. BECKER. (1972): “Market Insurance, Self-Insurance, and Self-Protection,” *Journal of Political Economy*, 80, 623–48.
- FRONSTIN, P. (2002): “Sources of Health Insurance and Characteristics of the Uninsured: Analysis of the March 2002 Current Population Survey.,” *EBRI Issue Brief*, 242.
- GARRETT, B., AND J. HUDMAN (2002.): “Women Who Left Welfare: Health Care Coverage, Access, and Use of Health Services,” *The Henry J Kaiser Family Foundation*, No. 4058.
- GLIED, S., AND M. STABILE (2001): “Generation vexed: age-cohort differences in employer-sponsored health insurance coverage.,” *Health Affairs*, 20(1), 184–191.
- GOLLIER, C. (2001): *The Economics of Risk and Time*. MIT Press.
- HADAR, J., AND T. K. SEO (1990): “The Effects of Shifts in a Return Distribution on Optimal Portfolios,” *International Economic Review*, 31, 721–736.
- HADLEY, J. (2002): “Sicker and Poorer: The Consequences of Being Uninsured.,” *The Henry J Kaiser Family Foundation*.
- HANOCH, G., AND H. LEVY (1969): “The Efficiency Analysis of Choices Involving Risk,” *Review of Economic Studies*, 36, 335–346.
- JANG, Y., AND J. HADAR (1995): “A Note on Increased Probability of Loss and the Demand for Insurance,” *The Geneva Papers on Risk and Insurance Theory*, 20, 213–216.
- JENSEN, G. A. (1992): “The Dynamics of Health Insurance Among the Near Elderly,” *Medical Care*, 30 (7), 598–614.
- JONES, A. (2000): “Health Econometrics,” in *Handbook of Health Economics*, ed. by A. Culyer, and J. Newhouse. Elsevier, Amsterdam.
- JOVANOVIC, B. (1982): “Selection and Evolution of Industry,” *Econometrica*, 50 (3).
- KESSLER, D., AND E. N. WOLFF (1991): “A Comparative Analysis of Household Wealth Patterns in France and the United States,” *Review of Income and Wealth*, Series 37, No. 3, 249–66.

- LEVY, H. (1998): *Stochastic Dominance*. Kluwer Academic Publishers, Norwell, MA.
- MEYER, J., AND M. B. ORMISTON (1983): “The Comparative Statics of Cumulative Distribution Function Changes for the Class of Risk Averse Agents,” *Journal of Economic Theory*, 31, 153–169.
- (1985): “Strong Increases in Risk and Their Comparative Statics,” *International Economics Review*, 26, 425–437.
- (1989): “Deterministic Transformations of Random Variables and the Comparative Statics of Risk,” *Journal of Risk and Uncertainty*, 2, 179–188.
- MILGROM, P. (1981): “Good News and Bad News: Representation Theorems and Applications,” *Bell Journal of Economics*, 12, 380–391.
- MILLS, R. (2002): “Health insurance coverage: 2000.,” *US Census Bureau Current Population Reports*, 20(1), 169–177.
- PAULY, M. V., AND B. HERRING (1999): *Pooling Health Insurance Risks*. AEI Press, Washington DC.
- PHELPS, C. E. (1997): *Health Economics*. Addison-Wesley.
- ROTHSCHILD, M., AND J. STIGLITZ (1970): “Increasing Risk: I. A Definition,” *Journal of Economic Theory*, 2, 225–243.
- SANDMO, A. (1970): “The Effects of Uncertainty of Saving Decisions,” *Review of Economic Studies*, 37, 353–360.
- (1971): “On the Theory of the Competitive Firm Under Price Uncertainty,” *American Economic Review*, 61, 65–73.
- SCHAEFER, S., AND L. MCGINLEY (2003): “Number of Americans Who Lack Health-Care Coverage Is Rising: Census Bureau Counts 43.6 Million As Employer-Based Plans Shrink,” *Wall Street Journal*.
- SLOAN, F. A., AND C. J. CONOVER (1998): “Life Transitions and Health Insurance Coverage of the Near Elderly,” *Medical Care*, 36 (2), 110–125.
- SWARTZ, K. (1994): “Dynamics of People Without Health Insurance. Don’t let the Numbers Fool You,” *The Journal of the American Medical Association*, 274(1), 64–66.

- SWARTZ, K., J. MARCOTTE, AND T. D. MCBRIDE (1993): “Personal Characteristics and Spells Without Health Insurance,” *Inquiry*, 30, 64–76.
- SWARTZ, K., T. D. MCBRIDE, AND J. MARCOTTE (1993): “Spells Without Health Insurance: The Distribution of Durations When Left-Censored Spells are Included,” *Inquiry*, 30, 77–83.
- TESFATSION, L. (1976): “Stochastic Dominance and the Maximization of Expected Utility,” *Review of Economic Studies*, 43(2), 301–15.
- (1981): “Dynamic Investment, Risk Aversion, and Foresight Sensitivity,” *Journal of Economic Dynamics and Control*, 3, 65–96.
- VENEZIA, I., AND H. LEVY (1983): “Optimal Multiperiod Insurance Contract,” *Insurance: Mathematics and Economics*, 2, 199–208.
- YAARI, M. (1969): “Some remarks on measures of risk aversion and their uses,” *Journal of Economic Theory*, 1, 315–329.
- YEGIAN, J., D. G. POCKELL, AND E. K. MURRAY (2003): “To Buy or Not To Buy: A Profile of California’s Non-Poor Uninsured,” *A Survey by California HealthCare Foundation and the Field Research Corporation*.

# Appendix

## A Posterior Distributions

### Posterior Distribution of $\theta$

Let  $\eta^n = \{\eta_1, \eta_2, \dots, \eta_n\}$  be  $n$  observations drawn from  $N(\theta, \sigma_\epsilon^2)$ ,  $\theta$  unknown and  $\sigma_\epsilon^2$  known.  $\theta$  is drawn from  $N(\bar{\theta}, \sigma_\theta^2)$ .

$$p(\eta^n | \theta, \sigma_\epsilon^2) = \prod_{i=1}^n p(\eta_i | \theta, \sigma_\epsilon^2) = \frac{1}{(2\pi\sigma_\epsilon^2)^{\frac{n}{2}}} \exp \left[ -\frac{1}{2\sigma_\epsilon^2} [(n-1)s^2 + n(\theta - \bar{\eta}_n)^2] \right]$$

where  $\bar{\eta}_n = \frac{1}{n} \sum_{i=1}^n \eta_i$  and  $s^2 = \frac{1}{n-1} \sum_{i=1}^n (\eta_i - \bar{\eta}_n)^2$ . Following Bayes' Theorem,  $p(\theta | \eta^n, \sigma_\epsilon^2) \propto p(\theta)p(\eta^n | \theta, \sigma_\epsilon^2)$ , therefore,

$$p(\theta | \eta^n, \sigma_\epsilon^2) \propto \exp \left[ -\frac{1}{2} \left( \frac{\sigma_\theta^2 + \frac{\sigma_\epsilon^2}{n}}{\sigma_\theta^2 \frac{\sigma_\epsilon^2}{n}} \right) \left( \theta - \frac{\bar{\eta}_n \sigma_\theta^2 + \bar{\theta} \frac{\sigma_\epsilon^2}{n}}{\sigma_\theta^2 + \frac{\sigma_\epsilon^2}{n}} \right)^2 \right]$$

from which we see that  $\theta$  is normally distributed, a posteriori, with mean and variance

$$\begin{aligned} E(\theta | \eta^n) &= \lambda(n)\bar{\eta}_n + (1 - \lambda(n))\bar{\theta} \\ \text{Var}(\theta | \eta^n) &= \frac{\sigma_\epsilon^2 \sigma_\theta^2}{n\sigma_\theta^2 + \sigma_\epsilon^2} \end{aligned}$$

$$\text{where } \lambda(n) = \frac{n\sigma_\theta^2}{n\sigma_\theta^2 + \sigma_\epsilon^2}$$

### Posterior Distribution of $x_n$

We know that  $\eta$  is normally distributed with posterior mean and variance as in (2) and (3).  $x_n = \xi(\eta_n) = e^{\eta_n}$ , so  $\ln(x_n) = \eta_n$ . Thus  $x_n$  follows log normal distribution, with posterior mean and variance a function of posterior mean and variance of  $\eta_n$ .

$$\begin{aligned} E[x_n | \text{past}] &= \tilde{x}_n = \exp \left( \lambda(n)\bar{\eta}_n + \bar{\theta}(1 - \lambda(n)) + \frac{\sigma_\theta^2 \sigma_\epsilon^2}{n\sigma_\theta^2 + \sigma_\epsilon^2} \right) \\ \text{Var}[x_n | \text{past}] &= \left[ \exp \left( \lambda(n)\bar{\eta}_n + \bar{\theta}(1 - \lambda(n)) + \frac{\sigma_\theta^2 \sigma_\epsilon^2}{n\sigma_\theta^2 + \sigma_\epsilon^2} \right) \right]^2 \\ &\quad \times \left[ \exp \left( \frac{\sigma_\theta^2 \sigma_\epsilon^2}{n\sigma_\theta^2 + \sigma_\epsilon^2} \right) - 1 \right] \end{aligned}$$

$$\text{where } \lambda(n) = \frac{n\sigma_\theta^2}{n\sigma_\theta^2 + \sigma_\epsilon^2}$$

## B Proof of Propositions

To prove the two propositions, we need to show that an increase in  $\tilde{x}$ , for a given  $n$ , induces first order stochastic dominance (FOSD) and an increase in  $n$ , for a given  $\tilde{x}$  induces second order stochastic dominance (SOSD).

## B.1 Proof of Proposition 1

The posterior distribution of  $\theta$  has *monotone likelihood ratio property*<sup>21,22</sup>. Densities with the strict MLRP have the property that all signals are *comparable*<sup>23</sup> and imply First Order Stochastic Dominance (Milgrom (1981)). In **Proposition 3** pp 384 of Milgrom (1981), he shows that when signals are comparable then any real valued variable may be modeled with MLRP. Using **Proposition 2 and 3** pp 383-4, we get the following lemma.

**Lemma 1** *Let  $\xi : \mathfrak{R} \rightarrow \mathfrak{R}$  be any increasing function and define  $\tilde{x}$  as follows:*

$$\tilde{x} = \int \xi(\theta) dF(\theta|\bar{\eta})$$

where  $F$  is any non degenerate prior for  $\theta$ , which has MLRP property in that the signals are comparable, then densities of  $\tilde{x}$  have strict MLRP

**Proof.** Since signals are comparable,  $\tilde{x}_h > \tilde{x}_l$  if and only if  $\bar{\eta}_h > \bar{\eta}_l \Rightarrow F(\theta|\bar{\eta}_h) < F(\theta|\bar{\eta}_l)$ . Applying **Proposition 2** from Milgrom (1981), it follows that the densities of  $\tilde{x}$  have strict MLRP.

### Proof of Proposition 1

$$E[u_0|\tilde{x}_n, n] = \int_0^\infty u(w - m\delta^n x) dF(x|\tilde{x}_n, n) \quad (18)$$

By assumption  $u(\cdot)$  is continuous and decreasing in  $x$  and  $f(x|\tilde{x}_n, n)$  is a posterior distribution of  $x$  with sufficient statistics  $\tilde{x}_n$  and  $n$ . Thus, the function inside the integral sign is continuous in  $x$ . Monotone likelihood ratio property with respect to  $\tilde{x}_n$  implies first order stochastic dominance<sup>24</sup>,  $\tilde{x}_h > \tilde{x}_l \Rightarrow F(x|\tilde{x}_h) < F(x|\tilde{x}_l)$

$$\Delta_{\tilde{x}_n} E[u_0] = \int_0^\infty u(w - m\delta^n x) [dF(x|\tilde{x}_n^h, n) - dF(x|\tilde{x}_n^l, n)]$$

Integrating by parts and using the assumption that  $u(w - m\delta^n x)$  is finite and continuous for all finite  $x$

$$\begin{aligned} \Delta_{\tilde{x}_n} E[u_0] &= [F(x|\tilde{x}_n^h, n) - F(x|\tilde{x}_n^l, n)] u(w - m\delta^n x) \Big|_0^\infty \\ &+ m\delta^n \int_0^\infty u'(w - m\delta^n x) [F(x|\tilde{x}_n^h, n) - F(x|\tilde{x}_n^l, n)] dx < 0 \end{aligned} \quad (19)$$

The first term above is zero, and second is negative because  $u'(\cdot)$  is positive and  $F(x|\tilde{x}_h) < F(x|\tilde{x}_l) \forall x$ .

## B.2 Proof of Proposition 2

The following definitions follow Bawa (1975):

<sup>21</sup>Let  $\eta \subset \mathfrak{R}$ . The densities  $\{f(\cdot|\theta)\}$  have *strict* monotone likelihood ratio property (**MLRP**) if for every  $\eta_h > \eta_l$  and  $\theta_h > \theta_l$ , the following holds:  $\frac{f(\eta_h|\theta_h)}{f(\eta_h|\theta_l)} > (\geq) \frac{f(\eta_l|\theta_h)}{f(\eta_l|\theta_l)}$ . If strict inequality is changed to weak inequality, then it is called (weak) monotone likelihood ratio property.

<sup>22</sup> $\frac{f(\eta|\theta_h)}{f(\eta|\theta_l)} = \exp\left\{-\frac{1}{2\sigma_\varepsilon^2} [\theta_h - \theta_l] [(\theta_h + \theta_l) - 2\eta]\right\}$ , which is increasing in  $\eta$  if  $\theta_h > \theta_l$ .

<sup>23</sup>Write the defn of comparable.

<sup>24</sup>The upper limit of the integration in equation (18) is not finite, therefore an additional assumption is  $\int u dF(x|\tilde{x}_n^h, n) - \int u dF(x|\tilde{x}_n^l, n)$  is not of the form  $\infty - \infty$  (see Tsefatson (1976), pp 301-302) required.

**Definition 1** Let  $F_0$  be the class of distribution functions that are characterized by a location parameter ( $l$ ) and a scale parameter ( $s$ ) as follows

$$F_0 = \left\{ F(x) \mid x \in \mathfrak{R}, F(x) = \psi \left( \frac{x - l_F}{s_F} \right) \forall x \text{ and } s_F > 0 \right\}$$

and

**Definition 2** Let  $F_1$  be the class of distribution functions in which the monotonic transformation of random variables belong to the location and scale parameter family defined by  $F_0$ .

$$F_1 = \left\{ F(x) \mid x \in \mathfrak{R}, F(x) = \psi \left( \frac{\phi(x) - l_F}{s_F} \right) \text{ with } \phi'(x) \geq 0 \forall x \text{ and } s_F > 0 \right\}$$

The posterior distribution of  $\theta$  belongs to the class of distributions defined by  $F_0$ , with scale parameter ( $s_F$ ) which is a decreasing function of  $n$ . Similarly, the posterior distribution of  $x$  belongs to the class of distributions defined by  $F_1$ , with  $\phi(x) = \ln(x)$ <sup>25</sup>.

**Lemma 2** Let  $F(x|n+1)$  and  $F(x|n)$  be two distribution functions, with the same mean ( $\tilde{x}$ ) belonging to the class  $F_1$ . If  $s_{n+1} \neq s_n$ , then they intersect only once.

**Proof.** By definition,  $F(x|n+1) = \psi \left( \frac{\phi(x) - l_{n+1}}{s_{n+1}} \right)$  and  $F(x|n) = \psi \left( \frac{\phi(x) - l_n}{s_n} \right)$ . Let  $x'$  solve  $F(x'|n+1) = F(x'|n)$ . Then  $\phi(x') = \frac{l_{n+1}s_n - l_n s_{n+1}}{s_n - s_{n+1}}$ . Since  $\phi(\cdot)$  is strictly increasing,  $x'$  is unique. Thus the two distribution functions intersect only once.

**Lemma 3** If  $s_{n+1} < s_n$ , then  $F(x|n+1)$  cuts  $F(x|n)$  from below.

**Proof.** From Lemma 2  $\phi(x') = \frac{l_{n+1}s_n - l_n s_{n+1}}{s_n - s_{n+1}}$  and since  $\phi(\cdot)$  is strictly increasing, then  $\phi(x) < \phi(x')$ ,  $\forall x < x'$ .  $F(x|n+1) - F(x|n) = \psi \left( \frac{\phi(x) - l_{n+1}}{s_{n+1}} \right) - \psi \left( \frac{\phi(x) - l_n}{s_n} \right)$ . Therefore the sign of  $F(x|n+1) - F(x|n)$  depends on the sign of  $\frac{\phi(x) - l_{n+1}}{s_{n+1}} - \frac{\phi(x) - l_n}{s_n} = \frac{1}{[s_{n+1}s_n]} \frac{\phi(x') - \phi(x)}{[s_{n+1} - s_n]}$ . From above we know that the numerator is positive  $\forall x < x'$ , so the sign of  $F(x|n+1) - F(x|n)$  depends on the sign of  $s_{n+1} - s_n$ . Hence, if  $s_{n+1} < s_n$ , then  $F(x|n+1) < F(x|n) \forall x < x'$ .  $F(x|n+1)$  cuts  $F(x|n)$  from below.

**Proof of Proposition 2**

$$E[u_0|\tilde{x}_n, n] = \int_0^\infty u(w - mx) dF(x|\tilde{x}, n) \quad (20)$$

$u(\cdot)$  is continuous, twice differentiable with  $u'(\cdot) > 0$  and  $u''(\cdot) < 0$ , and  $F(x|\tilde{x}, n)$  is a posterior distribution of  $x$  with sufficient statistics  $\tilde{x}$  and  $n$ , belonging to the class of distributions defined by  $F_1$ .

---

<sup>25</sup>Bawa (1975) has shown that comparing distributions which are characterized by two parameters amounts to comparing *some* two parameters, but these parameters are *neither* mean and variance, *nor* location and scale. He has shown that a larger mean is a necessary condition for dominance for all the distribution functions and when variance is a monotonic function *only* of the scale parameter, variance may be used to compare two distribution functions, as in the case of normal distribution. If, however, variance is a function of other parameters as well, as in the case of log normal distribution where variance is a function of mean and the scale parameter, scale parameter is the appropriate parameter to be used to compare the distribution functions (see Bawa (1975), Theorem 5, 8, 9 and 11). Levy (1998) has shown, for example, that if  $x$  has a log-normal distribution, the appropriate parameter to compare is the coefficient of variation,  $\frac{\sqrt{Var(x)}}{E(x)}$  which is proportional to the variance of  $\ln(x)$  (see Levy (1998) pp 201-202).

The necessary and sufficient condition for  $F(x|n+1)$  to dominate  $F(x|n)$  in the sense of second order stochastic dominance for every  $u(\cdot)$  which is non decreasing and concave is  $\int_0^x [F(t|n) - F(t|n+1)] dt \geq 0 \forall x$  with strict inequality for some  $x$  and  $F_{n+1} \neq F_n$  for some  $x$  (Hanoch and Levy (1969, Theorem 2 pp. 338-340)). From Lemma 2, for a given mean  $\tilde{x}$ ,  $F(x|n+1)$  and  $F(x|n)$  intersect only once (so there is some  $x$  for which  $F_{n+1} \neq F_n$ ) and from Lemma 3  $F(x|n+1)$  cuts  $F(x|n)$  from below, which implies that  $\int_0^x [F(t|n) - F(t|n+1)] dt \geq 0 \forall x$ , with strict inequality for some  $x$ <sup>26</sup>. Hence  $E[u_0|n]$  increases in  $n$  for a given  $\tilde{x}$ .

## C Proof of Proposition 4

### C.1 Proposition 4 a

**Proof:**<sup>27</sup> Insurance premium is independent of  $\tilde{x}$ . Let  $\alpha_a^*(\tilde{x}_l, n)$  solve

$$\int_0^\infty [u'(w - m\delta^n[(1 - \alpha_a^*(\tilde{x}_l, n))x + \alpha_a^*(\tilde{x}_l, n)\lambda p])m\delta^n(x - \lambda p)] f(x|\tilde{x}_l, n) dx = 0$$

to show  $\alpha_a^*(\tilde{x}_h, n) > \alpha_a^*(\tilde{x}_l, n) \forall \tilde{x}_h > \tilde{x}_l$ . For brevity, normalize  $m$  and  $\delta$  to one. Since  $n$  is constant, we drop it from the notation.

Let  $l(x) = \frac{f(x|\tilde{x}_h)}{f(x|\tilde{x}_l)}$ ,  $\tilde{x}_h > \tilde{x}_l$ . From Lemma 1,  $l(x)$  (strictly) increases in  $x$ . Let  $g(x)$ <sup>28</sup> =  $\left[ u'(w - [(1 - \alpha_a^*(\tilde{x}_l))x + \alpha_a^*(\tilde{x}_l)\lambda p]) (x - \lambda p) \right]$ . By assumption,  $u'(\cdot)$  is non-negative. Therefore,  $g(x) < 0 \forall x < \lambda p$  and  $g(x) > 0 \forall x > \lambda p$ , i.e.  $g(\cdot)$  has single crossing property<sup>29</sup>, and  $\alpha_a^*(\tilde{x}_l)$  solves  $\int_0^\infty g(x)f(x|\tilde{x}_l)dx = 0$ . Therefore,

$$\begin{aligned} \int_0^\infty g(x)f(x|\tilde{x}_h)dx &= \int_0^\infty g(x) \left[ \frac{f(x|\tilde{x}_h)}{f(x|\tilde{x}_l)} \right] f(x|\tilde{x}_l)dx = \int_0^\infty g(x)l(x)f(x|\tilde{x}_l)dx \\ &= \int_0^{\lambda p} g(x)l(x)f(x|\tilde{x}_l)dx + \int_{\lambda p}^\infty g(x)l(x)f(x|\tilde{x}_l)dx. \end{aligned}$$

Since  $l(x) < l(\lambda p)$  and  $g(x) < 0 \forall x < \lambda p$ , this implies,  $|g(x)l(x)| < |g(x)l(\lambda p)| \forall x < \lambda p$  or  $-|g(x)l(x)| > -|g(x)l(\lambda p)| \forall x < \lambda p$ . Similarly, since  $l(x) > l(\lambda p)$  and  $g(x) > 0 \forall x > \lambda p$ , therefore  $g(x)l(x) > g(x)l(\lambda p) \forall x > \lambda p$ . Using this we get,

$$\begin{aligned} \int_0^\infty g(x)f(x|\tilde{x}_h)dx &> -l(\lambda p) \int_0^{\lambda p} |g(x)| f(x|\tilde{x}_l)dx + l(\lambda p) \int_{\lambda p}^\infty g(x)f(x|\tilde{x}_l)dx \\ &= l(\lambda p) \int_0^\infty g(x)f(x|\tilde{x}_l)dx \end{aligned} \quad (21)$$

Since  $\int_0^\infty g(x)f(x|\tilde{x}_l)dx = 0$  and  $l(\lambda p)$  is positive constant, we infer that at  $\alpha = \alpha_a^*(\tilde{x}_l; p)$

$$\int_0^\infty [u'(w - [(1 - \alpha_a^*(\tilde{x}_l))x + \alpha_a^*(\tilde{x}_l)\lambda p]) (x - \lambda p)] f(x|\tilde{x}_h)dx > 0.$$

Therefore,  $\alpha_a^*(\tilde{x}_h) > \alpha_a^*(\tilde{x}_l)$ .

<sup>26</sup>See Levy (1998) pp. 212 for the discussion.

<sup>27</sup>We exploit the *log super modularity* of probability density function with respect to  $\tilde{x}$  (see Lemma 1) and *single crossing property* of marginal utility. Athey (2002) has shown that these two conditions are necessary and sufficient for monotone comparative statics results (see Lemma 5 pp 201 and Theorem 2 pp 200).

<sup>28</sup>Since premium is independent of  $\tilde{x}$ ,  $g(x)$  does not depend on  $\tilde{x}$ .

<sup>29</sup>This implies that  $u(\cdot)$  satisfies single crossing in two variables  $(\alpha, x)$ , i.e.  $[u(\alpha_h, x) - u(\alpha_l, x)]$  satisfies single crossing. In other words, the *incremental* returns to  $\alpha$  cross zero at most once, from below (see Athey(2002) Definition 2 pp 190).

## C.2 Proposition 4 b

**Lemma 4** If  $u(\cdot)$  is DARA,  $h(x) = \frac{u'(z_h)}{u'(z_l)}$ , where  $z_i = w - m\delta^n[(1-\alpha)x + \alpha\lambda\tilde{x}_i]$ ,  $i = h$  and  $l$ , then  $\frac{\partial h}{\partial x} > 0$ .

**Proof:** Since  $\tilde{x}_h > \tilde{x}_l$ , it follows that  $z_l > z_h, \forall x$  and  $\frac{\partial z_h}{\partial x} = \frac{\partial z_l}{\partial x} = -m\delta^n(1-\alpha)$ . Taking log and differentiating  $h(x)$  with respect to  $x$  we get

$$h'(x) = D \left[ \left( -\frac{u''(z_h)}{u'(z_h)} \right) - \left( -\frac{u''(z_l)}{u'(z_l)} \right) \right]$$

where  $D = m\delta^n(1-\alpha) \left[ \frac{u'(z_h)}{u'(z_l)} \right] > 0$ . Since it is a DARA utility function and  $z_h < z_l$ , we have  $\left[ \left( -\frac{u''(z_h)}{u'(z_h)} \right) - \left( -\frac{u''(z_l)}{u'(z_l)} \right) \right] > 0$ . Consequently  $h(x)$  increases in  $x$ .

### Proof of Proposition 4 b

**Proof:** Insurance premium =  $m\delta^n\lambda\tilde{x}$ . Let  $\alpha^*(\tilde{x}_l)$  solve

$$\int_0^\infty [u'(w - m\delta^n[(1-\alpha^*(\tilde{x}_l)x + \alpha^*(\tilde{x}_l)\lambda\tilde{x}_l])]m\delta^n(x - \lambda\tilde{x}_l)] f(x|\tilde{x}_l, n)dx = 0$$

For brevity, normalize  $m$  and  $\delta$  to one. Since  $n$  is constant, we drop it from the notation.

Let  $l_1(x) = \frac{f(x|\tilde{x}_h)}{f(x|\tilde{x}_l)}$ ,  $l_2(x) = \frac{u'(z_h)}{u'(z_l)}$ , where  $z_i = w - m\delta^n[(1-\alpha)x + \alpha\lambda\tilde{x}_i]$  and  $l(x) = l_1(x) \cdot l_2(x)$ . From Lemma 1,  $l_1(x)$  increases in  $x$  and from Lemma 4,  $l_2(x)$  increases in  $x$ , therefore  $l(x)$  increases in  $x$ .

Let  $g(x) = \left[ u'(z_l)(x - \lambda\tilde{x}_l) \right]$ . By assumption,  $u'(\cdot)$  is non-negative. Therefore,  $g(x) < 0 \forall x < \lambda\tilde{x}_l$  and  $g(x) > 0 \forall x > \lambda\tilde{x}_l$ , i.e.,  $g(\cdot)$  has a single crossing property, and  $\alpha^*(\tilde{x}_l)$  solves  $\int_0^\infty g(x)f(x|\tilde{x}_l)dx = 0$ . Therefore,

$$\begin{aligned} \int_0^\infty u'(z_h)(x - \lambda\tilde{x}_h)f(x|\tilde{x}_h)dx &= \int_0^\infty \frac{u'(z_h)}{u'(z_l)} \frac{f(x|\tilde{x}_h)}{f(x|\tilde{x}_l)} u'(z_l)(x - \lambda\tilde{x}_l)f(x|\tilde{x}_l)dx \\ &\quad - \lambda(\tilde{x}_h - \tilde{x}_l) \int_0^\infty u'(z_h)f(x|\tilde{x}_h)dx \\ &= \int_0^\infty g(x)l(x)f(x|\tilde{x}_l)dx \\ &\quad - \lambda(\tilde{x}_h - \tilde{x}_l) \int_0^\infty u'(z_h)f(x|\tilde{x}_h)dx \end{aligned}$$

From (21), the first term is positive. The second term is also positive, since  $\tilde{x}_h > \tilde{x}_l$ ,  $u'(\cdot) > 0$  and  $f(x|\tilde{x}_h) > 0 \forall x$ . Therefore  $\alpha^*(\tilde{x}_h)$  is greater than or less than  $\alpha^*(\tilde{x}_l)$ .

## D Proof of Proposition 5

**Lemma 5** If  $u(\cdot)$  is DARA, and  $h(x) = \frac{u'(z_1)}{u'(z_0)}\delta$ , where  $z_i = w - m\delta^{n+i}[(1-\alpha)x + \alpha\lambda\tilde{x}]$ ,  $i = 0 \text{ \& } 1$ , then  $\frac{\partial h}{\partial x} > 0$ .

**Proof:** Since  $\delta > 1$ , it follows that  $z_0 > z_1$ ,  $\forall x$  and  $\frac{\partial z_i}{\partial x} = -m\delta^{n+i}(1-\alpha)$ . Taking log and differentiating  $h(x)$  with respect to  $x$  we get

$$\begin{aligned} h'(x) = \frac{\partial h(x)}{\partial x} &= \left[ \delta \frac{u'(z_1)}{u'(z_0)} \right] \left[ -\frac{\partial z_0}{\partial x} \right] \left[ \left( -\frac{u''(z_1)}{u'(z_1)} \delta \right) - \left( -\frac{u''(z_0)}{u'(z_0)} \right) \right] \\ &= D \left[ \left( -\frac{u''(z_1)}{u'(z_1)} \delta \right) - \left( -\frac{u''(z_0)}{u'(z_0)} \right) \right] \end{aligned}$$

where  $D = m\delta^{n+1}(1-\alpha) \left[ \frac{u'(z_1)}{u'(z_0)} \right] > 0$ . Since it is a DARA utility function and  $z_1 < z_0$ , we have  $\left[ \left( -\frac{u''(z_1)}{u'(z_1)} \right) - \left( -\frac{u''(z_0)}{u'(z_0)} \right) \right] > 0$ . Again since  $\delta > 1$ ,  $\left[ \left( -\frac{u''(z_1)}{u'(z_1)} \right) \delta - \left( -\frac{u''(z_0)}{u'(z_0)} \right) \right] > 0$ . Consequently  $h(x)$  increases in  $x$ .

**Proof of Proposition 5:**

Since the posterior distribution is fixed, we write it as  $f(x)$ .  $\alpha_d^*(n)$  solves

$$\int_0^\infty [u'(z_0)\delta^n(x - \lambda\tilde{x})] f(x)dx = 0 \quad (22)$$

where  $z_i = w - m\delta^{n+i}[(1-\alpha_d^*(n))x + \alpha_d^*(n)\tilde{x}]$ . We have to show that for a given  $\tilde{x}$  and fixed posterior distribution,  $\alpha_d^*(n+1) > \alpha_d^*(n)$ .

Let  $h(x) = \frac{u'(z_1)}{u'(z_0)}\delta$ . Since  $u(\cdot)$  is strictly concave,  $h(x) > 1$  and from Lemma 5,  $h(x)$  increases in  $x$ .

$$\begin{aligned} \int_0^\infty u'(z_1)\delta^{n+1}(x - \lambda\tilde{x})f(x)dx &- \int_0^\infty u'(z_0)\delta^n(x - \lambda\tilde{x})f(x)dx \\ &= \int_0^\infty [u'(z_1)\delta - u'(z_0)]\delta^n(x - \lambda\tilde{x})f(x)dx \\ &= \int_0^\infty [h(x) - 1]u'(z_0)\delta^n(x - \lambda\tilde{x})f(x)dx \end{aligned}$$

Let  $g(x) = [u'(z_0)\delta^n(x - \lambda\tilde{x})]$ , and let  $x_0 = \lambda\tilde{x}$ ; then  $g(x_0) = 0$ . By assumption  $u'(\cdot)$  is non-negative, therefore  $g(x) < 0 \forall x < x_0$  and  $g(x) > 0 \forall x > x_0$ , i.e.,  $g(\cdot)$  has a single crossing property. Therefore,

$$\begin{aligned} \int_0^\infty [h(x) - 1]u'(z_0)\delta^n(x - \lambda\tilde{x})f(x)dx &= \int_0^\infty [h(x) - 1]g(x)f(x)dx \\ &= \int_0^{x_0} [h(x) - 1]g(x)f(x)dx + \int_{x_0}^\infty [h(x) - 1]g(x)f(x)dx. \end{aligned}$$

Since  $h(x) < h(x_0)$  and  $g(x) < 0 \forall x < x_0$ , this implies,  $|g(x)|[h(x) - 1] < |g(x)|[h(x_0) - 1] \forall x < x_0$  or  $-|g(x)|[h(x) - 1] > -|g(x)|[h(x_0) - 1] \forall x < x_0$ . Similarly, since  $h(x) > h(x_0)$  and  $g(x) > 0 \forall x > x_0$ , this implies,  $g(x)[h(x) - 1] > g(x)[h(x_0) - 1] \forall x > x_0$ . Using this we get,

$$\begin{aligned} \int_0^\infty [h(x) - 1]u'(z_0)\delta^n(x - \lambda\tilde{x})f(x)dx & > -[h(x_0) - 1] \int_0^{x_0} |g(x)| f(x)dx + [h(x_0) - 1] \int_{x_0}^\infty g(x)f(x)dx \\ &= [h(x_0) - 1] \int_0^\infty g(x)f(x)dx = 0 \end{aligned} \quad (23)$$

Hence, at  $\alpha = \alpha_d^*(n)$

$$\int_0^\infty u'(z_1)\delta^{n+1}(x - \lambda\tilde{x})f(x)dx > \int_0^\infty u'(z_0)\delta^n(x - \lambda\tilde{x})f(x)dx = 0$$

and therefore,  $\alpha_d^*(n+1) > \alpha_d^*(n)$ .

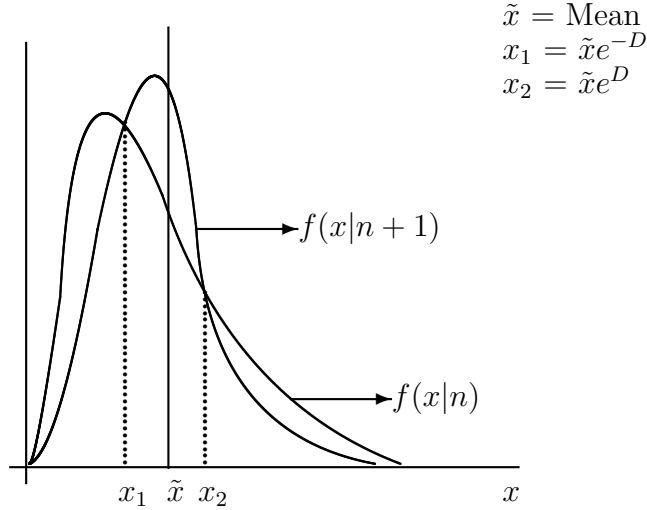


Figure 4: Relatively Strong Decrease in Risk

## E Proof of Proposition 6

**Definition 3 (Black and Bukley (1989), pp 120)** Let  $F(x|n)$  and  $F(x|n+1)$  be two cumulative distribution functions of a random variable  $x$ , then  $F(x|n+1)$  represents relatively strong decrease in risk compared to  $F(x|n)$  if

- (a)  $\int_0^{\infty} [F(x|n) - F(x|n+1)] dx = 0$
- (b) For all points in  $[x_1, x_2]$ ,  $f(x|n+1) > f(x|n)$  and for all the points outside this interval  $f(x|n+1) < f(x|n)$ , i.e. the two probability distribution functions (p.d.f.'s) intersect twice at  $x_1$  and  $x_2$ .
- (c)  $r(x; n) = \frac{f(x|n+1)}{f(x|n)}$  is non decreasing (increasing) in the interval  $[0, x_1]$ .
- (d)  $r(x; n) = \frac{f(x|n+1)}{f(x|n)}$  is non increasing (decreasing) in the interval  $(x_2, \infty)$ .

**Lemma 6** For a given mean,  $\tilde{x}$ , let the posterior distribution of log normal  $x$  with  $n$  observations be denoted by  $F(x|n)$  and the posterior distribution with  $n+1$  observations be represented by  $F(x|n+1)$ ; then  $F(x|n+1)$  represents a relative strong decrease in risk compared to  $F(x|n)$ .

### Proof

Since mean is constant, (a) is satisfied. The probability density function can be written as

$$f(x|i) = \frac{1}{x\sigma_i\sqrt{2\pi}} \exp \left[ -\frac{1}{2} \left( \frac{\log(x) - \mu_i}{\sigma_i} \right)^2 \right] \quad i = n, n+1$$

With mean constant, let  $\log(\tilde{x}) = \mu_n + \frac{1}{2}\sigma_n^2 = \mu_{n+1} + \frac{1}{2}\sigma_{n+1}^2$ . Let  $a(x) = \log(x) - \log(\tilde{x})$ . Using this write the probability density functions as

$$f(x|i) = \frac{1}{x\sigma_i\sqrt{2\pi}} \exp \left[ -\frac{1}{2} \left( \frac{a(x) + \frac{1}{2}\sigma_i^2}{\sigma_i} \right)^2 \right] \quad i = n, n+1$$

We have to find  $x_1$  and  $x_2$  such that  $f(x|n) = f(x|n+1)$  at  $x = x_1, x_2$ . Since  $x$  and  $\sqrt{2\pi}$  are nonnegative, we have

$$\frac{1}{\sigma_n} \exp \left[ -\frac{1}{2} \left( \frac{a(x) + \frac{1}{2}\sigma_n^2}{\sigma_n} \right)^2 \right] = \frac{1}{\sigma_{n+1}} \exp \left[ -\frac{1}{2} \left( \frac{a(x) + \frac{1}{2}\sigma_{n+1}^2}{\sigma_{n+1}} \right)^2 \right]$$

Taking log and noting that  $\sigma_n > \sigma_{n+1}$ , we get

$$\log(x) = \log(\tilde{x}) \pm \sqrt{\left[ \frac{\log(\sigma_n^2) - \log(\sigma_{n+1}^2)}{\sigma_n^2 - \sigma_{n+1}^2} + \frac{1}{4} \right] \sigma_n^2 \sigma_{n+1}^2} \quad (24)$$

Therefore, the points of intersection,  $x_1$  and  $x_2$ , are as follows

$$\begin{aligned} x_1 &= \tilde{x} e^{[-D]} < \tilde{x} \\ x_2 &= \tilde{x} e^{[+D]} > \tilde{x} \end{aligned}$$

where  $D = \left\{ \sqrt{\left[ \frac{\log(\sigma_n^2) - \log(\sigma_{n+1}^2)}{\sigma_n^2 - \sigma_{n+1}^2} + \frac{1}{4} \right] \sigma_n^2 \sigma_{n+1}^2} \right\} > 0$ , hence (b) is satisfied.

For (c) and (d), write  $\frac{f(x|n+1)}{f(x|n)}$  explicitly as follows,

$$\frac{f(x|n+1)}{f(x|n)} = \frac{\sigma_n}{\sigma_{n+1}} \exp \left( -\frac{1}{2} \left[ (\log(x) - \log(\tilde{x}))^2 \left( \frac{\sigma_n^2 - \sigma_{n+1}^2}{\sigma_n^2 \sigma_{n+1}^2} \right) - \frac{1}{4} (\sigma_n^2 - \sigma_{n+1}^2) \right] \right) \quad (25)$$

Differentiating with respect to  $x$  we get,

$$\frac{\partial(f(x|n+1)/f(x|n))}{\partial x} = - \left( \frac{\log(x) - \log(\tilde{x})}{x} \right) \left( \frac{\sigma_n^2 - \sigma_{n+1}^2}{\sigma_n^2 \sigma_{n+1}^2} \right) \frac{f(x|n+1)}{f(x|n)}$$

From above we see that  $\frac{\partial(f(x|n+1)/f(x|n))}{\partial x} > 0 \forall x < \tilde{x}$  and hence  $\forall x < x_1$ , since  $x_1 < \tilde{x}$ .

Similarly,  $\frac{\partial(f(x|n+1)/f(x|n))}{\partial x} < 0 \forall x > \tilde{x}$  and hence  $\forall x > x_2$ , since  $x_2 > \tilde{x}$ . Thus conditions (c) and (d) are also satisfied. Therefore  $F(x|n+1)$  represents a *relative strong decrease in risk* compared to  $F(x|n)$ .

**Proof of Proposition 6** Let  $\alpha_i^*(n)$  be the solution of,

$$\int_0^\infty u'(z)(x - \lambda\tilde{x})f(x|n)dx = 0 \quad (26)$$

where  $z = w - \delta^n[(1 - \alpha_i^*(n))x + \alpha_i^*(n)\lambda\tilde{x}]$ . From Lemma 6 we know that  $F(x|n+1)$  represents a *relative strong decrease in risk* compared to  $F(x|n)$ . To show that  $\alpha_i^*(n+1) < \alpha_i^*(n)$

$$\int_0^\infty u'(z)(x - \lambda\tilde{x}) [f(x|n) - f(x|n+1)] dx > 0 \quad (27)$$

Let  $s(x) = f(x|n) - f(x|n+1)$ . Then (27) is  $\int_0^\infty u'(z)(x - \lambda\tilde{x})s(x)dx > 0$ . For convenience, let  $p = \lambda\tilde{x}$

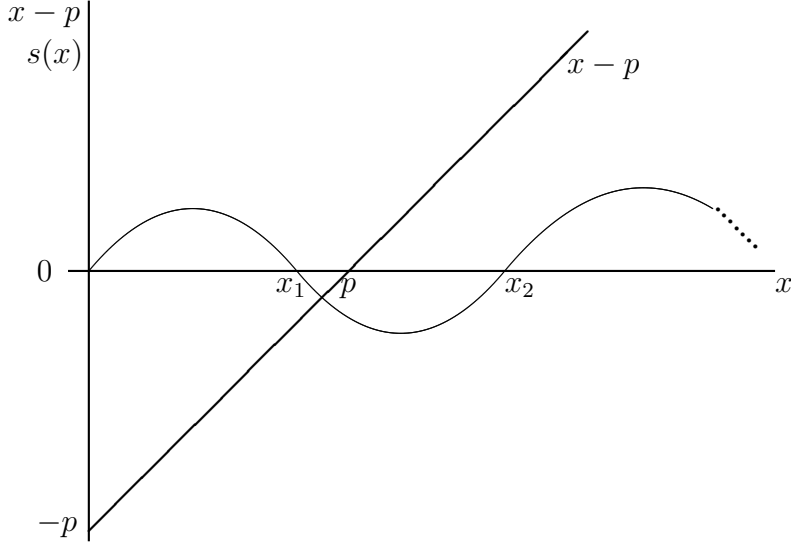


Figure 5: Representation of  $s(x)$  and  $[x - p]$  when  $p \in [x_1, x_2]$

Case A:  $p \in [x_1, x_2]$  (see Fig 5)

$$(x - p)s(x) \begin{cases} < 0 & \forall x \in (0, x_1) \\ > 0 & \forall x \in (x_1, p) \\ < 0 & \forall x \in (p, x_2) \\ > 0 & \forall x \in (x_2, \infty) \end{cases} \quad (28)$$

Therefore (27) may be written as

$$\begin{aligned} & \int_0^\infty u'(z)(x - \lambda \tilde{x}) [f(x|n) - f(x|n+1)] dx = \int_0^{x_1} u'(z)(x - p)s(x) dx \\ & + \int_{x_1}^p u'(z)(x - p)s(x) dx + \int_p^{x_2} u'(z)(x - p)s(x) dx + \int_{x_2}^\infty u'(z)(x - p)s(x) dx \end{aligned}$$

Using (28), note that the first and third terms are negative and second and fourth are positive. By assumption,  $u(\cdot)$  strictly increases and concave; hence  $u'(\cdot)$  increases in  $x$ . Let  $z_i = w - \delta^n[(1 - \alpha^*)x_i + \alpha^*p]$  for  $i = 1, 2$ . Thus,  $u'(z) < u'(z_1) \forall x \in [0, x_1]$  and hence  $-u'(z)|(x - p)s(x)| > -u'(z_1)|(x - p)s(x)| \forall x \in [0, x_1]$ . Similarly, since  $u'(z) > u'(z_1) \forall x \in (x_1, p)$ , hence  $u'(z)(x - p)s(x) < u'(z_1)(x - p)s(x) \forall x \in (x_1, p)$ . Thus  $\int_0^{x_1} u'(z)(x - p)s(x) dx + \int_{x_1}^p u'(z)(x - p)s(x) dx > u'(z_1) \int_0^p (x - p)s(x) dx$ . With the same logic,  $\int_p^{x_2} u'(z)(x - p)s(x) dx + \int_{x_2}^\infty u'(z)(x - p)s(x) dx > u'(z_2) \int_p^\infty (x - p)s(x) dx$ . Putting them together we have

$$\begin{aligned} \int_0^\infty u'(z)(x - p)s(x) dx & > u'(z_1) \int_0^p (x - p)s(x) dx \\ & + u'(z_2) \int_p^\infty (x - p)s(x) dx \end{aligned} \quad (29)$$

with  $u'(z_1) < u'(z_2)$ . Since by assumption, the mean is constant, condition (a), we have  $\int_0^\infty (x - p)s(x) dx = 0$ , or  $\int_0^p (x - p)s(x) dx = -\int_p^\infty (x - p)s(x) dx$ . Using this in (29) we get

$$\int_0^\infty u'(z)(x - p)s(x) dx > [u'(z_1) - u'(z_2)] \left[ \int_0^p (x - p)s(x) dx \right] \quad (30)$$

The term in the first bracket above is negative. To complete the proof we need to show that the term in the second bracket in (30) is negative as well. Integrating by parts

$$\begin{aligned} \int_0^p (x-p)s(x)dx &= \left[ (x-p)[F(x|n) - F(x|n+1)] \right]_0^p \\ &\quad - \int_0^p [F(x|n) - F(x|n+1)] dx \\ &= - \int_0^p [F(x|n) - F(x|n+1)] dx < 0 \end{aligned}$$

which is negative by definition of relatively strong decrease in risk.

Case B:  $p \in (x_2, \infty)$

$$(x-p)s(x) \begin{cases} < 0 & \forall x \in (0, x_1) \\ > 0 & \forall x \in (x_1, x_2) \\ < 0 & \forall x \in (x_2, p) \\ > 0 & \forall x \in (p, \infty) \end{cases} \quad (31)$$

Suppose,

$$\int_{x_2}^{\infty} (x-p)s(x)dx > 0$$

Therefore,  $\int_0^{x_2} (x-p)s(x)dx < 0$ , since by assumption  $\int_0^{\infty} (x-p)s(x)dx = 0$ . This implies  $[u'(z_1) - u'(z_2)] [\int_0^{x_2} (x-p)s(x)dx] > 0$  and therefore  $\int_0^{\infty} u'(z)(x-p)s(x)dx > 0$ , (see (29)). On the other hand, suppose,

$$\int_{x_2}^{\infty} (x-p)s(x)dx < 0$$

Hence,  $\int_0^{x_2} (x-p)s(x)dx > 0$ . By assumption,  $u'(\cdot) > 0$ , therefore,  $\int_0^{x_2} u'(z)(x-p)s(x)dx > 0$ . This implies,

$$\begin{aligned} \int_0^{\infty} u'(z)(x-p)s(x)dx &= \int_0^{x_2} u'(z)(x-p)s(x)dx + \int_{x_2}^{\infty} u'(z)(x-p)s(x)dx \\ &> \int_{x_2}^{\infty} u'(z)(x-p)s(x)dx \end{aligned}$$

From condition (d) of the definition of relatively strong decrease in risk,  $\frac{f(x|n+1)}{f(x|n)}$  decreases in  $x$  and is less than 1  $\forall x \in [x_2, \infty)$ . We may therefore write  $s(x) = f(x|n) - f(x|n+1)$  as  $s(x) = \left[1 - \frac{f(x|n+1)}{f(x|n)}\right] f(x|n)$ , or  $s(x) = [1 - r(x)]f(x|n)$ , where  $r(x) = \frac{f(x|n+1)}{f(x|n)}$ . Thus  $r(x) < 1$  and  $r'(x) < 0$ .

$$\begin{aligned} \int_{x_2}^{\infty} u'(z)(x-p)s(x)dx &= \int_{x_2}^{\infty} u'(z)(x-p)[1 - r(x)]f(x|n)dx \\ &= \int_{x_2}^p u'(z)(x-p)[1 - r(x)]f(x|n)dx \\ &\quad + \int_p^{\infty} u'(z)(x-p)[1 - r(x)]f(x|n)dx \end{aligned}$$

Since  $r(x)$  strictly decreases in  $x \in (x_2, \infty)$ ,  $[1 - r(p)] > [1 - r(x)] \forall x \in (x_2, p)$  and therefore  $-[1 - r(x)]|(x - p)f(x|n)| > -[1 - r(p)]|(x - p)f(x|n)| \forall x \in (x_2, p)$ . Similarly,  $[1 - r(x)](x - p)f(x|n) > [1 - r(p)](x - p)f(x|n) \forall x \in (p, \infty)$ . This implies,

$$\int_{x_2}^{\infty} u'(z)(x - p)[1 - r(x)]f(x|n)dx > [1 - r(p)] \int_{x_2}^{\infty} u'(z)(x - p)f(x|n)dx$$

Also from (26),  $\int_0^{\infty} u'(z)(x - p)f(x|n)dx = 0$ . Therefore  $\int_{x_2}^{\infty} u'(z)(x - p)f(x|n)dx > 0$ . Thus

$$\begin{aligned} \int_0^{\infty} u'(z)(x - p)s(x)dx &> \int_{x_2}^{\infty} u'(z)(x - p)s(x)dx \\ &> [1 - r(p)] \int_{x_2}^{\infty} u'(z)(x - p)f(x|n)dx > 0. \end{aligned}$$

Case C:  $p \in [0, x_1)$  The proof is similar to Case B and condition (c) is used to get the desired result.

## F Data

### F.1 New Information

#### F.1.1 Family of Distribution and their Mean Variance relationship

<i>Distribution</i>	<i>Mean-Variance Relationship</i>
Gaussian	constant variance
Poisson	variance is proportional to mean
Gamma	variance is proportional to square of mean
Inverse Gaussian	variance is proportional to cube of mean

Table 1: Family of Distribution and their Mean Variance relationship

#### F.1.2 Modified Hosmer Lemshaw Test

Variables	Estimates	Standard Errors	p-value
XBETA1	-141.9082	100.6117	0.1586
XBETA2	-25.3545	172.1569	0.8829
XBETA3	-81.1299	137.7078	0.5559
XBETA4	317.0586	200.9011	0.1147
XBETA5	302.1217	328.4401	0.3578
XBETA6	114.098	267.0186	0.6692
XBETA7	-172.782	200.7944	0.3897
XBETA8	-65.1925	296.6735	0.8261
XBETA9	-121.5707	458.8804	0.7911
XBETA10	-301.9304	845.2578	0.721

Table 2: Modified Hosmer Lemshaw Test

### F.1.3 Two Part Model - All Variables

	Estimates	Std Error	p-value
Intercept	5.5199	112.7	0.9609
1 if White	0.3718	0.1055	0.0004
1 if Male	-0.5498	0.0957	<.0001
Age	0.01	0.00381	0.0087
Any Limitation	0.2548	0.1304	0.0507
Paid Leave to see Doc	-0.1874	0.2044	0.3593
1 if Health Poor	0.1744	0.4991	0.7268
1 if Health Exe	-0.3468	0.1058	0.001
Coglim	-0.3693	0.2923	0.2064
Irr work hr	-0.0207	0.1524	0.892
Hours worked	-0.00377	0.00304	0.2147
No prob bending	-4.5416	112.7	0.9679
Paid sick	0.4007	0.2068	0.0526
Elementary Edu	-0.7881	0.2359	0.0008
Some High	-0.3477	0.0961	0.0003
Wage (1000)	0.00662	0.00259	0.0107
1 if Data 98	0.1325	0.1139	0.2444
1 if Data 99	0.0128	0.1025	0.9006

**# of Observations 1550**

Table 3: Part 1: Probability of Positive medical Expenditure (All Variables).

	Estimates	Std Error	p-value
Intercept	7.048	0.292	<.0001
1 if White	0.0627	0.0848	0.46
1 if Male	-0.1602	0.065	0.0137
Age	0.0265	0.0028	<.0001
Any Limitation	0.5398	0.0831	<.0001
Paid Leave to see Doc	-0.1734	0.1251	0.1656
1 if Health Poor	0.5686	0.2297	0.0133
1 if Health Exe	-0.2707	0.0719	0.0002
Coglim	0.573	0.2168	0.0082
Irr work hr	0.0145	0.1166	0.9013
Hours worked	-0.0054	0.0023	0.0166
No prob bending	-0.3739	0.2273	0.1
Paid sick	0.2572	0.1339	0.0548
Elementary Edu	0.5384	0.2304	0.0194
Some High	-0.1929	0.0726	0.0079
Wage (1000)	0.0009	0.0016	0.5599
1 if Data 98	-0.1149	0.0797	0.1493
1 if Data 99	0.0773	0.0771	0.3165

**# of Observations 1358**

Table 4: Part 2: Estimating GLM for Positive Expenditures (All Variables)

### F.1.4 Two Part Model - Only Race, Gender and Age

	Estimates	Std Error	p-value
Intercept	0.7262	0.1834	<.0001
1 if White	0.3803	0.1011	0.0002
1 if Male	-0.6325	0.0912	<.0001
Age	0.0124	0.00344	0.0003
1 if Data 98	0.1817	0.1098	0.0979
1 if Data 99	0.0251	0.098	0.7976

**# of Observations 1550**

Table 5: Part 1: Probability of Positive medical Expenditure (Only Race, Gender and Age).

	Estimates	Std Error	p-value
Intercept	5.3693	0.3757	<.0001
1 if White	0.0166	0.0882	0.8509
1 if Male	-0.1894	0.0654	0.0038
Age	0.0841	0.0188	<.0001
Age Sq	-0.0006	0.0002	0.014
1 if Data 98	-0.1581	0.0835	0.0585
1 if Data 99	-0.0344	0.0792	0.6641

**# of Observations 1358**

Table 6: Part 2: Estimating GLM for Positive Expenditures (Only Race, Gender and Age)

## F.2 Descriptive Statistics

	All		Dropped Coverage in T2		Continued Coverage in T2	
	Mean	Std.Dev	Mean	Std.Dev	Mean	Std.Dev
Age	40.01	12.52	35.36	11.96	40.22	12.51
Elementary Education	2.39%	0.15	2.90%	0.17	2.36%	0.15
Some High School	31.55%	0.46	43.48%	0.50	30.99%	0.46
Some College	65.81%	0.47	53.62%	0.50	66.37%	0.47
Employed T1	91.94%	0.27	73.91%	0.44	92.78%	0.26
Pvt Sector	66.90%	0.47	65.22%	0.48	66.98%	0.47
Fed Gov	3.29%	0.18	2.90%	0.17	3.31%	0.18
State Gov	14.58%	0.35	1.45%	0.12	15.19%	0.36
Part Time in T1	10.06%	0.30	21.74%	0.42	9.52%	0.29
Part Time in T2	10.00%	0.30	27.54%	0.45	9.18%	0.29
Poverty Level T1	422.27	302.46	221.55	160.96	431.62	304.29
Poverty Level T2	452.03	341.34	225.82	155.00	462.56	344.02
Wage in T1	31882.85	23559.72	16185.60	11042.70	32614.19	23734.21
Income T1	35281.12	25316.43	18542.38	13457.31	36060.98	25471.75
HI Offered in T1	73%	0.44	22%	0.42	76%	0.43
HI Offered in T2	70%	0.46	14%	0.35	73%	0.45
Lost Job in T2	3.16%	0.18	11.59%	0.32	2.77%	0.16
Left Pvt Sector	14.77%	0.35	2.90%	0.17	15.33%	0.36
Left Govt Sector	3.68%	0.19	2.90%	0.17	3.71%	0.19
Δ Poverty Level	29.76	287.02	4.27	218.20	30.95	289.83
Δ Wage	3405.33	18233.82	1108.76	9186.71	3512.33	18542.86
Δ Income	3451.82	24379.87	834.77	18371.55	3573.74	24622.09
Offer Withdrawn	3.42%	0.18	5.80%	0.24	3.31%	0.18
Rel Shock	-0.04	1.71	-0.55	1.46	-0.01	1.71
# of Observations	1550		69		1481	

Table 7: Descriptive Statistics

	All		Dropped Coverage in T2		Continued Coverage in T2	
	Mean	Std.Dev	Mean	Std.Dev	Mean	Std.Dev
Age	27.75	4.64	25.14	4.58	27.90	4.60
Elementary Education	1.35%	0.12	2.86%	0.17	1.26%	0.11
Some High School	25.45%	0.44	40.00%	0.50	24.64%	0.43
Some College	73.05%	0.44	57.14%	0.50	73.93%	0.44
Employed	94.16%	0.23	91.43%	0.28	94.31%	0.23
Pvt Sector	74.70%	0.44	82.86%	0.38	74.25%	0.44
Fed Gov	1.65%	0.13	0.00%	0.00	1.74%	0.13
State Gov	13.02%	0.34	2.86%	0.17	13.59%	0.34
Part Time in T1	11.68%	0.32	34.29%	0.48	10.43%	0.31
Part Time in T2	10.93%	0.31	37.14%	0.49	9.48%	0.29
Poverty Level T1	365.06	252.92	210.29	143.64	373.62	254.95
Poverty Level T2	402.99	304.02	231.35	145.23	412.48	307.72
Wage in T1	28067.82	19704.13	15639.04	10484.72	28755.03	19869.89
Income T1	30491.27	21171.02	17566.13	11994.15	31205.93	21343.19
HI Offered in T1	72.60%	0.45	22.86%	0.43	75.36%	0.43
HI Offered in T2	68.41%	0.47	11.43%	0.32	71.56%	0.45
Lost Job in T2	3.14%	0.17	8.57%	0.28	2.84%	0.17
Left Pvt Sector	13.32%	0.34	5.71%	0.24	13.74%	0.34
Left Govt Sector	2.25%	0.15	0.00%	0.00	2.37%	0.15
Δ Poverty Level	37.93	265.81	21.06	201.29	38.86	269.02
Δ Wage	4396.15	17529.06	809.90	11137.23	4594.44	17800.56
Δ Income	4007.69	22534.67	2182.54	16844.73	4108.60	22813.91
Offer Withdrawn	5.39%	0.23	11.43%	0.32	5.06%	0.22
Rel Shock	-0.13	1.52	-0.87	0.98	-0.09	1.54
# of Observations	668		35		633	

Table 8: Descriptive Statistics - Age < 37

	All		Dropped Coverage in T2		Continued Coverage in T2	
	Mean	Std.Dev	Mean	Std.Dev	Mean	Std.Dev
Age	49.29	7.71	45.88	6.96	49.43	7.71
Elementary Education	3.17%	0.18	2.94%	0.17	3.18%	0.18
Some High School	36.17%	0.48	47.06%	0.51	35.73%	0.48
Some College	60.32%	0.49	50.00%	0.51	60.73%	0.49
Employed	90.25%	0.30	55.88%	0.50	91.63%	0.28
Pvt Sector	61.00%	0.49	47.06%	0.51	61.56%	0.49
Fed Gov	4.54%	0.21	5.88%	0.24	4.48%	0.21
State Gov	15.76%	0.36	0.00%	0.00	16.39%	0.37
Part Time in T1	8.84%	0.28	8.82%	0.29	8.84%	0.28
Part Time in T2	9.30%	0.29	17.65%	0.39	8.96%	0.29
Poverty Level T1	465.59	328.72	233.14	178.48	474.91	330.00
Poverty Level T2	489.16	362.88	220.14	166.46	499.95	364.50
Wage in T1	34772.24	25740.74	16748.23	11720.46	35494.91	25889.53
Income T1	38908.80	27513.22	19547.35	14929.96	39685.09	27622.96
HI Offered in T1	73.92%	0.44	20.59%	0.41	76.06%	0.43
HI Offered in T2	71.54%	0.45	17.65%	0.39	73.70%	0.44
Lost Job in T2	3.17%	0.18	14.71%	0.36	2.71%	0.16
Left Pvt Sector	15.87%	0.37	0.00%	0.00	16.51%	0.37
Left Govt Sector	4.76%	0.21	5.88%	0.24	4.72%	0.21
$\Delta$ Poverty Level	23.57	302.09	-13.01	236.14	25.04	304.45
$\Delta$ Wage	2654.92	18724.94	1416.41	6776.22	2704.58	19048.49
$\Delta$ Income	3030.82	25693.97	-552.63	19980.73	3174.49	25895.76
Offer Withdrawn	1.93%	0.14	0.00%	0.00	2.00%	0.14
Rel Shock	0.04	1.83	-0.22	1.78	0.05	1.83
# of Observations	882		34		848	

Table 9: Descriptive Statistics - Age  $\geq$  37

### F.3 New Information against age

Variable	New Information <sup>2</sup>	New Information <sup>2</sup>
Intercept	1.49926 [1.57928]	2.98121 [ 5.53294]
Age	0.03438 [ 0.03683]	-0.04327 [ 0.28030]
Age Square	-	0.00092612 [ 0.00331]
$\beta = 0$ (F-Value)	0.87	0.47
Adj - $R^2$	-0.0001	-0.0007
Obs	1550	1550

Standard Errors in brackets

Table 10: Regression of Square of New Information against age

### F.4 Multivariate Regression

	(1)	Standard Error	(2)	Standard Error
Intercept	3.938 ***	1.5139	4.693 ***	1.6016
Age	-0.221 ***	0.084	-0.250 ***	0.084
Age Sq	0.003 ***	0.0015	0.003 ***	0.001
1 if White	0.672 **	0.3139	0.681 **	0.3144
1 if Male	-0.165	0.2798	-0.150	0.2814
New Information	0.242 *	0.13	-	-
New Info * [1 if Age < 37]	-		0.571 ***	0.2457
New Info * [1 if Age ≥ 37]	-		0.067	0.1119
Δ Income (\$ 1,000's)	-0.005	0.0063	-0.005	0.00624
Lost offer	0.716	0.5713	0.737	0.5803
Left Private Sector	1.302 *	0.7705	1.325 *	0.7778
Lost Job	-0.609	0.4766	-0.650	0.4791
Now Part Time	-0.697 **	0.3303	-0.705 **	0.3336
Δ Ind wise Divv rate	-6.331 *	3.7538	-6.643 *	3.7993
Tax Rate * [1 if off]	0.177 ***	0.0278	0.177 ***	0.0279
Tax Rate * [1 if not off]	0.057 ***	0.0216	0.057 ***	0.0216
-2 Log L	410.83		406.56	
HL Lack of fit (p-value)	0.261		0.311	
# of Observations	1550		1550	

\*\*\*: Statistically significant at 1%; \*\*: at 5%; \*: at 10%;

Table 11: Multivariate Regression - Study Sample

	Age < 37	Standard Error	Age ≥ 37	Standard Error
Intercept	1.0558	1.3127	-5.6976 ***	1.7451
Age	0.0258	0.0465	0.1165 ***	0.0285
1 if White	0.8413 *	0.4687	0.4517	0.4542
1 if Male	-0.4025	0.4111	0.00305	0.4121
New Information	0.7587 ***	0.2674	0.0348	0.0995
Δ Income (\$ 1,000's)	-0.00427	0.00999	-0.00541	0.0078
Lost offer	0.4775	0.6457	13.6721	617.7
Left Private Sector	1.0204	0.8488	11.2039	201.9
Lost Job	-0.1625	0.7394	-1.1871 *	0.6572
Now Part Time	-1.1405 **	0.4554	0.045	0.6055
Δ Ind wise Div rate	0.6896	4.6561	-16.5461 ***	6.5674
Tax Rate * [1 if off]	0.1383 ***	0.0434	0.191 ***	0.0377
Tax Rate * [1 if not off]	0.0116	0.0353	0.0836 ***	0.0291
-2 Log L	194.34		191.42	
HL Lack of fit (p-value)	0.444		0.646	
# of Observations	668		882	

\*\*\* : Statistically significant at 1%; \*\* : at 5%; \* : at 10%;

Table 12: Multivariate Regression - Split Sample

## F.5 Magnitude of the effect of New Information

Study Sample						
	Percentile		Logit Coeff	Change in Probability (% points)	95% C I (% points)	
	5 <sup>th</sup>	95 <sup>th</sup>			Lower	Upper
All Individuals	-2.188	6.964	0.242*	1.57	-0.13	3.01
Age < 37	-2.461	6.623	0.571***	5.33	0.39	13.60
Age ≥ 37	-2.120	7.282	0.067	0.37	-1.62	1.20

Split Data						
	Percentile		Logit Coeff.	Change in Probability (% points)	95% C I (% points)	
	5 <sup>th</sup>	95 <sup>th</sup>			Lower	Upper
Age < 37	-2.461	6.623	0.7587***	5.80	1.47	18.29
Age ≥ 37	-2.120	7.282	0.0348	0.04	-0.31	0.18

Author's calculations using regressions in Table 11 and 12

\*\*\*: Statistically significant at 1%; \*\*: at 5%; \*: at 10%;

Table 13: Magnitude of the effect of New Information

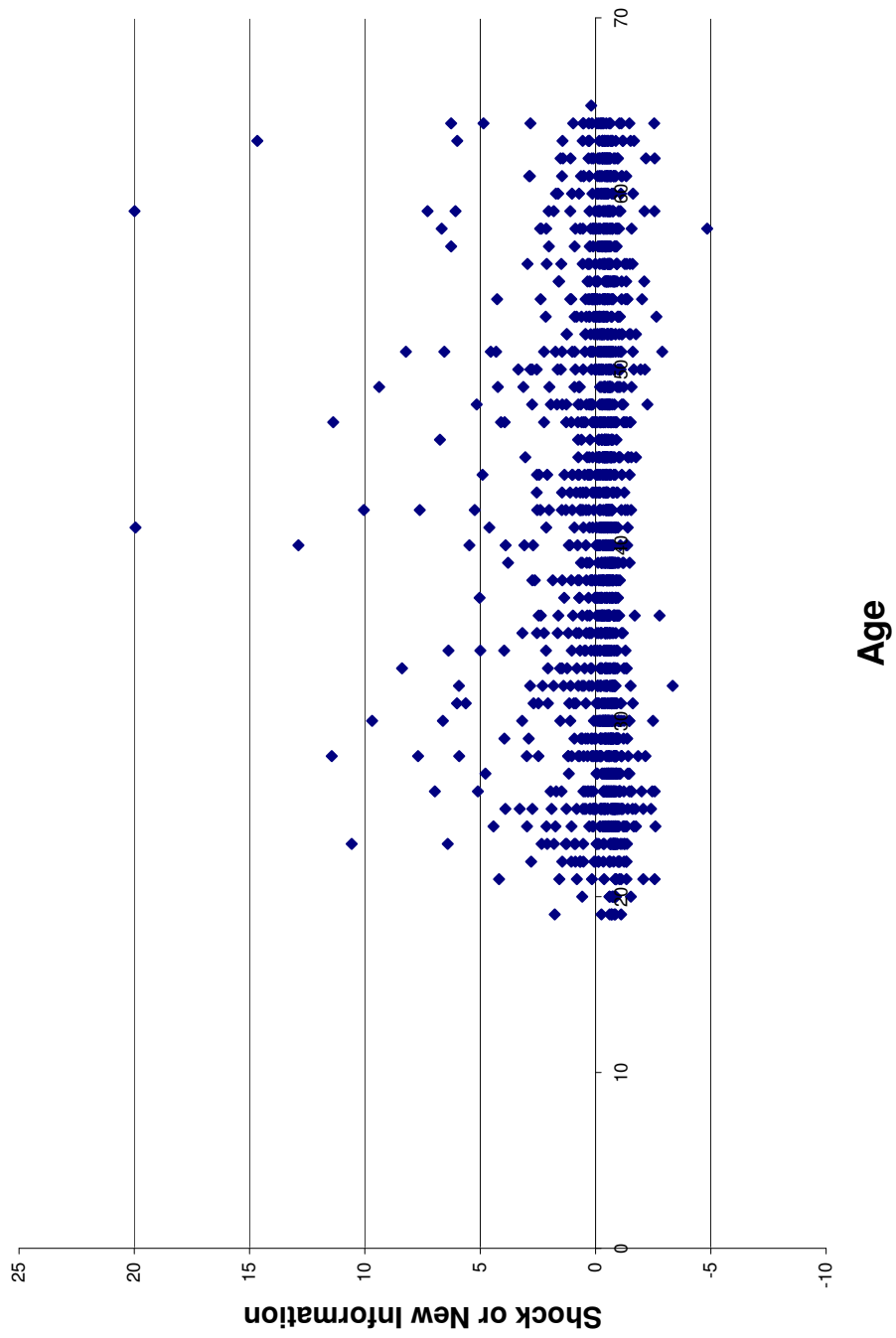


Figure 6: New Information against Age