Health Insurance, Access to Care, and Risk-Aversion: Separating Selection and Incentive Effects

Job Market Paper

Dzmitry Asinski†
Iowa State University
November 2005

Abstract

This paper provides new evidence on the existence and nature of imperfect information in health insurance markets that has important implications for assessing market-based health care reform proposals. Empirical research has traditionally concentrated on testing for positive correlation between coverage and risk predicted by standard moral hazard and adverse selection models. Recent theoretical advances demonstrate that the lack of such correlation does not signal the absence of informational asymmetries and is consistent with more complex models with unobserved risk-aversion. I extend the empirical literature by testing for selection on more than one type of latent information - risk type and risk-aversion type. In order to separate an incentive (moral hazard) effect from effects of selection on multiple types of unobservables, I specify a hybrid endogenous treatment model with explicit modeling of indicators of latent attitudes toward risk-taking behavior. The model is estimated using Markov Chain Monte Carlo (MCMC) methods. I use data from the 2000 Medical Expenditure Panel Survey (MEPS), which contains a set of attitudinal variables necessary to estimate the model. Although the overall selection effect appears to be insignificant, the results indicate that individuals do, in fact, possess private information which increases their propensity to be insured and to utilize health care. This suggests that lack of conditional correlation between insurance coverage and health care utilization results from informational asymmetries of multiple types inducing selection in opposite directions. In addition, I find strong evidence of moral hazard in outpatient and office-based health care utilization but not in inpatient or emergency room utilization.

*I wish to express my sincere gratitude to Brent Kreider, Justin Tobias, and Steven Hill for their invaluable help and advice. All remaining errors are my own.
†dasinski@iastate.edu
1 Introduction

The problem of asymmetric information in insurance markets is of considerable interest from both theoretical and policy perspectives. The informational problems related to health insurance (or lack thereof) have been on the leading edge of public interest because of concerns over access to care by the uninsured and rapidly rising medical care costs. It is well-known that adverse selection can potentially completely destroy the market or leave a very large portion of the population uninsured (Rothschild and Stiglitz (1976)). The other major type of market failure, moral hazard, leads to inefficiently high consumption and is often credited as one of the engines of the fast growth of medical care costs in recent decades. In addition, the extent and nature of asymmetric information in health insurance markets has important implications for evaluating various market-based health care reform proposals.

The main goal of this paper is to test for existence of multiple types of private information in health insurance markets. I show that asymmetric information does exist and that it is multidimensional. Moreover, failure to account for the multidimensional nature of private information may lead to the erroneous conclusion that there are no informational asymmetries.

The empirical literature addressing informational asymmetries in insurance markets has traditionally been focused on testing for positive correlation, conditional on observables, between coverage and risk predicted by standard adverse selection and moral hazard models. Adverse selection, on the one hand, predicts that (unobservably) riskier individuals should be more likely to self-select into better coverage. Moral hazard, on the other hand, predicts that individuals with better coverage take fewer precautions because of the reduced penalties for risky behavior thereby increasing their risk level. Both types of informational asymmetries can cause positive correlation between coverage and risk in the data, yet the direction of causality is different. The empirical evidence on the existence of such correlation in health care and health insurance markets, especially coming from adverse selection, is not
overwhelming.\(^1\) A possible explanation for the absence of positive correlation between risk and coverage in some studies is provided by recent theoretical advances in contract theory. In particular, more complex asymmetric information models that include unobserved risk-aversion do not generally produce the positive correlation prediction generated by simpler theoretical models.\(^2\)

The innovation of this paper is to consider two distinct types of selection in the context of health insurance: selection based on the latent *risk type* and selection based on the latent *risk-aversion type*. The distinction between these two types of private information is crucial because the selection on risk-aversion type can generate negative correlation between coverage and risk. Relatively more risk-averse individuals may have higher incentives to select better coverage because they have higher demand for consumption smoothing provided by insurance. At the same time, more risk-averse individuals may choose to be more cautious and to be more actively engaged in prevention activities, which, in turn, reduces risk. In other words, individuals with higher risk-aversion have higher disutility from both negative health and negative income shocks, which induces them to insure in all available ways – buying (costly) insurance or engaging in (costly) prevention activities. Models that do not consider such type of selection in addition to the traditional selection on the risk type may erroneously imply that the lack of correlation between coverage and risk signals the lack of asymmetric information (this possibility was pointed out by Finkelstein and McGarry (2003) and Gardiol et al. (2003)). Testing for selection on multiple types of private information has implications for assessing health care reform proposals that emphasize consumer choice.

\(^{1}\) The sizeable empirical literature attempting to test for asymmetric information and to separate effects of moral hazard and adverse selection produced diverse results with two important messages emerging from it – the choice of control variables and the population under consideration matter. For example, Cardon and Hendel (2001), Coulson and Stuart (1995), Dionne et al. (1998), Dionne et al. (2003), Dowd et al. (1991) find no evidence of adverse selection in different insurance markets after properly accounting for risk classification. Bradley (2002), Deb and Trivedi (2004), Ettner (1997), Holly et al. (1998), Hurd and McGarry (1997) do find such evidence.

\(^{2}\) Indeed, Chiappori et al. (2002) demonstrate that positive correlation between risk and coverage persists when the basic adverse selection model of Rothschild and Stiglitz (1976) is extended to more general environments. In particular, a model with both adverse selection and moral hazard has this property. There are, however, two important dimensions in which the positive correlation property generally does not extend: non-competitive environments and unobserved risk-aversion.
and competition among insurers. Instead of offsetting each other’s effects, different dimensions of private information together with increased competition among insurers may lead to further segmentation of the population. Yet, despite the availability of theoretical literature incorporating both types of selection, applied work has generally been hindered by the lack of data allowing researchers to separate different types of private information. Finkelstein and McGarry’s study of long-term care insurance markets is the only study I am aware of that tests and finds evidence of different types of asymmetric information. They use an individual’s subjective evaluation of the probability of using a nursing home within the next five years to show that, controlling for observables, it is a significant predictor of future nursing home use. In addition, the preventive care measures are found to be negatively correlated with subsequent nursing home use and positively correlated with the propensity to have long-term care insurance. The authors argue that these two types of private information cancel each other to produce an insignificant estimate of the overall selection effect.

I follow Gardiol et al. (2003) in calling the two main effects of interest the selection effect and the incentive effect. The selection effect occurs if individuals select into different insurance states (plans) based on private information. The presence of the selection effect is a necessary but not sufficient condition for adverse selection in the market. The incentive effect is comprised of ex-ante moral hazard and ex-post moral hazard. The former effect is present if the probability of experiencing an adverse health event (sickness) is increased because of the reduced benefits from engaging in preventive activities. The latter effect appears because the insured individuals face a greatly reduced price of health care.

---

3See, for example, Araujo and Moreira (2001), de Meza and Webb (2001).
4It is also interesting to note that Meer and Rosen (2004) estimate the demand for health care using instrumental variables (IV) techniques to control for potential endogeneity of health insurance status. Their main contribution to the literature is to introduce a new instrument – self-employment status. They find that after accounting for endogeneity, health insurance status is still an important predictor of health care utilization. What is more interesting is that compared to un-instrumented estimates, the IV estimates are larger. One possible explanation is that self-employed individuals are more likely to be more risk-loving than the general populace. It suggests that this particular instrument may control for some part of risk-aversion.
5The adverse selection is an equilibrium concept, and I only model individual behavior in this paper. Nevertheless, the term is often used in the applied literature to denote selection.
6It is also worth noting that from a theoretical standpoint, the classical moral hazard model requires some unobservable action by an agent which directly influences the outcome of interest. For example, an agent’s
In order to test for the presence of multiple sources of private information, I extend the standard two-equation endogenous treatment model along the lines of the general discrete choice model of Walker and Ben-Akiva (2002) by explicitly modeling indicators of latent risk-aversion. The resulting hybrid model includes the endogenous treatment system and a latent variable measurement equation. The endogenous treatment model has been used extensively in the literature to separate the effects of moral hazard and adverse selection system, and it consists of an outcome equation (health care utilization, which serves as a proxy for risk) and a treatment (insurance coverage) equation. I estimate the resulting system of equations jointly using Markov Chain Monte Carlo (MCMC) methods. The choice of methodology is strongly motivated by the computational advantages of Bayesian methods over classical procedures in related models (see Munkin and Trivedi (2003)).

In this paper I use a unique set of variables available in the Medical Expenditure Panel Survey (MEPS) that measure individual responses to a series of questions about risk-taking behavior in general and attitudes toward health insurance in particular.\footnote{The MEPS variables used in this paper first appeared in 2000. The precursor to the MEPS – the 1987 National Medical Expenditure Survey (NMES) – also contained these variables.} I assume that these self-reported attitudes measure one dimension of private information available to the individuals – latent risk-aversion. One of the goals of this paper is to provide additional insights into what responses to these attitudinal questions actually convey in the context of medical care utilization and health insurance choice. The importance of such insights for future research is underlined by the resurgence of interest in the role played by individual preferences in making decisions related to health insurance.\footnote{For example, Monheit and Vistnes (2004) show that self-reported attitudes toward health insurance play an important role in job search.}

I use binary indicators for both medical care utilization and health insurance status. The most important advantage of the binary utilization and insurance variables is that they lend themselves naturally to the analysis of disparities in access to care by the insured and engaged in preventive activities is unobserved, and these preventive activities can directly influence health status and health care expenditures.
uninsured.\textsuperscript{9} A multinomial variable reflecting the choice of a contract from an available menu is an often used alternative. Unfortunately, the MEPS dataset I use does not contain reliable information on the actual menus that people face.\textsuperscript{10} Furthermore, studies analyzing the choice of a contract from an available menu typically restrict their samples to employed individuals with offers of multiple contracts, which affects the generality of their results.

The estimates of the model indicate that there is no selection on unobservables in my sample. At the same time, the latent information measured by the attitudinal questions in my data significantly affects both health insurance and utilization decisions. The lack of an estimated overall selection effect combined with the presence of private information affecting both decisions suggests that there must be other distinct types of private information inducing selection in the opposite direction. Therefore, the lack of an overall selection effect does not signal the absence of informational asymmetries in my sample. Instead, it signals the presence of multiple types of private information that act in opposite directions to cancel out. In addition, I find a strong incentive effect (moral hazard) of being insured on the access to outpatient and office-based care. Consistent with the previous literature, I do not find any significant incentive effect on access to emergency room or inpatient utilization.

The paper proceeds as follows. Section 2 outlines the econometric model and estimation details, Section 3 discusses the data, Section 4 provides results, and Section 5 concludes.

2 The Model

2.1 General Outline

Individuals make two choices in the model: whether to be insured and whether to utilize medical care. The two decisions can be thought of as corresponding to two distinct stages of

\textsuperscript{9}In addition, there is no consensus in the literature as to which continuous measure of health care utilization better reflects the demand for health care.

\textsuperscript{10}In fact, there is a restricted-use file in the MEPS containing information on the actual menu of contracts facing each employed individual. Unfortunately, a very high non-response rate renders these data unusable for the purposes of this paper.
decision making. In the first stage, individuals observe a private signal about their expected future health care needs and choose an insurance option according to this expectation. In the second stage, health care needs are realized and, conditional on chosen insurance status, the medical care decision is made. The important feature of the decision making process is the fact that the propensity to consume health care, which would obviously influence both the health insurance decision and the medical care consumption decision, is unobserved.

Both the selection and the incentive effects are expected to generate positive correlation between the presence of health insurance and health care utilization conditional on observables. The direction of causality is very different, however, for the two effects. The selection effect induces people who expect to be relatively heavy users of medical care to select into the insured status. The incentive effect induces insured individuals to consume more health care, conditional on their health status, because they face a reduced price.

With $i$ indexing individuals, I denote by $U_i^*$ and $HI_i^*$ the latent indexes (utilities) governing the choice of health care utilization and the choice of health insurance status, respectively. The binary observed outcome variables $U_i$ and $HI_i$ are obtained from the latent indexes associated with each choice in the following manner:

$$U_i = 1[U_i^* > 0]$$
$$HI_i = 1[HI_i^* > 0],$$

(2.1)

where $1[.]$ is an indicator function.

I then introduce two latent factors reflecting risk type and risk-aversion type of individual $i$ and denote them by $R_i^*$ and $RL_i^*$, respectively. I assume that higher $RL_i^*$ reflects lower risk-aversion (higher risk-loving).\(^{11}\) Key to this model, I also observe an indicator of the

\(^{11}\)This reflects the way risk-aversion is measured in my data.
latent risk-loving denoted by $RL_i$, that is obtained from the latent index $RL_i^*$ as

$$RL_i = 1[RL_i^* > 0].$$

(2.2)

The latent utilities governing the two decisions in the model are assumed to be equal to:

$$U_{i}^* = x_i \theta_1 + \gamma HI_i + \alpha_1 R_i^* + \zeta_1 RL_i^* + \epsilon_{1i},$$

$$HI_i^* = x_i \theta_2 + z_i \nu + \alpha_2 R_i^* + \zeta_2 RL_i^* + \epsilon_{2i},$$

(2.3)

where $x_i$ is a vector of exogenous variables; $z_i$ is a vector of instruments; $\theta_1$ and $\theta_2$ are vectors of parameters, $\alpha_1$, $\alpha_2$, $\zeta_1$, $\zeta_2$, $\nu$ and $\gamma$ are scalar parameters; the error terms $\epsilon_{1i}$, $\epsilon_{2i}$ are assumed to be independently and identically distributed $\sim N(0, \sigma_\epsilon)$; the latent risk type $R_i^*$ is assumed to have standard normal distribution. Finally, I assume that the latent risk aversion $RL_i^*$ is determined as follows

$$RL_i^* = x_i \delta + \eta_i,$$

(2.4)

where $x_i$ is the vector of exogenous variables; $\delta$ is a scalar parameter; and the error term $\eta_i$ is assumed to have a standard normal distribution. Note that although the errors $\epsilon_{1i}$, $\epsilon_{2i}$ are uncorrelated, the problem of endogeneity of insurance choice still remains because of the presence of latent variables $R_i^*$ and $RL_i^*$. Substituting the expression for the latent risk

---

12Note that this structure explicitly imposes the restriction that the two types of private information are independent. Given that I only have data on risk-aversion, I can only identify the additive effects. To model interactions between different types of private information one would need data on all of them.

13Identification of the parameter $\gamma$ in a model with only two binary equations cannot be achieved through non-linearity of the treatment (health insurance) equation alone and exclusion restrictions would be required (see Maddala (1986)). Since my model includes the third equation, the identification by non-linear form is possible. Nevertheless, I include a vector of instruments $z_i$ for more robust identification. I discuss identification of the model in more detail later. 

---
aversion $RL^*_i$ into (2.3) I obtain the three equation system:

\[
\begin{align*}
U^*_i &= x_i \beta_1 + \gamma HI_i + \alpha_1 R^*_i + \zeta_1 \eta_i + \epsilon_{1i} \\
HI^*_i &= x_i \beta_2 + z_i \nu + \alpha_2 R^*_i + \zeta_2 \eta_i + \epsilon_{2i} \\
RL^*_i &= x_i \delta + \eta_i,
\end{align*}
\]

(2.5)

where $\beta_1 = \theta_1 + \zeta_1 \delta$, $\beta_2 = \theta_2 + \zeta_2 \delta$ are the identified reduced-form coefficients on the observables in the first two equations of the model.

The vector of error terms of the system (2.5) is given by $\epsilon_i = (\epsilon_{1i}, \epsilon_{2i}, \epsilon_{3i})'$, where $\epsilon_{1i} = \alpha_1 R^*_i + \zeta_1 \eta_i + \epsilon_{1i}$, $\epsilon_{2i} = \alpha_2 R^*_i + \zeta_2 \eta_i + \epsilon_{2i}$, $\epsilon_{3i} = \eta_i$. The error covariance matrix $\Sigma$ is equal to

\[
\Sigma = \begin{pmatrix}
\alpha_1^2 + \zeta_1^2 + \sigma_\epsilon^2 & \alpha_1 \alpha_2 + \zeta_1 \zeta_2 & \zeta_1 \\
\alpha_1 \alpha_2 + \zeta_1 \zeta_2 & \alpha_2^2 + \zeta_2^2 + \sigma_\epsilon^2 & \zeta_2 \\
\zeta_1 & \zeta_2 & 1
\end{pmatrix} = \begin{pmatrix}
1 & \rho_{12} & \rho_{13} \\
\rho_{12} & 1 & \rho_{23} \\
\rho_{13} & \rho_{23} & 1
\end{pmatrix},
\]

(2.6)

where I normalize the scale of all three equations to 1 by setting $\alpha_1^2 + \zeta_1^2 + \sigma_\epsilon^2 = 1$ and $\alpha_2^2 + \zeta_2^2 + \sigma_\epsilon^2 = 1$. The signs of covariance terms of the error covariance matrix are very important for this analysis, and they are unaffected by these scale normalizations.

Without loss of generality, I assume that higher values of $R^*_i$ reflect higher risk as measured by the higher expected probability of utilizing any medical care. I therefore expect $\alpha_1 > 0$. The standard selection on unobservables hypothesis states that $\alpha_2 > 0$ because individuals with greater propensity to utilize health care are more likely to select into the insured state. It then follows that $\alpha_1 \alpha_2 > 0$. In the absence of any other types of private information; i.e., $\eta_i = 0$, I expect to observe positive correlation between health care utilization and insurance status conditional on observables; i.e., $\rho_{12} = \alpha_1 \alpha_2 > 0$.

If latent risk-aversion is an important factor in decision making, the two-equation endogenous treatment model can lead to incorrect conclusions about the existence of imperfect information. Without loss of generality, I assume that higher values of $\eta_i$ reflect lower
risk-aversion (higher $RL_i^*$). Then the theory predicts that $\zeta_1 > 0$ because more risk-loving individuals (higher $\eta_i$) will be less cautious and will be less actively engaged in prevention activities, thereby worsening health status and increasing the probability of using health care.\footnote{It is also conceivable that more risk averse individuals will be more likely to use health care because much of what we call prevention consists of frequent visits to a doctor to perform various check-ups. In this paper I consider various measures of health care utilization to try to distinguish between different possible effects of risk aversion on health care utilization.}

At the same time, more risk-loving individuals are less likely to be insured, i.e. $\zeta_2 < 0$. It then follows that $\zeta_1\zeta_2 < 0$, which in turn implies that $\text{sign}(\rho_{12}) = \text{sign}(\alpha_1\alpha_2 + \zeta_1\zeta_2)$ is ambiguous. More specifically, the estimate of $\rho_{12} = \alpha_1\alpha_2 + \zeta_1\zeta_2$ can be very close to zero, implying that there is no private information and no selection effects. However, a conclusion based on $\text{sign}(\rho_{12})$ alone could clearly be erroneous as there is, in fact, private information ($\alpha_1\alpha_2 > 0$ and $\zeta_1\zeta_2 < 0$) of two different kinds that induce selection in the opposite directions.

The inclusion of an observed index of latent risk-aversion allows me to distinguish between two very different scenarios: (1) no private information and (2) private information of multiple types.\footnote{As a matter of practical implementation, it is not clear that the attitudinal questions used in this paper actually do measure risk aversion type and not, say, risk type or some combination of risk type and risk aversion type. Still, the important message of my argument is that the attitudinal questions do measure some facet of a potentially multifaceted private information set. By comparing correlations it is possible to reach certain conclusions about the presence of imperfect information of multiple types and also gain insights into what these questions are asking.}

This can be achieved by combining the information on the estimate of $\rho_{12}$ with information on the estimates of $\rho_{13}$ and $\rho_{23}$. Given the theoretical prediction discussed above, I should obtain $\text{sign}(\rho_{13}) = \text{sign}(\zeta_1) = +$ and $\text{sign}(\rho_{23}) = \text{sign}(\zeta_2) = -$. If the estimate of $\rho_{12}$ is insignificant but the estimates of both $\rho_{13}$ and $\rho_{23}$ are significantly different from zero, I will conclude that there is private information instead of reporting no informational asymmetries.

The next subsection will outline estimation of the identified system of reduced-form
The parameter $\gamma$ measures the incentive effect of having health insurance on the probability of utilizing any health care. The correlations between the error terms of the three equations are designed to capture the unobserved private information about the latent risk type and risk-aversion type. The theoretical predictions of the signs of the error correlation matrix elements are as follows:

\[
\Sigma = \begin{pmatrix} 1 & +? & + \\ +? & 1 & - \\ + & - & 1 \end{pmatrix}.
\]  

(2.8)

2.2 Estimation Details

2.2.1 Derivation of the Posterior Distribution

I use Bayesian methods to fit the model defined by 2.1, 2.2, and 2.7. The joint posterior distribution of the parameters of the model has to be simulated because it does not have a convenient analytical form. The parameter set is split into blocks and a variant of Gibbs sampling algorithm (see, for example, Tierney (1994)) is used to iteratively draw values from posterior distribution of each block of parameters conditional on other parameters of the model. The posterior output of this Markov Chain is used to make inferences about the parameters of the interest. Because the conditional posterior distribution of error correlations is not of a standard form, a tailored Metropolis-Hastings algorithm is used to sample from

An alternative approach would be to model the choice of insurance status and the choice of health care utilization with added explanatory variable $RL_i$. The coefficients on the latter would tell how observed indicators of latent factors affect the insurance status and the health care utilization decisions. However, such an approach ignores the potential endogeneity of self-reported indicators to the actual decisions made. It also ignores the potentially continuous nature of the latent risk-aversion index.
it (Metropolis et al. (1953), Hastings (1970)). Lastly, I follow Albert and Chib (1993) by augmenting the parameter set with the latent data. The three equations for each individual are stacked in the following manner

\[
y^*_i = \begin{pmatrix} U^*_i \\ HI^*_i \\ RL^*_i \end{pmatrix}_{3 \times 1}, y_i = \begin{pmatrix} U_i \\ HI_i \\ RL_i \end{pmatrix}_{3 \times 1}, \varepsilon_i = \begin{pmatrix} \varepsilon_{1i} \\ \varepsilon_{2i} \\ \varepsilon_{3i} \end{pmatrix}_{3 \times 1},
\]

\[
X_i = \begin{pmatrix} x_i & HI_i & 0 & 0 & 0 \\ 0 & 0 & x_i & z_i & 0 \\ 0 & 0 & 0 & 0 & x_i \end{pmatrix}_{3 \times k}, \text{and } \beta = \begin{pmatrix} \beta_1 \\ \gamma \\ \beta_2 \\ \nu \\ \delta \end{pmatrix}_{k \times 1},
\]

where \( k \) is the total number of explanatory variables in all three equations. The system can then be expressed as

\[
y^*_i = X_i \beta + \varepsilon_i
\]

\[
\varepsilon_i \sim N(0, \Sigma).
\]

The observations are then stacked together as

\[
y^* = \begin{pmatrix} y^*_1 \\ y^*_2 \\ \vdots \\ y^*_n \end{pmatrix}_{3n \times 1}, y = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}_{3n \times 1}, X = \begin{pmatrix} X_1 \\ X_2 \\ \vdots \\ X_n \end{pmatrix}_{3n \times k}, \varepsilon = \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{pmatrix}_{3n \times 1}.
\]
to produce
\[ y^* = X\beta + \varepsilon \]
\[ \sim N(X\beta, I_n \otimes \Sigma). \] (2.9)

For computational simplicity, I follow Albert and Chib (1993) and treat the latent data \( y^* \) as additional parameters of the model. The augmented posterior \( p(y^*, \beta, \Sigma| y) \), which also contains the latent data, is proportional to

\[
p(y^*, \beta, \Sigma| y) \propto p(y|y^*, \beta, \Sigma) p(y^*|\beta, \Sigma) p(\beta, \Sigma)
\]
\[
\propto p(\beta, \Sigma) \prod_{i=1}^{n} p(y_i|y_i^*) p(y_i^*|\beta, \Sigma)
\]
\[
\propto p(\beta, \Sigma) \prod_{i=1}^{n} 1[U_i = 1[U_i^* > 0]] 1[HI_i = 1[HI_i^* > 0]] \times
\]
\[
1[RL_i = 1[RL_i^* > 0]] p(y_i^*|\beta, \Sigma), \] (2.10)

where the second line follows from the assumed independence across individuals. Conditional on the parameters of the model, the augmented likelihood can be expressed as\textsuperscript{17}

\[
p(y^*|\beta, \Sigma) = (2\pi)^{-\frac{n}{2}|I_n \otimes \Sigma|^{-\frac{1}{2}} \exp \left( -\frac{1}{2} (y^* - X\beta)'(I_n \otimes \Sigma)^{-1}(y^* - X\beta) \right)
\]
\[
\propto |\Sigma|^{-\frac{n}{2}} \exp \left( -\frac{1}{2} \sum_{i=1}^{n} (y_i^* - X_i\beta)'\Sigma^{-1}(y_i^* - X_i\beta) \right). \] (2.11)

I place the following independent Normal prior distribution on \( \beta \textsuperscript{18} \)

\[
\beta \sim N(\mu_{\beta 0}, V_{\beta 0}), \] (2.12)

where \( \mu_{\beta 0} \) and \( V_{\beta 0} \) denote the prior mean and covariance matrix of \( \beta \). Because identification

\textsuperscript{17}Refer to Appendix A for the details of derivations.

\textsuperscript{18}I use proper priors for all parameters. The exact hyperparameters used in estimation are given in the next subsection.
requires the use of the normalized covariance matrix $\Sigma$, priors are placed directly on the identified correlation parameters $\rho_{12}$, $\rho_{13}$, and $\rho_{23}$, and are chosen to ensure that $\Sigma$ is positive definite with probability 1. Specifically,

\[
\begin{align*}
\rho_{12} &\sim UNIF(-1, 1) \\
\rho_{13} &\sim UNIF(-1, 1) \\
\rho_{23} | \rho_{12}, \rho_{13} &\sim UNIF\left(\rho_{12}\rho_{13} - \sqrt{(1 - \rho_{12}^2)(1 - \rho_{13}^2)}, \rho_{12}\rho_{13} + \sqrt{(1 - \rho_{12}^2)(1 - \rho_{13}^2)}\right).
\end{align*}
\]

**2.2.2 Posterior Simulation**

The conditional posteriors of both $\beta$ and $\Sigma$ are proportional to the product of likelihood and the respective prior distribution. It turns out that the conditional posterior for $\beta$ is also Normal:\(^{19}\)

\[
p(\beta | y^*, \Sigma) \sim N(\mu_{\beta1}, V_{\beta1})
\]

\[
V_{\beta1} = \left(\sum_{i=1}^{n} X_i^\prime \Sigma^{-1} X_i + V_{\beta0}^{-1}\right)^{-1}
\]

\[
\mu_{\beta1} = V_{\beta1} \left(\sum_{i=1}^{n} X_i^\prime \Sigma^{-1} y_i^* + V_{\beta0}^{-1} \mu_{\beta0}\right).
\]

The conditional posterior distribution of $(\rho_{12}, \rho_{13}, \rho_{23})$ is given by

\[
p(\rho_{12}, \rho_{13}, \rho_{23} | y^*, \beta) \propto |\Sigma|^{-\frac{n}{2}} \exp\left(-\frac{1}{2} \sum_{i=1}^{n} (y_i^* - X_i \beta)^\prime \Sigma^{-1} (y_i^* - X_i \beta)\right) \times
\]

\[
\frac{1}{\sqrt{(1 - \rho_{12}^2)(1 - \rho_{13}^2)}}.
\]

This distribution does not take any convenient standard form. To sample from (2.15), I employ a Metropolis-Hastings (MH) step within Gibbs algorithm. I choose the proposal density to be Normal centered at the sample correlation of errors, which are “known” given

---

\(^{19}\)The reader is referred to Appendix A for details.
\( \beta \) and \( y^* \). The variance of the proposal density is chosen so that the acceptance rate is approximately 35\%.\(^{20}\) Finally, the data augmentation step draws the values of latent variables \( U_i^* \), \( HI_i^* \), and \( RL_i^* \) conditional on the observed data \( y_i \) and parameters of the model \( \beta, \Sigma \). The distribution of latent data is truncated normal:

\[
y_i^* \mid \beta, \Sigma, y \sim TN_{C(y_i)}(X_i\beta, \Sigma), \tag{2.16}
\]

where \( TN_R(\mu, \Omega) \) denotes the multivariate normal distribution with mean \( \mu \) and covariance matrix \( \Omega \) truncated to the region \( R \). For each individual \( i \) the normal density is truncated to the region \( C(y_i) = C(U_i) \times C(HI_i) \times C(RL_i) \) with

\[
C(U_i) = \begin{cases} [0, \infty) & \text{if } U_i^* = 1 \\ (-\infty, 0) & \text{if } U_i^* = 0 \end{cases}
\]

\[
C(HI_i) = \begin{cases} [0, \infty) & \text{if } HI_i^* = 1 \\ (-\infty, 0) & \text{if } HI_i^* = 0 \end{cases}
\]

\[
C(RL_i) = \begin{cases} [0, \infty) & \text{if } RL_i^* = 1 \\ (-\infty, 0) & \text{if } RL_i^* = 0. \end{cases}
\]

I follow Geweke (1991) to sample from this truncated multivariate normal distribution. I sample each latent index from a univariate truncated normal density conditional on the current values of all other latent indices using the inverse distribution function method.

The details of the algorithm are as follows:

**Step 0:** Set \( (y_i^*)^0 = [(U_i^*)^0 \ (HI_i^*)^0 \ (RL_i^*)^0]' = [U_i \ HI_i \ RL_i]' \) and \( \Sigma^0 = I_3 \), where \( I_j \) is the identity matrix of dimension \( j \).

**Step 1:** Draw \( \beta^1 \) from the distribution given in (2.14) conditional on \( (y_i^*)^0 \) and \( \Sigma^0 \) (I use

\(^{20}\)The general rule of thumb for acceptance rate is about 44\% when drawing one parameter and is about 23\% when drawing a large number of parameters (see Koop (2003) and Train (2003)).
the following hyperparameters for the prior distributions: \( \mu_\beta = 0, V_\beta = 1000 * I_k \).

**Step 2:** Draw the elements of the covariance matrix \( \Sigma \) using the Metropolis-Hastings algorithm:

a. Let \( \Sigma^0 = \Sigma^0 \);

b. Compute the errors \( \varepsilon_i \) given the realized \( \beta^1 \) from **Step 1** and latent data \( (y_i^*)^0 \);

c. Compute the sample correlations \( \rho_{ij}^{\text{smpl}} \) between errors in equations \( i \) and \( j \);

d. Draw a candidate value of \( (\rho_{12}^{\text{cand}}, \rho_{13}^{\text{cand}}, \rho_{23}^{\text{cand}}) \) from the proposal density \( p^{\text{prop}}(\cdot|y^*, \beta) \), given by: \(^{21}\)

\[
\begin{align*}
\rho_{12} &\sim TN_{[-1,1]} \left( \rho_{12}^{\text{smpl}}, \sigma^2_\rho \right) \\
\rho_{13} &\sim TN_{[-1,1]} \left( \rho_{13}^{\text{smpl}}, \sigma^2_\rho \right) \\
\rho_{23} &\sim TN_{ \left[ \rho_{12}\rho_{13} - \sqrt{(1-\rho_{12}^2)(1-\rho_{13}^2)}, \rho_{12}\rho_{13} + \sqrt{(1-\rho_{12}^2)(1-\rho_{13}^2)} \right]} \left( \rho_{23}^{\text{smpl}}, \sigma^2_\rho \right),
\end{align*}
\]

Then construct the candidate covariance matrix \( \Sigma^{\text{cand}} \) as

\[
\Sigma^{\text{cand}} = \begin{pmatrix}
1 & \rho_{12}^{\text{cand}} & \rho_{13}^{\text{cand}} \\
\rho_{12}^{\text{cand}} & 1 & \rho_{23}^{\text{cand}} \\
\rho_{13}^{\text{cand}} & \rho_{23}^{\text{cand}} & 1
\end{pmatrix}.
\]

e. Accept the candidate covariance matrix \( \Sigma^{\text{cand}} \) and assign \( \Sigma^1 = \Sigma^{\text{cand}} \) with probability equal to

\[
\text{prob(accept)} = \min \left[ \frac{p(\rho_{12}^{\text{cand}}, \rho_{13}^{\text{cand}}, \rho_{23}^{\text{cand}} | (y^*)^0, \beta^1)p^{\text{prop}}(\rho_{12}^{0}, \rho_{13}^{0}, \rho_{23}^{0})}{p(\rho_{12}^{0}, \rho_{13}^{0}, \rho_{23}^{0} | (y^*)^0, \beta^1)p^{\text{prop}}(\rho_{12}^{0}, \rho_{13}^{0}, \rho_{23}^{0})}, 1 \right],
\]

where \( p(\cdot|y^*, \beta) \) is the conditional posterior density given in (2.15), otherwise \(^{21}\)

\(^{21}\)The variance of the proposal density was chosen so that the acceptance rate was roughly 35%. To determine the average acceptance rate the algorithm was initially run with 100 MH step replications. Subsequently, the number of MH steps was reduced to 10.
assign $\bar{\Sigma}^1 = \bar{\Sigma}^0$;

f. Repeat steps [d] and [e] $T$ times and assign $\Sigma^1 = \bar{\Sigma}^T$.

**Step 3:** Data augmentation step. Draw the latent data $(y_i^*)^1 = [(U_i^*)^1 (HI_i^*)^1 (RL_i^*)^1]^T$ conditional on $\beta^1$, and $\Sigma^1$:

a. Compute the errors $\varepsilon_2i$ and $\varepsilon_3i$ given $\beta^1$ from Step 1 and latent data $(HI_i^*)^0$ and $(RL_i^*)^0$;

b. Draw $(U_i^*)^1$ from

$$TN_{0,\infty}(x_i, \beta^1_1 + \gamma^1 HI_i + \left( \begin{array}{c} \rho_{12}^1 \\ \rho_{13}^1 \end{array} \right) \left( \begin{array}{cc} 1 & \rho_{23}^1 \\ \rho_{23}^1 & 1 \end{array} \right)^{-1} \left( \begin{array}{c} \varepsilon_2i \\ \varepsilon_3i \end{array} \right),$$

$$1 - \left( \begin{array}{c} \rho_{12}^1 \\ \rho_{13}^1 \end{array} \right) \left( \begin{array}{cc} 1 & \rho_{23}^1 \\ \rho_{23}^1 & 1 \end{array} \right)^{-1} \left( \begin{array}{c} \rho_{12}^1 \\ \rho_{13}^1 \end{array} \right),$$

if $U_i \geq 0$.

$$TN_{(-\infty,0)}(x_i, \beta^1_1 + \gamma^1 HI_i + \left( \begin{array}{c} \rho_{12}^1 \\ \rho_{13}^1 \end{array} \right) \left( \begin{array}{cc} 1 & \rho_{23}^1 \\ \rho_{23}^1 & 1 \end{array} \right)^{-1} \left( \begin{array}{c} \varepsilon_2i \\ \varepsilon_3i \end{array} \right),$$

$$1 - \left( \begin{array}{c} \rho_{12}^1 \\ \rho_{13}^1 \end{array} \right) \left( \begin{array}{cc} 1 & \rho_{23}^1 \\ \rho_{23}^1 & 1 \end{array} \right)^{-1} \left( \begin{array}{c} \rho_{12}^1 \\ \rho_{13}^1 \end{array} \right),$$

if $U_i \leq 0$.

c. Compute the errors $\varepsilon_1i$ given $\beta^1$ from Step 1 and latent data $(U_i^*)^1$;

---

22The Metropolis-Hastings is repeated to achieve faster convergence.
d. Draw \((HI_i^*)^1\) from
\[
TN_{[0,\infty)}(x_i\beta_2^1 + z_i\nu^1 + \left( \begin{array}{c} \rho_{12}^1 \\ \rho_{23}^1 \end{array} \right) \left( \begin{array}{cc} 1 & \rho_{13}^1 \\ \rho_{13}^1 & 1 \end{array} \right)^{-1} \left( \begin{array}{c} \varepsilon_{1i} \\ \varepsilon_{2i} \end{array} \right),
\]
\[
1 - \left( \begin{array}{c} \rho_{12}^1 \\ \rho_{23}^1 \end{array} \right) \left( \begin{array}{cc} 1 & \rho_{13}^1 \\ \rho_{13}^1 & 1 \end{array} \right)^{-1} \left( \begin{array}{c} \rho_{12}^1 \\ \rho_{23}^1 \end{array} \right),
\]
if \(HI_i \geq 0\).

\[
TN_{(-\infty,0]}(x_i\beta_2^1 + z_i\nu^1 + \left( \begin{array}{c} \rho_{12}^1 \\ \rho_{23}^1 \end{array} \right) \left( \begin{array}{cc} 1 & \rho_{13}^1 \\ \rho_{13}^1 & 1 \end{array} \right)^{-1} \left( \begin{array}{c} \varepsilon_{1i} \\ \varepsilon_{2i} \end{array} \right),
\]
\[
1 - \left( \begin{array}{c} \rho_{12}^1 \\ \rho_{23}^1 \end{array} \right) \left( \begin{array}{cc} 1 & \rho_{13}^1 \\ \rho_{13}^1 & 1 \end{array} \right)^{-1} \left( \begin{array}{c} \rho_{12}^1 \\ \rho_{23}^1 \end{array} \right),
\]
if \(HI_i \leq 0\).

e. Compute the errors \(\varepsilon_{2i}\) given \(\beta^1\) from Step 1 and latent data \((HI_i^*)^1\);

f. Draw \((RL_i^*)^1\) from
\[
TN_{[0,\infty)}(x_i\delta^1 + \left( \begin{array}{c} \rho_{13}^1 \\ \rho_{23}^1 \end{array} \right) \left( \begin{array}{cc} 1 & \rho_{12}^1 \\ \rho_{12}^1 & 1 \end{array} \right)^{-1} \left( \begin{array}{c} \varepsilon_{1i} \\ \varepsilon_{2i} \end{array} \right),
\]
\[
1 - \left( \begin{array}{c} \rho_{13}^1 \\ \rho_{23}^1 \end{array} \right) \left( \begin{array}{cc} 1 & \rho_{12}^1 \\ \rho_{12}^1 & 1 \end{array} \right)^{-1} \left( \begin{array}{c} \rho_{13}^1 \\ \rho_{23}^1 \end{array} \right),
\]
if \(RL_i \geq 0\).

\[
TN_{(-\infty,0]}(x_i\delta^1 + \left( \begin{array}{c} \rho_{13}^1 \\ \rho_{23}^1 \end{array} \right) \left( \begin{array}{cc} 1 & \rho_{12}^1 \\ \rho_{12}^1 & 1 \end{array} \right)^{-1} \left( \begin{array}{c} \varepsilon_{1i} \\ \varepsilon_{2i} \end{array} \right),
\]
\[
1 - \left( \begin{array}{c} \rho_{13}^1 \\ \rho_{23}^1 \end{array} \right) \left( \begin{array}{cc} 1 & \rho_{12}^1 \\ \rho_{12}^1 & 1 \end{array} \right)^{-1} \left( \begin{array}{c} \rho_{13}^1 \\ \rho_{23}^1 \end{array} \right),
\]
if \(RL_i \leq 0\).
**step 5:** repeat steps 1-4 $S$ times.

The Gibbs algorithm generates a sample of size $S$ from conditional posterior distribution of each of the parameters of the model. The first $S_0$ draws are discarded as burn-in because the Markov Chain has to converge to the joint posterior distribution of the parameters of the model. The remaining $S_1$ draws constitute the sample from the joint posterior distribution used for the analysis.

### 3 Data

I use the data from the Household Component (HC) of the Medical Expenditure Panel Survey (MEPS), a national survey of the US civilian non-institutionalized population administered by the Agency for Healthcare Research and Quality (AHRQ). The HC contains detailed information about individuals’ demographics, employment, income, health, health insurance status, and health care utilization. Although the MEPS is a panel survey, its longitudinal dimension is quite short. Respondents are interviewed five times over the course of two and a half years. It is an ongoing survey that began in 1996, with new panels introduced each year, resulting in annual files containing overlapping panels. I use the 2000 MEPS, the first year to contain a series of attitudinal questions that were asked in the paper-and-pencil Self-Administered Questionnaire (SAQ). There are four agree-disagree (ranging from 1-disagree strongly to 5-agree strongly) questions:\footnote{In the econometric implementation, I group individuals whose answers were ‘strongly disagree’ and ‘disagree somewhat’ into one class and the rest of responses into the other.}

1. *I do not need health insurance, I’m healthy enough;*
2. *Health insurance is not worth the money it costs;*
3. *I am more likely to take risks than the average person;*
4. *I can overcome illness without help from a medically trained person.*

The sample is restricted to the adult non-elderly population aged 18 to 64.\footnote{Age is computed as of July 1, 2000, the appropriate age for any analysis involving variables from SAQ.}
Monheit and Vistnes (2004) who used the same variables, I also exclude full-time students. Full-time students, children, and the elderly are likely to face substantially different choice sets regarding the insurance status. I also restrict the sample to individuals who were either insured the whole year or uninsured the whole year. The temporarily uninsured have been found to differ from both the full-year insured and the full-year uninsured in their health care utilization patterns (see Li and Trivedi (2004)). The group of part-year uninsured is highly heterogenous and, therefore, probably deserves special attention, which is beyond the scope of this paper. Following Monheit and Vistnes (2004), I also exclude individuals whose answers to the SAQ questionnaire were provided by a proxy (most often a spouse). After deleting observations with missing values, I arrived at a sample of 7,967 individuals. Definitions of all variables along with summary statistics are given in Table 1.\(^{25}\)

To study potential differences in incentive effects of health insurance on different types of health care utilization, I use four different dependent variables corresponding to four different types of health care utilization – office-based visits (OBVISIT), outpatient visits (OPVISIT), emergency room visits (ERVISIT), and hospital discharges (IPDIS). In addition, I use four different dependent variables that describe attitudes (NONEEDHI, HINOTWRTH, TAKERISK, and OVRCMILL), each corresponding to one of the SAQ questions described above. The system of equations is estimated 16 times, separately for each combination of the utilization variables and the attitudinal variables. Both private and public insurance were used in assigning the insurance status to each individual (INSURED).

The set of controls used in all three equations consists of demographic variables, education, family income, employment status, regional and MSA dummies, and a variety of measures of health status. It is well-known that in endogenous treatment models the treatment parameter \(\gamma\) is not non-parametrically identified. Identification in my model can be

\(^{25}\)Please refer to Appendix B.1.
achieved via the nonlinear functional form of the model, but for robustness I also impose a set of exclusion restrictions suggested by Li and Trivedi (2004). In particular, in the insurance equation I use a set of variables related to the demographic characteristics and health condition of dependents like children and spouse: spouse’s age (SPAGE), spouse’s education (SPEDU), a dummy for spouse’s any priority conditions (SPRICOND), and a dummy for presence of a child who easily gets sick (SICCHILD).26

4 Results

For each of the specifications, 25,000 draws from the posterior distribution were obtained. The first 5,000 were discarded as burn-in, and the remaining 20,000 were used for analysis. Posterior means and standard deviations of the parameters of interest (correlation coefficients and the endogenous treatment parameter $\gamma$) are given in Table 2.27

[Table 2 about here]

Two robust results emerge from Table 2. First, the posterior means and standard deviations of the treatment parameter $\gamma$ suggest that the probability of utilizing office-based and outpatient services is strongly affected by insurance status. At the same time, consistent with previous literature, I find no evidence of an incentive effect of health insurance in the consumption of inpatient or emergency room services. Second, the estimates of the covariance matrix $\Sigma$ do indicate the presence of private information. Although the overall selection effect, as measured by the correlation between unobservables in the utilization and insurance equations, $\rho_{12}$, is statistically not distinguishable from zero, the estimates of both $\rho_{13}$ and $\rho_{23}$ are negative and significant for most of the specifications. The presence

---

26The priority conditions include diabetes, asthma, high blood pressure, heart disease, stroke, emphysema, and joint pain.

27Please refer to Appendix B.2. The posterior statistics for all the remaining parameters of the model are not reported here to save space but are available upon request.
of selection on private information would not have been detected in the simpler endogenous
treatment model without incorporating the indicators of latent information.

The summary of results reported in Table 2 along with the theoretical predictions from
Section 2 are given below:

$$\Sigma(hypothesis) = \begin{pmatrix}
1 & +? & + \\
+? & 1 & - \\
+ & - & 1
\end{pmatrix}, \quad \Sigma(estimated) = \begin{pmatrix}
1 & 0 & - \\
0 & 1 & - \\
- & - & 1
\end{pmatrix}. \quad (4.1)$$

The presence of private information is indicated by the fact that, conditional on observ-
ables, both $\rho_{13}$ and $\rho_{23}$ are significantly different from zero. It appears that individuals
possess private information that is not completely captured by the observables and that is
used in making both health insurance and medical care decisions. The lack of overall se-
lection on private information combined with the estimated presence of private information
affecting both insurance and utilization decisions suggests that there are other types of pri-
vate information, which induce selection in the opposite direction. In particular, since the
aspect of private information measured by the SAQ indices used in this paper has negative
correlation with unobservables in both insurance and utilization, the correlation between
unobservables in the first and second equations should be positive. The fact that we do not
observe such a correlation implies that there are other sources of private information that
correlate with both decisions and that are not captured by the indexes used in this paper.

The negative posterior means of $\rho_{23}$ in all specifications are consistent with theory in
that more risk-loving individuals (or individuals who claim that they do not need health
insurance) tend to have a lower probability of having health insurance. The absolute values
of the posterior means are expectedly bigger in specifications involving the questions directly
related to the need for health insurance (“I do not need health insurance, I’m healthy enough”
and “Health insurance is not worth the money it costs”).

The correlation coefficient between unobservables in the utilization equation and risk-
aversion indicator equation, $\rho_{13}$, was hypothesized to be positive, which would indicate that more risk-loving individuals tend to have a higher probability of utilizing medical care. The posterior output on this correlation coefficient suggests that the contrary is true – more risk-loving individuals tend to be less likely to consume medical care. A notable exception is the specification involving the direct question on risk-aversion (“I am more likely to take risks than average person”) and emergency room utilization. A possible explanation of the negative estimates of $\rho_{13}$ is that the SAQ indices themselves reflect more than just attitudes toward risk taking behavior. Specifically, the first SAQ question (“I don’t need health insurance, I’m healthy enough”) makes a direct reference to the health state. Not surprisingly, the specifications involving this question show the consistently bigger estimates of $\rho_{13}$ in absolute value compared with specifications involving other questions. A similar argument applies to specifications involving the last SAQ question (“I can overcome illness without help of a medically trained person”). The direct reference to the ability to avoid medical help even in case of illness makes it logical to expect a negative correlation with observed utilization.

The lack of hypothesized positive correlation in specifications involving the direct risk-aversion question is harder to explain. A possible explanation is that the health care markets are more complex than basic contract theory would predict. In particular, standard asymmetric information models commonly assume that the outcome (accident, illness) is publicly observed and does not have to be verified (diagnosed). Furthermore, the consumption of medical care is driven not only by an individual’s objective health state, but also by her subjective perceptions of both her health state and of the potential usefulness of medical intervention, which may depend on risk-aversion. Although insignificant, the only positive estimate of $\rho_{13}$, which involves ER utilization and the direct risk-aversion question, is suggestive of the complex nature of imperfect information in health care markets. The pair of dependent variables (ERVISIT and TAKERISK) is probably least affected by other types of informational asymmetries inherent in medical care markets. On the one hand, some of
the ER visits result from easily observable injuries that require medical intervention by any subjective standards. On the other hand, TAKERISK provides the most direct measure of risk-aversion (not contaminated by attitudes toward other things).

It is also interesting to note that, although insignificant, the sign of the treatment coefficient $\gamma$ in the specifications involving emergency room utilization is negative. It contrasts with all other types of utilization and is potentially indicative of the proposition raised in health economics literature that some of the uninsured use the ER as a way to get non-emergency treatment. However, more research is needed on this question before solid conclusions can be drawn.

5 Concluding Remarks

In this paper I provide additional evidence on the selection and incentive effects of health insurance and extend the empirical literature by testing for the existence of multiple types (dimensions) of unobserved information. Using data on individual attitudes available in the MEPS survey, I am able to detect the presence of informational asymmetries even when the direct test does not signal any selection on unobservables. The evidence suggests that there are multiple types of private information inducing selection in opposite directions. In addition, I find strong evidence that insurance status affects the probability of using office-based and outpatient care. I also provide an additional interpretation of the indices of latent information available in the MEPS. Lastly, the results suggest that the role of risk-aversion in medical insurance markets is more complex than basic asymmetric information models suggest.

References


Appendix A

A.1 Likelihood of the Latent Data

Conditional on the parameters of the model, the likelihood can be expressed as

\[
p(y^*|\beta, \Sigma) = (2\pi)^{-\frac{n}{2}} |I_n \otimes \Sigma|^{-\frac{1}{2}} \exp \left( -\frac{1}{2} (y^* - X\beta)' (I_n \otimes \Sigma)^{-1} (y^* - X\beta) \right)
\]

\[
\propto \left| \Sigma \right|^{-\frac{1}{2}} \exp \left( -\frac{1}{2} (y^* - X\beta)' (I_n \otimes \Sigma^{-1}) (y^* - X\beta) \right)
\]

\[
\propto |\Sigma|^{-\frac{1}{2}} \exp \left( -\frac{1}{2} \sum_{i=1}^{n} \epsilon_i \Sigma^{-1} \epsilon_i \right)
\]

\[
\propto |\Sigma|^{-\frac{1}{2}} \exp \left( -\frac{1}{2} \sum_{i=1}^{n} (y^*_i - X_i\beta)' \Sigma^{-1} (y^*_i - X_i\beta) \right) \tag{A-1}
\]

A.2 Posterior of \( \beta \)

\[
p(\beta|y^*, \Sigma) \propto \exp \left( -\frac{1}{2} \sum_{i=1}^{n} (y^*_i - X_i\beta)' \Sigma^{-1} (y^*_i - X_i\beta) + (\beta - \mu_{\beta 0})' V_{\beta 0}^{-1} (\beta - \mu_{\beta 0}) \right)
\]

\[
\propto \exp \left( -\frac{1}{2} \sum_{i=1}^{n} (y^*_i' \Sigma^{-1} y^*_i - 2\beta' X_i' \Sigma^{-1} y^*_i + \beta' X_i' \Sigma^{-1} X_i\beta) \right.
\]

\[
\left. + (\beta' V_{\beta 0}^{-1} \beta - 2\beta' V_{\beta 0}^{-1} \mu_{\beta 0} + \mu_{\beta 0}' V_{\beta 0}^{-1} \mu_{\beta 0}) \right)
\]

\[
\propto \exp \left( -\frac{1}{2} \beta' \left( \sum_{i=1}^{n} X_i' \Sigma^{-1} X_i + V_{\beta 0}^{-1} \right) \beta - 2\beta' \left( \sum_{i=1}^{n} X_i' \Sigma^{-1} y^*_i + V_{\beta 0}^{-1} \mu_{\beta 0} \right) \right)
\]

\[
\propto \exp \left( -\frac{1}{2} \beta' V_{\beta 1}^{-1} \beta - 2\beta' V_{\beta 1}^{-1} \mu_{\beta 1} \right)
\]

\[
\propto \exp \left( -\frac{1}{2} \beta' V_{\beta 1}^{-1} \beta - 2\beta' V_{\beta 1}^{-1} \mu_{\beta 1} + \mu_{\beta 1}' V_{\beta 1}^{-1} \mu_{\beta 1} - \mu_{\beta 1}' V_{\beta 1}^{-1} \mu_{\beta 1} \right)
\]

\[
\propto \exp \left( -\frac{1}{2} (\beta - \mu_{\beta 1})' V_{\beta 1}^{-1} (\beta - \mu_{\beta 1}) \right) \tag{A-2}
\]

Therefore,

\[
p(\beta|y^*, \Sigma) = N(\mu_{\beta 1}, V_{\beta 1})
\]

\[
V_{\beta 1} = \left( \sum_{i=1}^{n} X_i' \Sigma^{-1} X_i + V_{\beta 0}^{-1} \right)^{-1}
\]

\[
\mu_{\beta 1} = V_{\beta 1} \left( \sum_{i=1}^{n} X_i' \Sigma^{-1} y^*_i + V_{\beta 0}^{-1} \mu_{\beta 0} \right) \tag{A-3}
\]
### Appendix B

#### B.1 Variable definitions

<table>
<thead>
<tr>
<th>variable</th>
<th>description</th>
<th>mean</th>
<th>st.dev.</th>
<th>min</th>
<th>max</th>
</tr>
</thead>
<tbody>
<tr>
<td>AGE</td>
<td>Age as of July, 2001 * 0.1</td>
<td>4.14</td>
<td>1.18</td>
<td>1.8</td>
<td>6.4</td>
</tr>
<tr>
<td>AGESQ</td>
<td>AGE squared</td>
<td>18.48</td>
<td>9.83</td>
<td>3.24</td>
<td>40.96</td>
</tr>
<tr>
<td>FEMALE</td>
<td>1 if female</td>
<td>0.57</td>
<td>0.50</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>FEMALEAGE</td>
<td>FEMALE*AGE</td>
<td>2.34</td>
<td>2.22</td>
<td>0</td>
<td>6.4</td>
</tr>
<tr>
<td>BLACK</td>
<td>1 if black</td>
<td>0.14</td>
<td>0.34</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>HISPANIC</td>
<td>1 if Hispanic</td>
<td>0.20</td>
<td>0.40</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>EDUYEARS</td>
<td>Years of education, when first entered MEPS</td>
<td>12.97</td>
<td>2.89</td>
<td>0</td>
<td>17</td>
</tr>
<tr>
<td>FAMSIZE</td>
<td>Family size</td>
<td>3.03</td>
<td>1.54</td>
<td>1</td>
<td>14</td>
</tr>
<tr>
<td>FAMINCOME</td>
<td>Family income in $10,000</td>
<td>5.91</td>
<td>4.3</td>
<td>-0.39</td>
<td>36.06</td>
</tr>
<tr>
<td>MARRIED</td>
<td>1 if married</td>
<td>0.61</td>
<td>0.49</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>EMPLOYED</td>
<td>1 if employed in the second-round interview</td>
<td>0.78</td>
<td>0.41</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>NORTHEAST</td>
<td>1 if resides in North-East Census Region</td>
<td>0.15</td>
<td>0.36</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>MIDWEST</td>
<td>1 if resides in Midwest Census Region</td>
<td>0.23</td>
<td>0.42</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>SOUTH</td>
<td>1 if resides in South Census Region</td>
<td>0.38</td>
<td>0.49</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>MSA</td>
<td>1 if resides in MSA</td>
<td>0.78</td>
<td>0.41</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>PHYSEXCEL</td>
<td>1 if perceived excellent physical health</td>
<td>0.25</td>
<td>0.43</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>PHYSPOOR</td>
<td>1 if perceived poor physical health</td>
<td>0.02</td>
<td>0.15</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>MENTEXCEL</td>
<td>1 if perceived excellent mental health</td>
<td>0.38</td>
<td>0.49</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>MENTPOOR</td>
<td>1 if perceived poor mental health</td>
<td>0.01</td>
<td>0.09</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>FUNCLIM</td>
<td>1 if reported functional limitations</td>
<td>0.07</td>
<td>0.25</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>SOCLIM</td>
<td>1 if reported social limitations</td>
<td>0.03</td>
<td>0.16</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>NPRIORCOND</td>
<td>Number of priority conditions</td>
<td>0.73</td>
<td>1.01</td>
<td>0</td>
<td>8</td>
</tr>
<tr>
<td>ILLINJ</td>
<td>1 if had an illness or injury requiring immediate medical attention</td>
<td>0.34</td>
<td>0.47</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>SPAGE</td>
<td>Spouse’s age</td>
<td>2.68</td>
<td>2.32</td>
<td>0</td>
<td>8.5</td>
</tr>
<tr>
<td>SPEDU</td>
<td>Spouse’s years of education</td>
<td>7.92</td>
<td>6.77</td>
<td>0</td>
<td>17</td>
</tr>
<tr>
<td>SPRICOND</td>
<td>1 if spouse has any priority conditions</td>
<td>0.29</td>
<td>0.46</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>SICCHILD</td>
<td>1 if has a child that easily falls sick</td>
<td>0.13</td>
<td>0.34</td>
<td>0</td>
<td>1</td>
</tr>
</tbody>
</table>

*continued on next page*
<table>
<thead>
<tr>
<th>variable</th>
<th>description</th>
<th>mean</th>
<th>st.dev.</th>
<th>min</th>
<th>max</th>
</tr>
</thead>
<tbody>
<tr>
<td>OBVISIT</td>
<td>1 if any office-based visits in 2000</td>
<td>0.71</td>
<td>0.45</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>OPVISIT</td>
<td>1 if any outpatient visits in 2000</td>
<td>0.15</td>
<td>0.36</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>ERVISIT</td>
<td>1 if any emergency room visits in 2000</td>
<td>0.11</td>
<td>0.31</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>IPDIS</td>
<td>1 if any hospital discharges in 2000</td>
<td>0.07</td>
<td>0.26</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>INSURED</td>
<td>1 if insured the whole year, =0, if uninsured the whole year</td>
<td>0.82</td>
<td>0.39</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>NONEEDHI</td>
<td>1 if didn’t disagree with ‘I do not need health insurance, I’m healthy enough’</td>
<td>0.14</td>
<td>0.35</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>HINOTWRTH</td>
<td>1 if didn’t disagree with ‘Health insurance is not worth the money it costs’</td>
<td>0.34</td>
<td>0.47</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>TAKERISK</td>
<td>1 if didn’t disagree with ‘I’m more likely to take risks than average person’</td>
<td>0.36</td>
<td>0.48</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>OVRCMILL</td>
<td>1 if didn’t disagree with ‘I can overcome illness without help from medically trained person’</td>
<td>0.34</td>
<td>0.47</td>
<td>0</td>
<td>1</td>
</tr>
</tbody>
</table>
### B.2 Results

Table 2: Results (standard deviations are in parentheses)

<table>
<thead>
<tr>
<th>dependent variables</th>
<th>$\gamma$</th>
<th>$\rho_{12}$</th>
<th>$\rho_{13}$</th>
<th>$\rho_{23}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>OBVISIT, INSURED, NONEEDHI</td>
<td>0.6811</td>
<td>-0.0263</td>
<td>-0.1507</td>
<td>-0.2571</td>
</tr>
<tr>
<td></td>
<td>(0.1811)</td>
<td>(0.0282)</td>
<td>(0.0256)</td>
<td></td>
</tr>
<tr>
<td>OBVISIT, INSURED, HINOTWRTH</td>
<td>0.6290</td>
<td>0.0044</td>
<td>-0.0635</td>
<td>-0.2721</td>
</tr>
<tr>
<td></td>
<td>(0.1635)</td>
<td>(0.0238)</td>
<td>(0.0220)</td>
<td></td>
</tr>
<tr>
<td>OBVISIT, INSURED, TAKERISK</td>
<td>0.7024</td>
<td>-0.0393</td>
<td>-0.0601</td>
<td>-0.1416</td>
</tr>
<tr>
<td></td>
<td>(0.1714)</td>
<td>(0.0218)</td>
<td>(0.0228)</td>
<td></td>
</tr>
<tr>
<td>OBVISIT, INSURED, OVRCMILL</td>
<td>0.7252</td>
<td>-0.0514</td>
<td>-0.1062</td>
<td>-0.1285</td>
</tr>
<tr>
<td></td>
<td>(0.1832)</td>
<td>(0.0218)</td>
<td>(0.0239)</td>
<td></td>
</tr>
<tr>
<td>OPVISIT, INSURED, NONEEDHI</td>
<td>0.5838</td>
<td>-0.1805</td>
<td>-0.0592</td>
<td>-0.2726</td>
</tr>
<tr>
<td></td>
<td>(0.2224)</td>
<td>(0.0253)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>OPVISIT, INSURED, HINOTWRTH</td>
<td>0.4820</td>
<td>-0.1189</td>
<td>-0.0592</td>
<td>-0.2726</td>
</tr>
<tr>
<td></td>
<td>(0.2173)</td>
<td>(0.0223)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>OPVISIT, INSURED, TAKERISK</td>
<td>0.5253</td>
<td>-0.1444</td>
<td>-0.0401</td>
<td>-0.1420</td>
</tr>
<tr>
<td></td>
<td>(0.2372)</td>
<td>(0.0232)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>OPVISIT, INSURED, OVRCMILL</td>
<td>0.5630</td>
<td>-0.1673</td>
<td>-0.0746</td>
<td>-0.1277</td>
</tr>
<tr>
<td></td>
<td>(0.2132)</td>
<td>(0.0239)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ERVISIT, INSURED, NONEEDHI</td>
<td>-0.0358</td>
<td>0.0485</td>
<td>-0.0857</td>
<td>-0.2556</td>
</tr>
<tr>
<td></td>
<td>(0.2397)</td>
<td>(0.0254)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ERVISIT, INSURED, HINOTWRTH</td>
<td>-0.0619</td>
<td>0.0632</td>
<td>-0.0620</td>
<td>-0.2722</td>
</tr>
<tr>
<td></td>
<td>(0.2302)</td>
<td>(0.0220)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ERVISIT, INSURED, TAKERISK</td>
<td>-0.0296</td>
<td>0.0442</td>
<td>0.0075</td>
<td>-0.1414</td>
</tr>
<tr>
<td></td>
<td>(0.2348)</td>
<td>(0.0230)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ERVISIT, INSURED, OVRCMILL</td>
<td>-0.0257</td>
<td>0.0422</td>
<td>-0.0680</td>
<td>-0.1280</td>
</tr>
<tr>
<td></td>
<td>(0.2400)</td>
<td>(0.0237)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>IPDIS, INSURED, NONEEDHI</td>
<td>0.1721</td>
<td>0.1158</td>
<td>-0.1027</td>
<td>-0.2561</td>
</tr>
<tr>
<td></td>
<td>(0.2948)</td>
<td>(0.0256)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>IPDIS, INSURED, HINOTWRTH</td>
<td>0.1393</td>
<td>0.1368</td>
<td>-0.1112</td>
<td>-0.2720</td>
</tr>
<tr>
<td></td>
<td>(0.2615)</td>
<td>(0.0217)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>IPDIS, INSURED, TAKERISK</td>
<td>0.0539</td>
<td>0.1802</td>
<td>-0.0972</td>
<td>-0.1414</td>
</tr>
<tr>
<td></td>
<td>(0.3030)</td>
<td>(0.0232)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>IPDIS, INSURED, OVRCMILL</td>
<td>0.1297</td>
<td>0.1412</td>
<td>-0.1020</td>
<td>-0.1277</td>
</tr>
<tr>
<td></td>
<td>(0.2560)</td>
<td>(0.0235)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>