ESSAYS IN RISK AND APPLIED BAYESIAN ECONOMETRIC MODELING

By

IBRAHIM LANDOLO CHEICK OUMAR NIANKARA

Bachelor of Sciences Business Economics with
Quantitative studies Options
Oklahoma State University
Stillwater, Oklahoma, USA
2007

Master of Science in Economics with a graduate Minor
in Statistics
Oklahoma State University
Stillwater, Oklahoma, USA
2009

Submitted to the Faculty of the
Graduate College of
Oklahoma State University
in partial fulfillment of
the requirements for
the Degree of
DOCTOR OF PHILOSOPHY
December, 2011

ESSAYS IN RISK AND APPLIED BAYESIAN ECONOMETRIC MODELING

Dissertation Approved:

Professor Lee C. Adkins

Dissertation Advisor

Associates Professor Mary N. Gade

Assistant Professor Harounan Kazianga

Associates Professor Melinda H. McCann

Outside Committee Member

Dr. Sheryl A. Tucker

Dean of the Graduate College

ACKNOWLEDGMENTS

My deepest gratitude goes to my advisor, Lee C. Adkins. He is truly a superior researcher, teacher and advisor. He guided me through this dissertation process with his broad and deep knowledge, insightful advice and constant encouragement. His devotion to excellence has always been inspiring. Without his support, this dissertation would be an impossible mission. I also want to thank Professors Jim Fain, Dan Rickman, J. B. Kim, for their advice and guidance all along.

I would like to express my genuine appreciation to Dr. Mary N. Gade, who has been a real mentor to me throughout my entire educational journey. I am also grateful to Dr. Harounan Kazianga for his valuable comments and suggestions. I want to thank Dr. McCann for providing me with a strong basis for grasping most of the tools I have put forth in my dissertation. Special thanks to Wilma Wehr, whose administrative assistance has helped make this dissertation process a smooth one. I also express my sincere gratitude to my adoptive family, the Holman's (Carole, Harley, Susan and Michael) without whom my educational dream would have been impossible. I'm also grateful to all my fraternity brothers, especially to Jordan Henricks, and Nick Jackson for allowing this educational dream to come true.

My family has always been the biggest source of my strength. I want to thank my parents, Niankara Karim, and Diarra Assita, for loving and caring for me. I am deeply indebted to my senior brother Aziz Niankara, my two senior sisters Aminata Niankara and Madiara Niankara, whom I look up to in so many aspects and owe every bit of my achievement to.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

The uncertain nature of our economic realm makes economic modeling a difficult exercise. There is hardly a situation where decisions are made with perfect information, therefore developing tools that can be used to model uncertainty is critical for both positive analysis (where the aim is at understanding human behavior), and normative analysis (where the aim is at making policy recommendations). In addition, recent developments in computing have led to the emergence of computationally intensive methods of analysis that provide the economic modeler with a greater freedom in addressing complex problems that were previously intractable. The goal of this dissertation is to model economic agents' behavior by considering the significance of uncertainty in decision making, and by putting into context the importance of Bayesian MCMC methods in expanding modeling possibilities. To this end, three essays are considered with emphasis on risk in scientific productivity, rationality and health insurance choice.

The first essay shows how risk affects the productivity of untenured faculty researchers. Junior faculty members under the tenure system are appointed on the basis of up-or-out contracts, requiring the faculty member to comply to vaguely specified departmental criteria. Failure to reach this standard leads to the loss of a job and its corresponding wage. On the basis of both, uncertainty in monetary value of publication (which is a price uncertainty) and publication uncertainty (which is an output uncertainty), a faculty member's output decision is modeled through the expected utility paradigm (VonNeumann and Morgenstern, 1944). This essay extends the analysis of Chen and Lee (2009), by looking at ex-ante incentive properties of tenure from the perspective of a junior faculty member, rather than a department as accustomed in the principal-agent framework. The results suggest that risk significantly affect faculty output decisions.

It was shown that a risk averse faculty member in the presence of publication value uncertainty, publishes at a point where the expected value of publication exceeds its marginal cost. Also a faculty member's scientific productivity was shown to be stimulated by increases in base salary when decreasing absolute risk aversion (DARA) described the faculty member's preference, while increased uncertainty in the value of publication provided a lesser incentive for scientific research output production. Moreover, when publication decision can be targeted at different journals, risk averse faculty had the economic incentive to target journals with negatively correlated per-unit returns.

In terms of academic field-wide effects, risk was shown to affect the distribution of faculty across journal tiers with the few most able faculty targeting publications at top-tier journals and most of the faculty community concentrated in the relatively lower tiers. This has implications for the observed increase in the volume of uncited published papers, which Mohamed (2004) suggested was due to the growing publish-or-perish emphasis of many academic institutions. This emphasis seems to have shifted faculty priorities away from risky quality publications in top-tier journals and toward relatively sure quantity publication in low-tier journals. Because top-tier journals are the ones more widely read and cited, this leaves a mass of uncited papers in low-tier publication outlets. Furthermore, risk was shown to affect faculty distribution across differently ranked academic departments, guaranteeing then viable matching between departments and faculty through its effect on faculty self selection mechanism. The most able faculty were shown to self select into top-ranked departments with high quality publication standards, while the less able faculty self selected into relatively lower ranked departments.

The second essay tests for the appropriateness of the rationality assumption in adults health insurance choice in the U.S., using the 2007 MEPS dataset. The question this essay attempts to answer is : given the sample evidence from the MEPS, is it reasonable to assume that adults in the U.S. behave rationally in their choices of health insurance? The rationality test is achieved here by looking at the consistency between respondents' stated preferences for health insurance in earlier rounds of the MEPS with their revealed health insurance choices at the end of the survey. The underlying assumption is that if agents are rational in their choices, then on average, we should observe some degree of consistency between stated preferences and revealed preferences for health insurance, accordingly with rational theory predictions. In fact, the results suggest rationality is appropriate in this context, therefore allowing one to use the discrete choice modeling framework to investigate health insurance choices by U.S. adults.

The evidence from the econometric estimation supports the two hypotheses. This implies consistency with the presented economic model's predictions in section 2 of the essay, and further suggests that the rationality assumption as a guiding mechanism for the 2007 MEPS data generating process, in the context of health insurance choice, is reasonable.

In addition to providing an answer to the intended question of interest, the study produces results that are consistent with the past literature, with respect to individuals' satisfaction (or optimism) toward health insurance consumption. The results in fact suggest relatively more skepticism (dissatisfaction) toward public only coverage compared to having some private coverage. This is because the strength of unlikeliness to seek health insurance for those that expressed health insurance to not be worth its cost relative to being uncertain, is stronger for public only coverage compared to having some private coverage.

Finally, the third essay addresses the issue of health insurance preference endogeneity in adults' health insurance enrollment decision in the U.S. within a Bayesian multinomial probit framework. The research goal here is to be able to say for a set of covariates, how the health insurance outcome probabilities vary based on differing health insurance preferences. The **R** package **endogMNP** (Burgette, 2010) is used to fit the model, with appropriate convergence diagnostics suggesting acceptable convergence of the chains. The sampler uses the marginalization principle of van Dyk (2010), producing marginal posterior distributions that are easy to interpret. For example we find that females are less likely than males to express "not-worthy" compared to "uncertain" as their preference for health insurance, while more likely than males to express "worthy" compared to "uncertain." In addition, the effect of college education on coverage through public only, is significant across all health insurance preference categories. Although this effect is almost similar for adults with uncertain, and Not-worthy preferences which are 0.2234 and 0.2218 respectively, it is relatively larger, 0.3681 for individuals with strong preference for health insurance (worthy). These coefficient values suggest that regardless of insurance preference, compared to adults with no college experience, those with at least one year of college experience are more likely to be publicly covered only, over being uninsured. Also, looking at the regional dummy variables, positive and significant coefficient values across all preference categories suggest that relative to southerners, adults from the Midwest are more likely to choose public coverage over being uninsured irrespective of insurance preference. Similarly, relative to southerners, adults from the Northeast are more likely to choose public coverage over being uninsured, however they only do so when they have Uncertain, or Not-Worthy preferences. Finally, the coefficient estimates for the health characteristics variables suggest that relative to having an EXCELLENT health condition, adults with GOOD or VERY-GOOD health conditions are more likely to have some private coverage, but only when their preference for health insurance is “Not-Worthy.” However, individuals with FAIR or POOR health conditions, relative to adults with EXCELLENT health condition, are more likely to have some private coverage when they are more decisive in the expression of their health insurance preference (worthy or Not-worthy).

CHAPTER 2

RISK AND JUNIOR FACULTY SCIENTIFIC PRODUCTIVITY INCENTIVES UNDER THE ACADEMIC TENURE SYSTEM

2.1 Introduction

Today more journals in a particular academic research field are published than anyone can reasonably keep up with. As a consequence many articles, both print and electronic, remain without a single citation for several years. This situation, as suggested by Mohamed (2004), is the mere result of the growing "publish-or-perish"emphasis of many academic institutions. In fact with the "up-or-out" rules that come with tenure-track appointment in institutions where scientific research output is the main and objective measure for tenure and promotion decisions, the emphasis on publishing introduces some degree of risk for faculty members and triggers behavioral response in terms of scientific productivity.

Kou and Zhou (2009) suggested that a university, by offering up-or-out contracts with probationary period and predetermined academic criterion, can ensure that professors produce knowledge through research activities. Formally up-or-out contracts are arrangements between a department and a faculty with the following features: (i) the department commits to retain the faculty for a pre-specified period; (ii) the faculty is considered for promotion only at the end of the probationary period, subject to satisfactory completion of some departmental criteria. If promoted, permanent retention is granted, otherwise the faculty member is permanently fired. Therefore the prospect of risk is prevalent during the probationary period junior faculty must go through to achieve tenure.

Freeman (1977) offers a risk-sharing explanation for the existence of the tenure system. In his opinion, the combination of tenure system and minimum wage policy is a risk-sharing mechanism encouraging risk-averse faculty to do risky but socially beneficial research projects. For McKenzie (1996), academic tenure is intended to guarantee the right to academic freedom, to allow original ideas to arise by giving scholars the intellectual autonomy to investigate the issues about which they are more passionate, and to report unbiased findings. While Carmichael (1988) argues that tenure provides older faculty members with the needed security to select new members of potentially greater ability, McKenzie (1996) on the other hand believes that what incumbents are really seeking is protection from their colleagues in a work environment operating under the rules of academic democracy, by increasing the costs predatory faculty members must incur to be successful in having more productive colleagues dismissed.

In more recent studies by Jing (2008) and Chen and Lee (2009), tenure-track appointment is found to play the role of a screening device by screening out low productivity faculty before tenure contract is signed. From the perspective of the junior faculty member seeking tenure such screening is risky because the knowledge specialization required for scientific progress puts researchers at risk of being misunderstood, and not rightly evaluated by other colleagues, especially in the short run (McPherson and Winston (1983)). Whereas a normative stand is taken in other studies to justify the existence of the tenure system, this paper takes a rather positive approach. Considering the fact that the tenure system is adopted by many academic institutions as an internal policy across the U.S., Canada and some European countries, I investigate the behavior of junior faculty prior to the tenure decision to see how sources of uncertainty affect scientific research output decisions.

This analysis is a follow up to Chen and Lee (2009), who model the ex-ante incentives produced by academic tenure under asymmetric information using a self-selection principal-agent model with unobservable type and action and examined the incentives in different institutions. As a principal-agent model, their analysis is focused on the interest of the principal (academic institution/ department), rather than that of the agent (faculty). Furthermore, none of the above referenced literature analyzes the tenure system from the perspective of the faculty member seeking tenure. Therefore in the current study I combine tools from the physics literature with modeling tools from probability theory and the economics of risk to focus on the interest of agents by analyzing ex-ante incentive properties of tenure from the perspective of a junior faculty member seeking tenure.

The modeling strategy implemented in this paper is the first of its kind on the topic. To this end, I start by first motivating risk under the academic tenure system in section 2, and then describe a way to quantify faculty scientific research output in section 3. In section 4, faculty publication decisions are analyzed in a single journal setting, first under uncertainty in the monetary value of publication, with some comparative statics, then under publication uncertainty with a general setting, followed by special cases. Section 5 extends the single journal setting to examine the implication of risk and risk aversion for faculty diversification strategies when publication can be targeted at more than one journal. Section 6 considers academic field wide implications of risk, while section 7 provides a discussion and model comparison, and Finally section 8 concludes the analysis.

2.2 Motivating Risk Under The Academic Tenure System

Risk is found in any situation where an event is not known with certainty (Chavas, 2004, p. 5). By this definition prospects for risk are widespread since the occurrence of any future event is almost always uncertain. In most academic institutions, where publication has become an imperative endeavor for the survival and prosperity of faculty members, the "up-or-out" rules that usually come with tenure-track appointment in such institutions introduces risk because of the long probationary period and faculty inability to fully control the publication process. Usually, a well established senior scholar is offered a tenure position directly while a junior faculty with an uncertain academic prospect has to experience a probationary period, by the end of which, he will obtain tenure if he has met some academic criterion and will be fired otherwise.

The prospect of being fired introduces an overall income uncertainty for the junior faculty member. Therefore, consistent with the expected utility hypothesis by Von-Neumann and Morgenstern (1944), it is assumed that faculty make research output decisions on the basis of the expected utility of this uncertain income. This overall income uncertainty can be motivated by two sources of uncertainty: a price uncertainty, and a quantity uncertainty. To see how, let I represent the junior faculty member's overall income, then it can be represented as the sum of a (nonrandom) base salary w, and the (random) incremental income py from publishing y research output each with monetary value p. Under such formulation, the income $I = w + py$ is also a random variable because of the randomness in the second term to the right of the equality. Randomness in this second term has two possible sources. It can be introduced through uncertainty in the monetary value of publication p, or through uncertainty in the publication of scientific research output y. For this reason, the analysis presented in this paper focuses on these two sources of tenure-track risk: Uncertainty in the monetary value of publication and Publication uncertainty.

2.2.1 Uncertainty in the monetary value of publication

Because of the specialized and highly sophisticated nature of academic work, its valuation is somewhat uncertain. According to McKenzie (1996) the benefits of scientific research projects undertaken by faculty are uncertain and sometimes may not be known for a long time. Moreover, the value may change with the passage of time, stressing the temporal dimension of risk. Therefore the use of probability theory as a formal structure describing and representing risky events allows us to model the monetary value of publication in this analysis as a random variable. For example, in a study of faculty compensation, Broder and Ziemer (1982) found that an additional American Journal of Agricultural Economics (AJAE) article published every other year increases faculty salary by $735/ year. Therefore publication has a monetary value attached to it, which as a random variable, implies that it can lead to different outcomes.

2.2.2 Publication uncertainty

because scientific research output production is the main and objective measure considered for tenure promotion, any randomness in the publication process introduces uncertainty in the final tenure outcome. In addition, the use of the h-index described in the next section as a measure of scientific productivity shows how publication uncertainty can affect faculty probability of getting tenure after the probationary period. The h-index is computed using the number of publication and the number of citations received by those publications, therefore the prospect of rejection from publishers and the lack of control over research output citations introduce uncertainty in the final tenure outcome. The randomness in the publication process is modeled using a stochastic publication function in later sections.

2.3 Quantifying Faculty Scientific Research Output

The h-index as adapted in this chapter was introduced in the physics literature by Hirsch (2005) as a useful index to characterize the scientific output of a researcher. Because junior faculty have a limited amount of time to prove their competence as researchers, this index allows for objective judgment of the impact and relevance of the faculty research work by the end of the probation period. Letting the probationary period set by the department be n years, then junior faculty publication records include among other things the number (N_p) of papers published over the n years, the number of citations (N_c^j) for each paper (j), the journals where the papers were published and the journal impact parameters. *A junior faculty is said to have index h if at the end of the probationary period, h of the faculty's N_p papers have at least h citations each, and the other $(N_p - h)$ papers have no more than h citations each.* Hirsch (2005) argues that two faculty with similar number of total papers or total citation count and very different h, the one with the higher h value is likely to be a more accomplished researcher. Therefore the h index measures the broad impact of the faculty's work and avoids disadvantages of the other single-number criteria commonly used to evaluate scientists research output such as: the total number of papers (N_p), the total number of citations($N_{c,tot}$), citations per paper, number of significant papers, etc.

To understand the h index, consider the case where faculty publishes p papers every year, and each publication receives c citations per year every subsequent year. The total number of citations when tenure decision is being made, that is at the $(n+1)$th year, is given by $N_{c,tot} = \sum_{j=1}^{n} pcj = \frac{pcn(n+1)}{2}$. If all papers up to year y contribute to the index, then we have

$$(n-y)c = h$$

$$py = h$$

where the left hand sides of the above two equations represent respectively the number of citations to the most recent of the papers contributing to h, and the total number of papers contributing to h. Combining those two equations yields $h = \frac{c}{1+c/p} n \approx m \cdot n$, where $m = \frac{c}{1+c/p}$.The total number of citations is approximately $N_{c,tot} \approx \frac{(1+c/p)^2}{2c/p} \cdot h^2 = a \cdot h^2$, where $a = \frac{(1+c/p)^2}{2c/p}$ therefore, the total number of citations received during the probationary period is proportional to the squared index value.

The linear relationship between the index h and the probationary period n should hold generally for junior faculty producing papers of similar quality at a steady rate over the course of the probationary period, however the slope m will vary from faculty to faculty, and so provides a useful measure for faculty comparison. This linear relationship beaks down however when the researcher slows down in paper production or stops publishing altogether. In such case a stretched exponential model may be more realistic as suggested by Hirsch (2005). Since the current analysis focuses on untenured risk averse faculty's productivity, it is reasonable to assume that faculty will not stop publishing during the probationary period, such that the linear relationship between h and n holds as a realistic model. As suggested by Hirsch (2005), a department may set a minimal value of h that a junior scholar must achieve by the end of the probationary period to secure tenure.

2.4 Faculty Decisions Under A Single Journal Setting

2.4.1 Uncertainty in the monetary value of publication

First, the market for academic employment is assumed to be competitive, with faculty research output targeted at a single journal in one's chosen field, such that publications have the same but uncertain monetary value. In the publication process, the faculty chooses the input vector $x = (x_1, \ldots, x_n)^{'}$, that include time, effort, and other resources used in the production of scientific research output. Faculty research output in terms of publications is denoted by y, and the publication technology represented by the function $y = f(x)$. The publication function $f(x)$ measures the largest feasible research output the faculty can obtain by committing the input vector $x = (x_1, \ldots, x_n)^{'}$. At this point, no uncertainty in the publication process (such case will be investigated in latter sections) is assumed. At the time research project decisions are made, the faculty tries to anticipate the uncertain monetary value she will receive from publishing her research output. As such, faculty treats the monetary value of publication p as a random variable, with a given subjective probability distribution.

Let $v = (v_1, \ldots, v_n)^{'}$ denote the respective prices paid for the inputs $x = (x_1, \ldots, x_n)^{'}$. Then the faculty cost of producing research output can be represented by $v^{'}x = \sum_{i=1}^{n} v_i x_i$, and the uncertain income generated is py. it follows that faculty net monetary benefit from publication can be represented by: $\tau = py - v^{'}x$. In addition, letting w denotes the base salary or initial wealth, then faculty terminal wealth is: $w + py - v^{'}x$. Given that the monetary value of publication is uncertain, this terminal wealth is also uncertain. Now assuming faculty behave in a way consistent with expected utility model, then the objective function of the faculty is

$$E[U(w + py - v^{'}x)] = E[U(w + \tau)] \tag{2.1}$$

where the E is the expectation operator based on the subjective probability distribution of the random variable p. It's assumed that faculty have risk-averse preferences represented by the utility function $U(\cdot)$ which satisfies $U^{'} \equiv \partial U/\partial w > 0$ and $U^{''} \equiv \partial^2 U/\partial w^2 < 0$. To see how the utility function $U(\cdot)$ summarizes all risk information relevant to faculty decisions, consider the following assumptions about faculty preferences among risky prospects b_1 and b_2 where,

$b_1 \sim^* b_2$ implies indifference between b_1 and b_2

$b_1 \geq^* b_2$ implies that b_2 is not preferred to b_1

$b_1 >^* b_2$ implies that b_1 is preferred to b_2

Assumption A1 *(ordering and transitivity)*

- For any random variables b_1 and b_2, exactly one of the following must hold: $b_1 >^* b_2, b_2 >^* b_1$ or $b_1 \sim^* b_2$.

- If $b_1 \geq^* b_2$ and $b_2 \geq^* b_3$ then $b_1 \geq^* b_3$. (transitivity)

Assumption A2 *(Independence)*

For any random variables b_1, b_2, b_3, and any $\alpha (0 < \alpha < 1)$, then $b_1 \leq^* b_2$ if and only if

$$[\alpha b_1 + (1-\alpha) b_3] \leq^* [\alpha b_2 + (1-\alpha) b_3].$$

(that is the preferences between b_1 and b_2 are independent of b_3)

Assumption A3 *(continuity)*

For any random variables b_1, b_2, b_3, where $b_1 <^* b_3 <^* b_2$, there exist numbers α and $\beta, (0 < \alpha < 1), (0 < \beta < 1)$, such that $b_3 <^* [\alpha b_2 + (1-\alpha) b_1]$ and $b_3 >^* [\beta b_2 + (1-\beta) b_1]$. (that is sufficiently small change in probabilities will not reverse a strict preference)

Assumption A4

For any risky prospects b_1, b_2 satisfying $Pr[b_1 \leq r : b_1 \leq^* r] = Pr[b_2 \geq r : b_2 \geq^* r] = 1$ for some sure reward r, then $b_2 \geq^* b_1$.

Assumption A5 *(ordering and transitivity)*

- For any number r, there exist two sequences of numbers $\alpha_1 \geq^* \alpha_2 \geq^* \ldots$ and $\beta_1 \leq^* \beta_2 \leq^* \ldots$ satisfying $\alpha_m \leq^* r$ and $r \leq^* \beta_n$ for some m and n.

- For any risky prospects b_1 and b_2, if there exists an integer m_o such that [b_1 conditional on $b_1 \geq \alpha_m : b_1 \geq^* \alpha_m$] $\geq^* b_2$ for every $m \geq m_o$, then $b_1 \geq^* b_2$. And if there exists an integer n_o such that [b_1 conditional on $b_1 \leq \beta_n : b_1 \leq^* \beta_n$] $\leq^* b_2$ for every $n \geq n_o$, then $b_1 \leq^* b_2$.

Under assumptions A1-A5, for any risky prospects b_1 and b_2, by *the expected utility theorem*, there exists a utility function $U(b)$ representing faculty risk preferences such that $b_1 \geq^* b_2$ if and only if $E[U(b_1)] \geq^* E[U(b_2)]$, with $U(b)$ defined up to *a positive linear transformation*. See (DeGroot, 1970, p. 113-114) for a proof of the expected utility theorem.

Therefore, under the assumption A1-A5, the expected utility hypothesis provides an accurate characterization of faculty behavior under risk in general, and so can be used to model faculty response to risk during the probationary period in tenure-track appointment.

Result 1. The definition of the utility function up to a *positive linear transformation*, implies that, if $U(b)$ is a utility function for a particular faculty, then so is the linear transformation $Z(b) = \alpha + \beta U(b)$ for any α and $\beta > 0$ scalars.

Proof. Starting from the equivalence between $b_1 \geq^* b_2$ and $E[U(b_1)] \geq^* E[U(b_2)]$, stated in the expected utility theorem, given $\beta > 0$, $E[U(b_1)] \geq^* E[U(b_2)]$ is equivalent to $\alpha + \beta E[U(b_1)] \geq \alpha + \beta E[U(b_2)]$, which is also equivalent to $E[Z(b_1)] \geq E[Z(b_2)]$. Therefore, $b_1 \geq^* b_2$ if and only if $E[Z(b_1)] \geq E[Z(b_2)]$, implying that $Z(\cdot)$ and $U(\cdot)$ provide equivalent representations of a faculty member's risk preferences. □

This characteristic further implies that without affecting a faculty member's preference ranking, the utility function $U(b)$ can be shifted by changing its intercept and/or by multiplying its slope by a positive constant. this special feature will be useful for our analysis.

Now, letting $\mu = E(p)$ be the expected monetary value of publication, then p can be represented as $p = \mu + \sigma e$. Where e is a random variable with zero mean, and introducing randomness in the monetary value of publication p. The random component e can exhibit any distribution for which both the mean and the variance exist. The standard deviation of the monetary value of publication, σ can be interpreted as a mean-preserving spread parameter for the distribution of p. Therefore, in this analysis, the probability distribution of p will be characterized by the mean μ and the mean preserving spread parameter σ following the analysis of firm production under uncertainty by Sandmo (1971).

Under the expected utility model, a faculty member's publication decision can be represented by

$$Max_{x,y}\{E[U(w+py-v^{'}x):y=f(x)]\}. \tag{2.2}$$

which states that publication decisions are made in a way consistent with expected utility maximization

Faculty costs minimizing behavior

In the absence of publication uncertainty, expected utility maximization implies cost minimization by faculty. Faculty will minimize the cost of the inputs used in the publication process, which include the effort cost, the opportunity cost of time allocated to research, and the costs of other used resources. To see that, note that the maximization problem in equation (2.2) can be written as

$$Max_y\{Max_x\{E[U(w+py-v^{'}x):y=f(x)]\}\}$$

$$=Max_y\{E[U(w+py+Max_x\{-v^{'}x:y=f(x)\}]\}$$

$$=Max_y\{E[U\left(w+py-Min_x\{v^{'}x:y=f(x)\}\right)]\}$$

$$=Max_y\{E[U(w+py-C(v,y)]\}. \tag{2.3}$$

where $C(v,y) = Min_x\{v^{'}x : y = f(x)\}$ is the publication cost function similar to the cost function in standard production theory under certainty. Therefore, for given input prices v, $C(v,y)$ measures the smallest possible cost of producing research output y. Where output here is measured in terms of publications. This shows that in the absence of publication risk, risk averse junior faculty has the incentive to behave in a cost minimizing fashion.

Equation (2.3) offers a convenient way of analyzing a junior faculty member's behavior, since it involves choosing only one variable: y, faculty scientific research output. Now assuming that the scholar decides on positive research output, $y > 0$, then using the chain rule, the first-order necessary condition associated with the optimal choice of y is given by:

$$F(y,\cdot) \equiv E[U^{'} \cdot (p - C_y)] = 0, \tag{2.4}$$

or recalling from rules of probability theory that $Cov(U^{'}, p) = E(U^{'} \cdot p) - E(U^{'}) \cdot E(p)$, then $E(U^{'} \cdot p) = E(U^{'}) \cdot \mu + Cov(U^{'}, p)$, and equation (2.4) can be rewritten as:

$$\mu - C_y + Cov(U^{'}, p)/EU^{'} = 0 \tag{2.5}$$

where $C_y \equiv \partial C/\partial y$ denotes the marginal cost of publication, $U^{'} \equiv \partial U/\partial y$, and $Cov(U^{'}, p) = E(U^{'}\sigma e)$. The associated second order condition for a maximum is

$$D \equiv \partial F/\partial y \equiv E[U^{'} \cdot (-C_{yy})] + E[U^{''} \cdot (p - C_y)^2] < 0. \tag{2.6}$$

Define R to be the Arrow-Pratt measure of risk premium, which shows the shadow cost of private risk bearing by faculty. Then R is the monetary value satisfying the indifference relationship $\{w + py - C(v, y)\} \sim^{*} \{w + E(p) \cdot y - C(v, y) - R\}$ and under the expected utility model R is the solution of the equation:

$$E[U[w + py - C(v, y)]] = U[w + E(p) \cdot y - C(v, y) - R] \tag{2.7}$$

So, given that $U[w+py-C(v, y)]$ is a strictly increasing function, its inverse always exists. Denoting the inverse by U^{-1}, then it follows that $U^{-1}\{E[U[w+py-C(v, y)]]\} = w + E(p) \cdot y - C(v, y) - R$, thus the risk premium can be written as

$$R(w, y, \cdot) = w + \mu \cdot y - C(v, y) - U^{-1}\{E[U[w + py - C(v, y)]]\}. \tag{2.8}$$

Maximizing the expected utility as shown in equation (2.3) is equivalent to maximizing *"the certainty equivalent"* $w + \mu \cdot y - C(v, y) - R(w, y, \cdot)$. It follows then that faculty publication decision can alternatively be written as:

$$Max_y[w + \mu \cdot y - C(v, y) - R(w, y, \cdot)], \tag{2.9}$$

with the associated first-order condition given by

$$\mu - C_y(v, y) - R_y(w, y, \cdot) = 0 \tag{2.10}$$

where $R_y(w, y, \cdot) \equiv \partial R/\partial y$ is the *marginal risk premium.* Comparing this result with the first order condition in equation(2.5), it follows that $R_y(w, y, \cdot) = -Cov(U^{'}, p)/E(U^{'})$, providing an intuitive interpretation for the covariance term: $[-Cov(U^{'}, p)/E(U^{'})]$ as *the marginal risk premium*, measuring the marginal effect of scientific research output production, on the implicit cost of private risk bearing by faculty.

The publication function

The publication function is the function $y^*(w, \mu, \sigma)$ satisfying the first-order condition in equation (2.5), or in equation (2.10):

$$\mu = C_y(v, y) + R_y(w, y, \cdot) \tag{2.11}$$

The condition in equation (2.11) implies that, at the optimum research output publication y^*, the expected monetary value of publication μ is equal to the marginal cost of publication C_y, plus the marginal risk premium R_y. This means that expressing the sum $(C_y + R_y)$ as a function of research output y gives *the publication function*, and generates the schedule of scientific research output production by the risk-averse faculty member, for each level of expected monetary value of publication μ.

Result 2. Under uncertainty in the monetary value of publication, and risk aversion, a junior scholar publishes at a point where the expected monetary value of publication μ exceeds the marginal cost of publication (C_y).

Proof. if $\partial U^{'}/\partial p < 0 (> 0)$, then $U^{'}$ and p move in the opposite(same) direction(s), implying a negative (positive) covariance, therefore the covariance term $Cov(U^{'}, p)$ is always of the sign of $\partial U^{'}/\partial p$. But $sign[\partial U^{'}/\partial p] = sign(U^{''} \cdot y)$. Thus, risk aversion (where $U^{''} < 0$) implies that $Cov(U^{'}, p) < 0$. And it follows that the marginal risk premium $R_y = -Cov(U^{'}, p)/E(U^{'}) > 0$ under risk aversion, which in turn implies that $\mu > C_y$ at the optimum. □

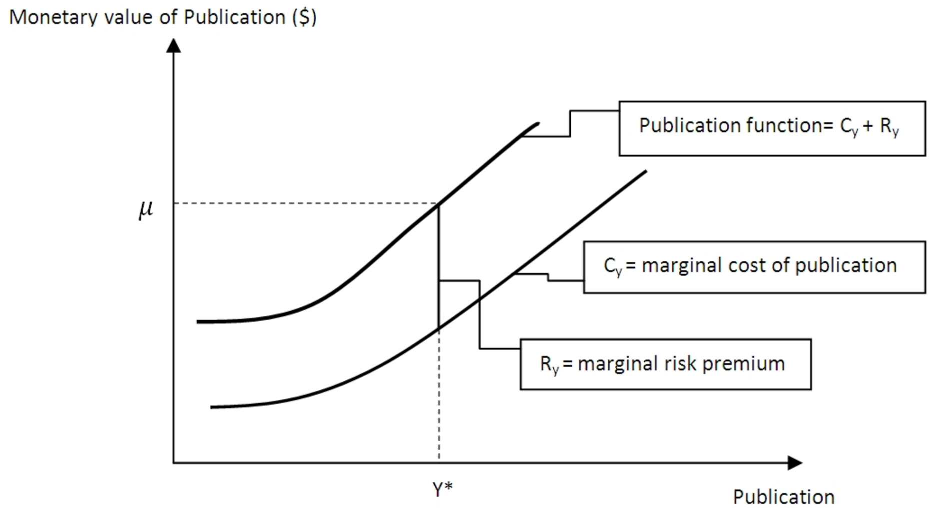

Figure 2.1: Publication function

The publication function is illustrated in figure (2.1) , and shows that under risk aversion, risk can have significant effects on a faculty member's resource allocation. The analysis shows that, while risk does not involve any explicit cost to faculty, its implicit cost (as measured by the marginal risk premium R_y) needs to be added to the marginal cost of publication C_y in the evaluation of optimal publication decisions.

Comparative statics analysis

Since risk affects faculty scientific productivity under risk aversion, this effect is investigated in more detail by conducting a comparative static analysis of the research output decision y in equation (2.4). Letting $\alpha = (w, \mu, \sigma)$ be the vector of parameters of the publication function $y^*(\alpha)$, then using the chain rule and total differentiating the first-order condition $F(y, \alpha) = 0$ at the optimum $y = y^*(\alpha)$ yields

$$\partial F/\partial \alpha + (\partial F/\partial y)(\partial y^*/\partial \alpha) = 0,$$

or, with $D = \partial F/\partial y < 0$, (from equation (2.6))

$$\begin{aligned}\partial y^*/\partial \alpha &= -D^{-1}\partial F/\partial \alpha \\ &= -D^{-1}\partial\{E[U' \cdot (p - C_y)]\}/\partial \alpha \\ &= sign(\partial\{E[U' \cdot (p - C_y)]\}/\partial \alpha).\end{aligned}$$

this result is used to analyze the properties of the publication function $y^*(\alpha)$, looking at changes in the elements of the parameter vector $\alpha = (w, \mu, \sigma)$.

4.1.3(a) The effect of a change in base salary, w

The effect of changing faculty base Salary (initial wealth) w is given by

$$\partial y^*/\partial w = -D^{-1}\{\partial\{E[U^{'} \cdot (p - C_y)]\}/\partial w\} = -D^{-1}\{E[U^{''} \cdot (p - C_y)]\}. \qquad (2.12)$$

Assuming faculty preferences exhibit decreasing absolute risk aversion (DARA), then the term $E[U^{''} \cdot (p - C_y)] > 0$. To see that, consider the Arrow-Pratt absolute risk aversion coefficient $r = -U^{''}/U^{'}$, and any risky return b then the result $R \approx -0.5(U^{''}/U^{'}) \cdot Var(b)$ by (Chavas, 2004, p. 36-37) provides a link between the risk premium R and the Arrow-Pratt coefficient of absolute risk aversion $r = -U^{''}/U^{'}$. Because $var(b) > 0$ for all b non degenerate, the sign of the risk premium R always is the same as that of r, and so for risk averse faculty ($R > 0$), the corresponding $r = -U^{''}/U^{'} > 0$. Therefore, letting τ_o denote the net benefit from publication τ, when evaluated at $p = C_y$, then under DARA,

$$r(\tau) \lesseqgtr r(\tau_o) \text{ if } p \gtreqless C_y,$$

it follows that

$$-U^{''}/U^{'} \lesseqgtr r(\tau_o) \text{ for } p \lesseqgtr C_y,$$

or

$$U^{''} \lesseqgtr -r(\tau_o) \cdot U^{'} \text{ for } (p - C_y) \gtreqless 0,$$

or

$$U^{''} \cdot (p - C_y) > -r(\tau_o) \cdot U^{'} \cdot (p - C_y),$$

and taking the expectation on both sides of the inequality yields

$$E[U^{''} \cdot (p - C_y)] > -r(\tau_o) \cdot E[U^{'} \cdot (p - C_y)] = 0,$$

with the last equality on the RHS coming from the first order conditions in equation (2.4). Therefore $E[U^{''} \cdot (p - C_y)] > 0$ as required and $\partial y^*/\partial w > 0$.

Result 3. Under uncertainty in the monetary value of publication, if DARA characterizes faculty preferences, increasing a junior faculty member's base salary w tends

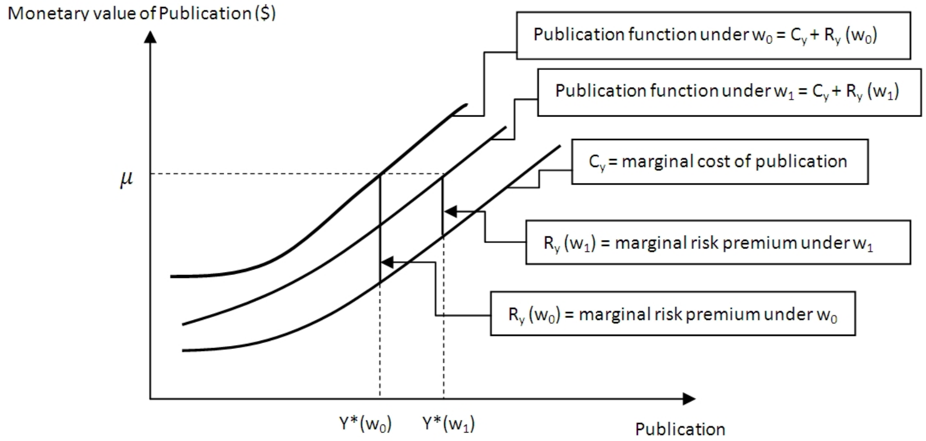

Figure 2.2: Effect of changing base salary w under DARA, with $w_1 > w_0$

to stimulate scientific research output production by decreasing the implicit cost of private risk bearing. This occurs because under DARA, private wealth accumulation and insurance motives are substitutes.

This result is illustrated in figure (2.2) and shows that, under DARA, increasing the faculty member's base salary w from w_0 to w_1 reduces the marginal risk premium $R_y(w)$. This is because under risk aversion and DARA, private wealth accumulation reduces the risk premium R, which is accompanied by a reduction in the marginal risk premium $R_y(w)$, which further generates a shift to the right of the publication function $(C_y + R_y)$.

4.1.3(b) The effect of a change in the expected monetary value of publication μ

The effect of changing expected monetary value of publication μ is given by

$$\partial y^*/\partial\mu = -D^{-1}\{\partial\{E[U^{'}\cdot(p-C_y)]\}/\partial\mu\} = -D^{-1}\{E[U^{'}+yE[U^{''}\cdot(p-C_y)]\}. \quad (2.13)$$

Now defining $\partial y^c/\partial\mu \equiv -D^{-1}[E(U^{'})]$ as the *compensated expected monetary value effect*, and given that $D < 0$ (from the second-order condition in equation (2.6)), it follows that $\partial y^c/\partial\mu > 0$. This in turn implies that the *"compensated" publication*

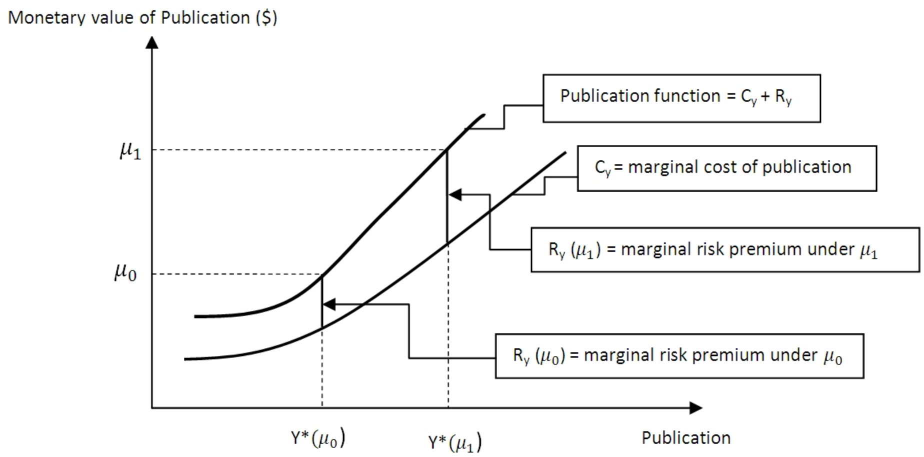

Figure 2.3: The effect of changing mean value of publication μ under DARA, with $\mu_1 > \mu_0$

function is always upward sloping with respect to changes in the expected monetary value of publication μ. Also, recalling from equation (2.12) that $\partial y^*/\partial w = -D^{-1}\{E[U'' \cdot (p - C_y))]$, then we have the following Slutsky equation:

$$\partial y^*/\partial \mu = \partial y^c/\partial \mu + (\partial y^*/\partial w) \cdot y^*, \tag{2.14}$$

which suggests that the slope of the uncompensated monetary value of publication $\partial y^*/\partial \mu$ is equal to that of the compensated monetary value $\partial y^c/\partial \mu$, plus an income effect $(\partial y^*/\partial w) \cdot y^*$. Under DARA, it was shown in result 3 that $(\partial y^*/\partial w) > 0$, therefore the income effect $(\partial y^*/\partial w) \cdot y^*$ is also positive. Given that the compensated effect is $\partial y^c/\partial \mu > 0$, from the Slutsky equation it follows that $\partial y^*/\partial \mu > 0$. This implies that the publication function exhibits a positive slope with respect to the uncompensated monetary value of publication.

Result 4. Under uncertainty in the monetary value of publication, if DARA characterizes faculty preferences, then an increase in expected monetary value of scientific research output stimulates a junior faculty member's publication effort.

This result is illustrated in figure (2.3) and shows that under DARA increasing the expected monetary value of research output μ from μ_0 to μ_1 increases publications $y^*(\mu_1) > y^*(\mu_0)$, since the publication function is upward sloping under DARA. Also, figure (2.3) suggests that the marginal cost C_y and the marginal risk premium R_y are higher at μ_1 compared to μ_0.

4.1.3(c) The effect of a change in publication risk σ

The effect of changing the mean-preserving parameter value σ, which also represents the risk in the monetary value of publication, is given as follows, with a standardized value of sigma($\sigma = 1$)

$$\partial y^*/\partial\sigma = -D^{-1}\{\partial\{E[U^{'} \cdot (p - C_y)]\}/\partial\sigma\}, \tag{2.15}$$

$$= -D^{-1}\{E(U^{'} \cdot e) + yE[U^{''} \cdot (p - \mu)(p - C_y)]\},$$

$$= -D^{-1}\{E(U^{'} \cdot e) + yE[U^{''} \cdot (p - C_y + C_y - \mu)(p - C_y)]\},$$

$$= -D^{-1}\{E(U^{'} \cdot e) + yE[U^{''} \cdot (p - C_y)^2] + y(C_y - \mu)E[U^{''} \cdot (p - C_y)]\}.$$

But $E(U^{'} \cdot e) = Cov(U^{'}, p)$, since $p = \mu + \sigma e$ with $E(e) = 0$. In addition, $Cov(U^{'}, p) = sign(U^{''}y) < 0$ under risk aversion, and $(U^{''} < 0)$ implies $E[U^{''} \cdot (p - C_y)^2] < 0$. Furthermore, it was shown in results (2) and (3) that under DARA, $C_y - \mu < 0$ and $E[U^{''} \cdot (p - C_y)] > 0$. It follows then from equation (2.15), that $\partial y^*/\partial\sigma < 0$, that is, increasing publication value risk (σ), decreases optimal research output.

Result 5. Under risk aversion, if DARA characterizes junior faculty preferences, for a given mean return to publication, an increase in publication value risk provides a general disincentive to publish.

This result is illustrated in Figure (2.4), where increasing σ from σ_0 to σ_1, shifts the publication function to the left. This increasing risk exposure increases the scholar's private risk bearing as measured by the risk premium R, and so increases the marginal risk premium R_y. The publication function $(C_y + R_y)$ increases such that for a given value of publication, research output falls.

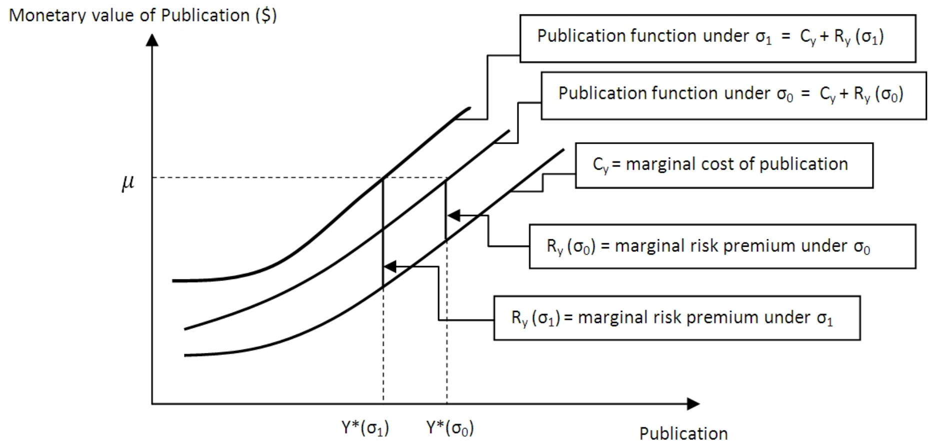

Figure 2.4: The effect of changing publication value risk σ under DARA, with $\sigma_1 > \sigma_0$

2.4.2 Publication uncertainty

So far, the analysis has focused on faculty behavior under uncertainty in the monetary value of publication alone. The sources of uncertainty are now extended by introducing the second source of tenure-track risk in the form of publication uncertainty, where risk is introduced through faculty research output as well rather than just the value of the output itself. This is motivated by the fact that while faculty have the power to make choices over the inputs committed in the publication process, the outcome of the process is beyond the faculty member's control. To this end, a general setting of such uncertainty, is considered, followed by special cases for practical considerations.

The general setting

Under general publication uncertainty, a junior faculty member's research output is a random variable at the time input choices are being made. The publication technology can be represented by a stochastic publication function. A generic specification of this function is $y(x, e)$, where y is the research output, x is a vector of inputs, and e is a random variable reflecting publication uncertainty.

The stochastic publication function $y(x, e)$ provides the maximum possible research output that can be obtained when the input vector x is committed and the random variable e is realized. The faculty member is assumed to have information about publication uncertainty, information represented by a subjective probability distribution of the random variable e.

Building on price uncertainty from the previous section, let p be the monetary value of the faculty member's research output, v, represent the prices of the inputs used in the publication process, and w the base salary received by faculty. Then, the income generated by the junior scholar's publications is $py(x, e)$, the cost of publication is $v^{'}x = \sum_{i=1}^{n} v_i x_i$,the net return from publication is $\tau = py(x, e) - v^{'}x$, and faculty terminal income (wealth) is $(w + \tau)$.

Allowing for uncertainty in the monetary value of publication, then the faculty member does not know either e or p at the time research input decisions are being made. In such setting, the faculty member faces both price and publication risks and so treats both e and p as random variables with a given subjective joint probability distribution. Under the expected utility model, the junior faculty's objective function is to choose inputs (time, effort, and other resources used in the publication process) so as to maximize expected utility of terminal income

$$Max_x\{E[U(w + py(x, e) - v^{'}x)]\} \tag{2.16}$$

where the expectation E is based on the joint subjective distribution of the random variables (p, e). Using the chain rule, the first order necessary conditions for the optimal choice of inputs is:

$$E[U^{'} \cdot (p\partial y(x, e)/\partial x - v)] = 0,$$

or

$$E[(p\partial y(x, e)/\partial x] = v - Cov[U^{'}, p\partial y(x, e)/\partial x]/E(U^{'}),$$

or

$$E(p)E[\partial y(x, e)/\partial x] + Cov[p, \partial y(x, e)/\partial x] = v - Cov[U^{'}, p\partial y(x, e)/\partial x]/E(U^{'}). \tag{2.17}$$

As shown in equation (2.9), maximizing expected utility is equivalent to maximizing the corresponding certainty equivalent. The certainty equivalent of faculty terminal income (wealth) is $w + E[py(x,e)] - v^{'}x - R(x,\cdot)$, where $R(x,\cdot)$ represents the Arrow-Pratt risk premium as before. Therefore the maximization problem in equation (2.16), can be rewritten as

$$Max_x\{w + E[py(x,e)] - v^{'}x - R(x,\cdot)\}$$

with first-order conditions

$$\partial E[py(x,e)]/\partial x - v - R_x(x,\cdot) = 0,$$

or

$$\partial E[py(x,e)]/\partial x = v + R_x(x,\cdot). \tag{2.18}$$

$R_x(x,\cdot) \equiv \partial R(x,\cdot)/\partial x$ represents the marginal risk premium. Comparing the first-order conditions in equation (2.17) to equation (2.18) indicates that the marginal risk premium takes the form: $R_x(x,\cdot) = -Cov[U^{'}, p\partial y(x,e)/\partial x]/E(U^{'})$. This equality allows us to characterize the covariance term $-Cov[U^{'}, p\partial y(x,e)/\partial x]/E(U^{'})$ as the marginal risk premium measuring the effects of the committed input vector x, on the implicit cost of private risk bearing by the junior faculty member. It also shows that at optimal input commitment by the junior faculty, the expected marginal value of research output, $\partial E[py(x,e)]/\partial x$, is equal to the per unit cost of committed input v plus the marginal risk premium, $R_x(x,\cdot)$.

In the general form of the stochastic publication function $y(x,e)$, the marginal risk premium can be either positive, negative, or zero. Whether a particular input increases or decreases the implicit cost of risk faced by the faculty member is largely an empirical matter. However, for risk-averse faculty, the risk premium $R(x,\cdot) > 0$ and so when the marginal risk premium $R_{x_i}(x,\cdot) > 0$, a faculty member will have an incentive to reduce the use of the i-th input because this input increases the implicit cost of risk bearing.

For example when the i-th input is "effort" exerted in the research process, then a positive marginal risk premium $R_{x_i}(x, \cdot) > 0$ provides the junior scholar with the incentive to reduce such"effort." Conversely when $R_{x_i}(x, \cdot) < 0$, the i-th input reduces the implicit cost of risk, and so gives the junior faculty member the incentive to increase the use of this input in the publication process. For example, if the i-th input is "time," then $R_{x_i}(x, \cdot) < 0$, gives the junior scholar the incentive to put more time in the research activity. For empirical tractability, possible specifications of the general form of the publication function $y(x, e)$ are now considered.

The stochastic publication function

There are two specifications of the stochastic publication function considered: one proposed by Just and Pope (1978, 1979) and a moment base approach based on Antle (1983).

Just-Pope specification: Just and Pope (1978, 1979) propose flexible specifications of stochastic production functions in general. Because of the parallelism between the publication function developed in this paper and the standard stochastic production function, a similar specification is adopted for the stochastic publication function, which can be represented as:

$$y(x, e) = f(x) + e[h(x)]^{1/2}, \tag{2.19}$$

where $E(e) = 0$ and $Var(e) > 0$. This specification of the publication function implies that $E(y) = f(x)$ and $\partial E(y)/\partial x = \partial f(x)/\partial x$, and that $Var(y) = Var(e)h(x)$ with $\partial Var(y)/\partial x = Var(e) \cdot \partial h(x)/\partial x$. Since the $var(e) > 0$, the sign of $\partial Var(y)/\partial x$ depends on that of $\partial h(x)/\partial x$.

The publication function as specified can be interpreted as a regression model exhibiting heteroskedasticity, where the interest is on identifying the effects of faculty inputs decision on the variance $Var(y)$ of the research output produced. A higher research output variance effect implies a riskier input choice. In this Just-Pope like specification of the publication function, depending on the functional form of $h(x)$, the marginal risk premium $\partial h(x)/\partial x$ can be negative, positive, or zero. And so, inputs used in the publication process will be identified as risk reducing, risk increasing, or risk neutral. In situations where the inputs affect publication risk, $\partial h(x)/\partial x \neq 0$, faculty can manage risk exposure through judicious choice of inputs. Under risk aversion, where $R(x, \cdot) > 0$, faculty have an incentive to use inputs which reduce risk exposure and its implicit cost and for which $\partial h(x)/\partial x < 0$. In such setting, risk has a direct impact on faculty input allocation and thus publication decisions.

Moment based approach: A moment based approach, which includes the Just-Pope specification as a special case, is adapted from Antle (1983). The Just-Pope specification, also referred to as mean-variance specification, is based on only the first two moments of the distribution of the stochastic publication function and as such is a special case of the more general moment based approach. The moment based approach allows for the empirical exploration of the role of higher-order moments, and can capture more interesting features of the stochastic publication function. Considering the stochastic publication function in its generic form $y(x, e)$, then the moment generating function(MGF) of the random research output y if it exists is given by $M_y(t) = E(\exp(ty))$. The r-th derivative of the MGF evaluated at $t = 0$ gives the r-th moment about the origin

$$M_Y^r(0) = E(Y^r) \; \forall \; r = 1, 2, \ldots$$

from which the first central moment, the mean of the publication function given the committed input vector x, is obtained by setting r=1. Let the mean be denoted as

$$\mu(x) = E[y(x, e)],$$

then the r-th moment about the mean of the stochastic publication function is given by

$$M_r(x) = E\{[y(x,e) - \mu(x)]^r\} \;\forall\; r = 2, 3, \ldots \tag{2.20}$$

From equation (2.20) we have $M_2(x) = Var(x)$, $M_3(x)$, and $M_4(x)$ respectively as the conditional variance, skewness, and kurtosis of the faculty member's stochastic publication function.

The skewness $M_3(x)$ for example, provides a measure of symmetry of the distribution and so can be adjusted to accommodate various forms of risk aversion. To make this approach empirically tractable, the following two specifications can be used:

$$y = \mu(x) + u \;\forall\; r = 1 \tag{2.21}$$

from which $[y - \mu(x)] = u$. Raising both sides of the equality to the r-th power gives the second specification as

$$[y - \mu(x)]^r \equiv u^r = M_r(x) + v_r, \;\forall\; r = 2, 3, \ldots \tag{2.22}$$

where $E(u) = E(v_r) = 0$ and $Var(u) = M_2(x)$ while

$$Var(v_r) \equiv E[u^r - M_r]^2 = E(u^{2r}) + M_r^2 - 2E(u^r)M_r = M_{2r} - M_r^2.$$

As specified, equations (2.21) and (2.22) are standard regression models that can be implemented in a non-parametric or semi-parametric estimation framework. One can also specify a parametric form for $\mu(x)$ and $M_r(x)$ and use generalized method of moments estimation(GMM), weighted least squares (WLS), or ordinary least squares with heteroskedasticity consistent covariance matrix estimator (OLS with HCCME) for estimation. All these are consistent while accounting for the non-constant variance (heteroskedasticity).

2.5 Faculty Decisions Under Multiple Journal Setting

The analysis of faculty publication decisions focused only on a single journal setting, allowing for the specification of a homogeneous research output with a unique but uncertain value. Now, the publication decision is allowed to be targeted at more than one journal, to account for publication diversification strategies that faculty might develop in order to reduce risk. To this end, consider faculty producing research output targeted at m different journals. Let $y = (y_1, \ \ldots, y_m)'$ be the publication vector with corresponding monetary value $p = (p_1, \ \ldots, p_m)'$. Under uncertainty in the monetary value of publication, due to publication lags, the monetary value of research output is not known at the time research decisions are being made. Let the random monetary value of publication be given by $p_i = \mu_i + \sigma_i e_i$, with $E(e_i) = 0$ for $i = 1, \ \ldots m$. Then under the expected utility model, a faculty member's risk preferences can be represented by the utility function $U(w + p'y - C(v, y))$, where w is the faculty member's base salary; $p'y = \sum_{i=1}^{m} p_i y_i$ is the income generated from publications; and $C(v, y)$ is the faculty member's cost of producing research output. Therefore, faculty publication decision are made in a way consistent with the maximization problem

$$Max_y[EU(w + p'y - C(v, y)],$$

where the expectation operator E is over the subjective probability distribution of the random vector p. Let y^* denote the optimal research output, the properties of y^* derived under the single journal setting are difficult to obtain here. The reason is that they now depend on both the joint probability distribution of the vector $p = (p_1, \ \ldots, p_m)'$ and on the multidimensional stochastic publication technology. Since such effects are difficult to predict in general, attention is focused on the slightly less general specification, in the form of the mean-variance model.

2.5.1 A mean-variance analysis

Consider a faculty member producing research output targeted at m different journals, with $y = (y_1, \ \ldots, y_m)^{'}$ representing the publication vector. Because each research output y_i has its corresponding monetary value p_i and cost v_i in terms of inputs used, the net return per unit from publishing in the i-th journal is $\dot{p}_i = p_i - v_i$, such that the vector of net return per unit of publication is $\dot{p} = (\dot{p}_1, \ \ldots, \dot{p}_m)^{'}$. A junior faculty member's net monetary benefit from publication is then given by $\tau = \dot{p}^{'} y = \sum_{i=1}^{m} \dot{p}_i y_i$. Because of the uncertainty in the monetary value of publication p, the net returns per unit of publications $\dot{p}_i = p_i - v_i$ is also random. Denote the mean vector of the later random vector by $\mu = (\mu_1 \ , \ldots, \mu_m)^{'} = E(\dot{p})$ and the variance of $\dot{p}$ by the $(m \times m)$ positive semi-definite matrix A shown below. The $\sigma_{ii} = Var(\dot{p}_i)$ and $\sigma_{ij} = Var(\dot{p}_i, \dot{p}_j)$ represent respectively the variance of $\dot{p}_i$ and the covariance between $\dot{p}_i$ and $\dot{p}_j$, with $i, j = 1, \ldots, m$.

$$A = Var(\dot{p}) = \begin{pmatrix} \sigma_{11} & \sigma_{11} & \ldots & \sigma_{1m} \\ \sigma_{12} & \sigma_{22} & \ldots & \sigma_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{1m} & \sigma_{2m} & \ldots & \sigma_{mm} \end{pmatrix}$$

.

In the mean variance framework, the faculty member's objective function is represented by the utility function $U[E(\tau), Var(\tau)]$, where $E(\tau) = E(\dot{p}^{'} \cdot y) = \mu^{'} y = \sum_{i=1}^{m} \mu_i y_i$ and $Var(\tau) = y^{'} A y = \sum_{i=1}^{m} \sum_{j=1}^{m} (y_i y_j \sigma_{ij})$, such that research output decisions are consistent with the maximization problem

$$Max_y \{U[E(\tau), Var(\tau)] : \tau = \dot{p}^{'} y, y \in Y\}.$$

or

$$Max_y \{U[\mu^{'} y, y^{'} A y] : y \in Y\}. \tag{2.23}$$

Y represents the feasible set for the research output vector y. Since the junior scholar is assumed to be risk averse, $\partial U/\partial E(\tau) > 0$ and $\partial U/\partial Var(\tau) < 0$. That is the faculty member's utility is increasing in the mean net monetary benefit from publication, while decreasing in the variance of the monetary value. Letting y^* denote the junior scholar's optimal research output choices, the properties of y^* are now investigated using the expected value-variance frontier paradigm.

The E-V frontier

An optimal allocation of resources across risky alternatives was considered by Markowitz (1952) whose solution was to find the set of allocations that maximize expected total return for different levels of the variance of total return. This is called the "expected value-variance efficient" set or **E-V frontier**. The E-V frontier allows for the decomposition of the mean-variance problem into two stages:

stage 1: This stage considers a junior faculty member's choice of research output y holding the expected return from publication $E(\tau) = \mu^{'}y$ constant at some level L:

$$Z(L) = Min_y[y^{'}Ay : \mu^{'}y = L, y \in Y]. \tag{2.24}$$

The indirect objective function is $Z(L) = y^+(L)^{'}Ay^+(L)$, provides the smallest possible variance attainable for given levels of expected publication return L, where $y^+(L)$ represents the solution to the optimization problem in equation (2.24) for given levels of expected return from publication L. The indirect objective function $Z(L)$ is called the "expected value-variance" (E-V) frontier and is the boundary of the feasible region in the mean-variance space.

Under risk aversion, a faculty member's utility maximizing behavior always implies the choice of a point on the E-V frontier. Risk aversion implies that $\partial U/\partial Var(\tau) < 0$, therefore, for any given expected publication return L, a faculty member would always prefer a reduction in return variability up to a point on the E-V frontier. Figure (2.5) illustrates this and shows that a point like A is feasible but generates a high return

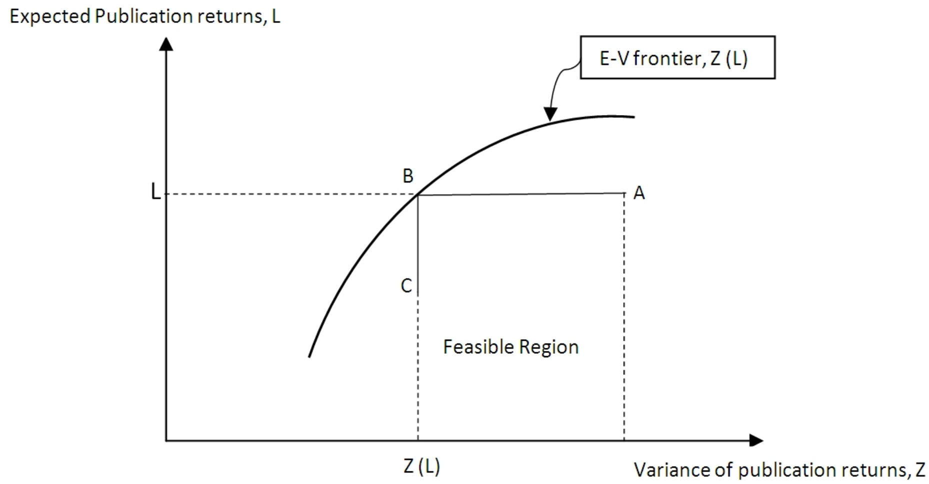

Figure 2.5: The E-V frontier

variance. Therefore holding expected publication return constant, from point A, a feasible reduction in return variability is always possible and improves the risk-averse faculty member's welfare.

A faculty member obtains the largest feasible return variability reduction by moving from point A to point B, which is located on the E-V frontier. Alternatively, considering the choice of expected monetary return from publication for a given risk exposure, with $\partial U/\partial Var(\tau) > 0$, a risk-averse faculty member will always choose a higher mean return up to a point on the E-V frontier. This is also shown in Figure (2.5), where the junior scholar can improve his/her welfare by moving from point C, which has a low expected publication return, to point B on the E-V frontier.

Stage 2: This stage considers choosing the optimal value for L, the expected publication return, which is fixed in the first stage. Consider the following optimization problem:

$$Max_L U(L, Z(L)) \tag{2.25}$$

Let L^* be the solution, then under differentiability and using the envelope theorem the first-order necessary condition is

$$\partial U/\partial L + (\partial U/\partial Z)(\partial Z/\partial L) = 0,$$

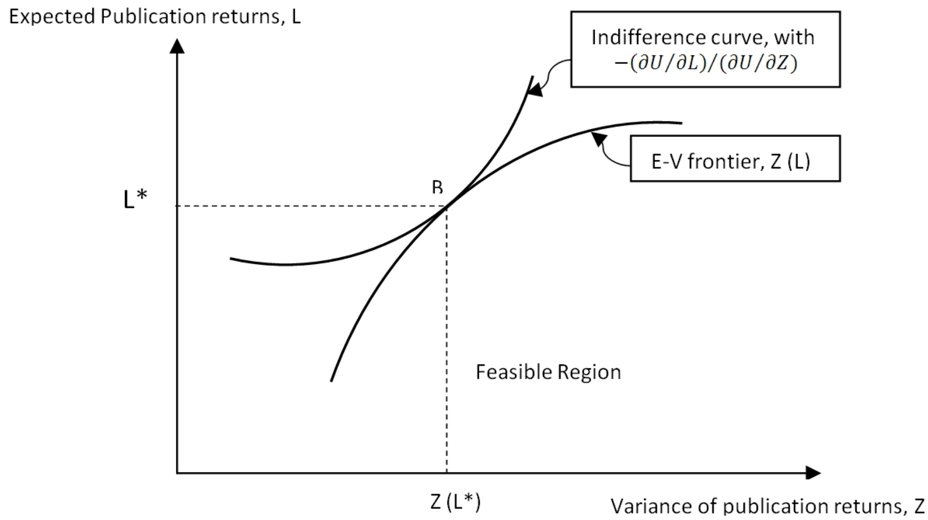

Figure 2.6: The E-V frontier with indifference curve

or

$$(\partial Z/\partial L) = -(\partial U/\partial L)/(\partial U/\partial Z).$$

Differentiating equation (2.25) with respect to the expected publication return L, gives the first order condition. It shows that, at the optimum, the slope of the E-V frontier, $\partial Z/\partial \mu$, is equal to the marginal rate of substitution between mean and variance of expected return to publication, $-(\partial U/\partial L)/(\partial U/\partial Z)$. This marginal rate of substitution is also the slope of the indifference curve between mean and variance of publication return.

Figure (2.6) illustrates this and shows that putting stage 1 and stage 2 together is always consistent with the original utility maximization problem, since $y^{+}(L)$ corresponds to the point on the E-V frontier where expected monetary return from publication is equal to L. This leads to the following result:

Result 6. Under the mean-variance analysis, with uncertainty in the monetary value of publication, when faculty can publish in more than one journal, a faculty member's optimal research output choice, y^{*}, is always the point on the E-V frontier corresponding to the optimal expected return from publication L^{*}.

Diversification

The mean-variance model also provides useful insights into faculty diversification strategies. To illustrate how, consider the simple case of $m = 2$ journals. In such case, $y = (y_1, y_2)$ is the vector of faculty research output, $\dot{p}_i$ represents the net return per unit of publication in the i-th journal, with $i = 1, 2$. Then, $\tau = \dot{p}_1 y_1 + \dot{p}_2 y_2$ is the net monetary benefit from publication. Let the mean return from the i-th journal be $\mu_i = E(\dot{p}_i)$ and its corresponding variance be $\sigma_i^2 = Var(\dot{p}_i)$. The correlation coefficient between the returns $\dot{p}_1$ and $\dot{p}_2$ can be represented by ρ with $-1 \leq \rho \leq +1$. Therefore, the mean and variance of the expected net monetary return from publication are given respectively by

$$E(\tau) = \mu_1 y_1 + \mu_2 y_2$$

and

$$Var(\tau) = \sigma_1^2 y_1^2 + \sigma_2^2 y_2^2 + 2\rho\sigma_1\sigma_2 y_1 y_2.$$

The stage-one optimization is of the form:

$$Z(L) = Min_y[\sigma_1^2 y_1^2 + \sigma_2^2 y_2^2 + 2\rho\sigma_1\sigma_2 y_1 y_2 : \mu_1 y_1 + \mu_2 y_2 = L, y \in Y]$$

or

$$Z(L) = Min_y[\sigma_1^2 y_1^2 + \sigma_2^2 (L - \mu_1 y_1)^2/(\mu_2)^2 + 2\rho\sigma_1\sigma_2 y_1 (L - \mu_1 y_1)/\mu_2 : y \in Y]. \quad (2.26)$$

First, consider the extreme case where there is perfect positive correlation between the returns $\dot{p}_1$ and $\dot{p}_2$. With $\rho = +1$, equation (2.26) can be rewritten as

$$Z(L) = Min_y[(\sigma_1 y_1 + \sigma_2 (L - \mu_1 y_1)/\mu_2)^2 : y \in Y] \quad (2.27)$$

which implies that $[Var(\tau)]^{\frac{1}{2}} = \sigma_1 y_1 + \sigma_2 (L - \mu_1 y_1)/\mu_2$; the standard deviation of the net monetary benefit τ from publication is a linear function of the publications in the

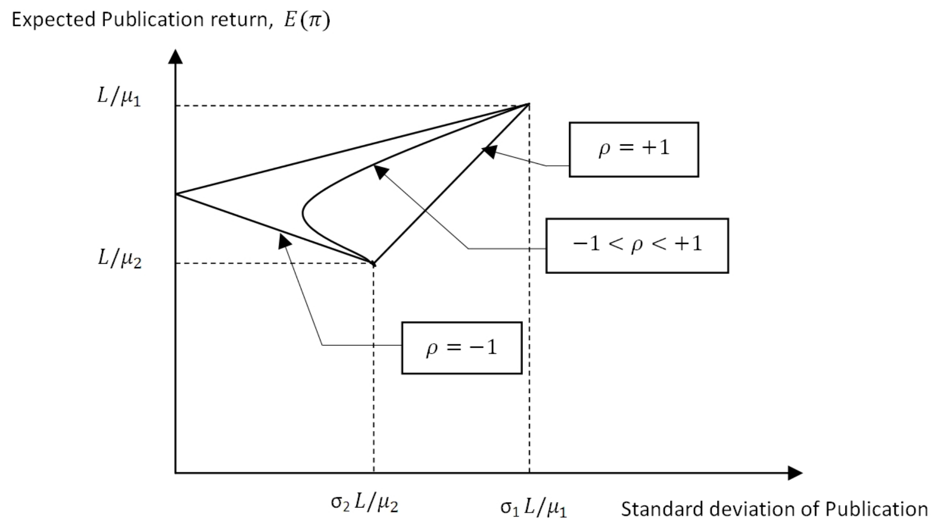

Figure 2.7: Faculty risk diversification

first journal y_1 and provides no possibility for diversification to reduce the variance of publication return.

The second extreme case is obtained when $\rho = -1$, suggesting a perfect negative correlation between the per unit returns from publication. Similarly, equation (2.26) can be rewritten as:

$$Z(L) = Min_y[(\sigma_1 y_1 - \sigma_2(L - \mu_1 y_1)/\mu_2)^2 : y \in Y] \tag{2.28}$$

Choosing $y_1 = L\sigma_2/(\mu_2\sigma_1 + \mu_1\sigma_2)$ implies $Var(\tau) = 0$, which means that there is a strategy that eliminates risk altogether. This suggest that $\rho = -1$ provides the greatest possibility to reduce risk exposure. Faculty in this case can better diversify by choosing two journals with opposite returns potential (for example one with a very high return potential and the other one with a relatively low return potential).

Now, consider the intermediate cases with $-1 < \rho < +1$, where the possibilities of diversification and risk reduction decrease with the correlation coefficient ρ between the two returns $\dot{p}_1$, and $\dot{p}_2$. This is illustrated in figure (2.7) which shows the trade off between expected net return to publication and the standard deviation of the net return under alternative correlation coefficients. This figure further suggests that diversification strategies cannot help a faculty member reduce risk exposure as much when the targeted journals have a strong positive correlation in their per unit return. In such situation, a faculty member's least risky strategy will be to target the journal with the least risk exposure. On the other hand, risk exposure can be greatly reduced, when faculty can target publications at journals with negative correlation in their per unit returns to publication.

Result 7. Under the mean-variance analysis, with uncertainty in the monetary value of publication,when faculty can diversify, then risk and risk aversion provide economic incentives for a faculty member to target journals that are negatively correlated in their per unit returns.

This result suggests that although a faculty member will target journals with potentially higher returns, i.e., the top tier journals in one's field, the faculty member will also consider publications in journals with a lower potential return. This is because acceptances in higher tier publications are constrained by the faculty member's ability to meet higher quality standards. Therefore a risk-averse faculty member may find it strategic to target lower ranked journals to increase publication numbers which increases the chances of a favorable tenure decision.

2.6 Field-Wide Implications Of Risk

2.6.1 Faculty distribution across journal tiers

In order to understand the implications of risk for the distribution of publications across journal tiers, assume that the m potential journals that the scholar can target publication at in a given field are ranked by level of riskiness from the most risky to the least risky, say $(J_1, \ldots J_m)$, with the top tier journals being the most risky and low tier journals the least. Then, the faculty member's problem can now be considered as that of publishing in a given tier, since differently ranked journals provide different publication returns with higher level publications providing the greatest return.

Under such a setting, then faculty tend to target publications at journal tiers that offer the greatest reward. The existence of different publication tiers at which faculty can target publication, parallels that of the alternative occupations in Roy (1951) and makes the results of the publisher's self-selection process more varied, but with more generalizations. Top tier publications will be more infrequent and carry greater risk, while low tier publications occur more often and carry lower risk. Everyone in the faculty community is assumed to be capable of publishing papers in either tier, though the probability of publishing at the highest level may be near zero for some.

Whatever the returns from publications and the correlation ρ between returns from each pair of outlets, top tier journals will always attract the most able faculty, and virtually none of those less endowed. As shown in equation (2.27) and equation (2.28), with perfectly negative correlation, or perfectly positive but small correlation between pairs of returns from tiers publications, every tier will attract a high proportion of the most able faculty and the proportion will decrease as less and less able faculty, in so far as that particular tier is concerned, are considered. in the intermediate tiers, the proportion of faculty of different levels of competence will not increase or decrease as steadily with competence, as is the case in top and lower tier, but over some ranges of potential research output, the proportion of faculty will rise and over others it will fall. The general effect is a high variance in the distribution of faculty with ability to publish in top-tier journals (top-tier publishers) within the field and a relatively more concentrated distribution for faculty with ability to publish in relatively lower-tier journals (low-tier publishers).

2.6.2 Faculty distribution across departments

The distribution of faculty across departments having different rankings will closely follow that of the distribution of members across journal tiers. This is because differently ranked department will have different standards of publication requirements for tenure. Top ranked departments will tend to require publications from top tier journals, while a relatively lower ranked department will have lower publication standards. In such situations, faculty self-selection mechanism as described above will guarantee that faculty with ability to publish in top-tier journals self select into top ranked departments, while faculty with relatively lower-tier publication ability self select into relatively lower ranked departments. Therefore risk operates to ensure compatible faculty/department matching. That is, risk as introduced by the up-or-out rules from tenure-track contracts, sufficiently guarantee a viable matching between faculty and department. This is because risk averse faculty with knowledge of their personal ability self-select into departments where the requirements for promotion (in terms of publication standards) are such that faculty have a fair chance of getting tenure after the probationary period. "Fair chance" is used to signify the fact that tenure is still uncertain and contingent on the faculty member's scientific research output meeting departmental standards.

2.7 Discussion and Model Comparison

In relation to the principal agent modeling framework under information asymmetry, this paper's results in terms of faculty ex-ante behavior under the tenure system are consistent with those of Jing (2008); Chen and Lee (2009). Tenure track contracts with up-or-out rules significantly distorts junior faculty productivity incentives but also allow for viable matching between department and faculty. A distinction however needs to be made between the two results. In this analysis, although information about faculty ability is not known to the department/university, a department through its ranking and publication quality standards, indirectly send information about tenure requirements. Such information is taken into account by risk averse junior faculty when looking for matching departments.

In the principal agent model, tenure track as a labor market contract, acts as a screening device providing department with information on faculty ability, and so helps reduce the information rent for department, and as such tenure track only operates after tenure-track contract is signed. In the current analysis however, tenure track with up-or-out rules regulates the self-selection mechanism of faculty into departments, long before the tenure track contract is signed, by affecting faculty decision on which departments to seek tenure track contracts from, based on faculty subjectively perceived probability of survival in such departments. Once a choice is made and tenure-track contract is obtained then tenure-track rules continue to operate affecting faculty scientific productivity incentives. Furthermore our analysis suggests that the information gain as suggested by principle-agent models may not be quite as much on the individual faculty member's type relative to the whole faculty community in the field, but only on whether faculty is willing and able to respond to the incentives provided in order to keep the position within the department by the end of the probationary period.

2.8 Conclusion

The growing emphasis on publication as a requirement for tenure has modified junior faculty members' scientific productivity incentives by raising the uncertainty level faced by untenured faculty. This paper has attempted to describe the effect of such risk on junior faculty scientific research output decisions and the implications of risk in terms of not only faculty distribution across publication tiers, but also across differently ranked academic departments. It was shown that a risk averse faculty member in the presence of publication value uncertainty, publishes at a point where the expected value of publication exceeds its marginal cost. Also a faculty member's scientific productivity was shown to be stimulated by increases in base salary when decreasing absolute risk aversion (DARA) described the faculty member's preference, while increased uncertainty in the value of publication provided a lesser incentive for scientific research output production. Moreover, when publication decision can be targeted at different journals, risk averse faculty had the economic incentive to target journals with negatively correlated per-unit returns.

In terms of academic field wide effects, risk was shown to affect the distribution of faculty across journal tiers with the few most able faculty targeting publications at top-tier journals and most of the faculty community concentrated in the relatively lower tiers. This has implications for the observed increase in the volume of uncited published papers, which Mohamed (2004) suggested was due to the growing publish-or-perish emphasis of many academic institutions. This emphasis seems to have shifted faculty priorities away from risky quality publications in top-tier journals and toward relatively sure quantity publication in low-tier journals. Because top-tier journals are the ones more widely read and cited, this leaves a mass of uncited papers in low-tier publication outlets. Furthermore, risk was shown to affect faculty distribution across differently ranked academic departments, guaranteeing then viable matching between departments and faculty through its effect on faculty self selection mechanism. The most able faculty were shown to self select into top-ranked departments with high quality publication standards, while the less able faculty self selected into relatively lower ranked departments.

Finally, it must be emphasized that this analysis was purely positive and intended to describe the effect of risk on junior faculty publication decisions, as faculty themselves perceive it. Nevertheless, a fairly detailed examination of this sort seems worth while to illuminate such a familiar and commonplace phenomena from a rather different angle, using already established tools from probability theory and the economics of risk.

CHAPTER 3

TESTING FOR RATIONALITY IN HEALTH INSURANCE CHOICE BY ADULTS IN THE U.S.: A PANEL-LIKE ERROR COMPONENTS MIXED LOGIT APPROACH

3.1 Introduction

Rational choice theory stipulates that individuals act as if balancing costs against benefits to arrive at actions that maximize personal well being. As such, patterns of behavior in society reflect the choices made by individuals as they try to maximize benefits and minimize costs. In economic decision making, a rational agent is one that has full information, and acts only if the marginal benefits of the action exceed its marginal costs. Economic models relying on rational choice theory often adopt the assumption that economic situations or collective behaviors are the result of individual actions alone, as they choose the best action according to unchanging and stable preference functions and constraints facing them.

This theoretical vision of rational choice theory has been subject to increasing criticism in the literature because of its failure to account for certain types of behavioral patterns expressed by individuals (Fernandez-Huerga, 2008; Schram and Caterino, 2006). These criticism have led to the development of the concept of bounded rationality, which explicitly recognizes the limited nature of information and the difficulties people have in processing information. Although viewed as limited, the rational choice approach remains an important paradigm because, assuming humans make decisions in a rational, rather than stochastic manner, implies that their behavior can be modeled; which further implies that predictions can be made about future actions. In addition, the mathematical formality of rational choice theory models allow economists and other social scientists to derive results from their models that may have otherwise not been observed, and to submit these theoretical results for empirical scrutiny.

In fact, the aim of this paper is to test the rational choice paradigm in relation to discrete choice modeling using a comparative statics result from an adapted version of the static economic model of health investments and health outcomes by Strauss and Thomas (2007, pp.3380-3385), which relies on the assumption that economic agents behave rationally in their choice of health inputs. Therefore, focusing on health insurance as a health input, two hypotheses are formulated from the model and tested using data from the 2007 Medical Expenditure Panel Survey (MEPS). The question this paper attempts to answer is:

Given the sample evidence from the 2007 MEPS, are individual adults in the U.S. rational in how they make their health insurance choices?

The MEPS is a nationally representative sample of the U.S. adult population that allows us to test for the adequacy of the rationality assumption in the context of health insurance choice. A respondent initially states her/his preference for health insurance by expressing her/his attitude towards health insurance cost worthiness as "uncertain" or "worthy" or "not worthy" in the Self Administered Questionnaire (SAQ) of the MEPS. In the last round, the actual revealed health insurance status by the individual over the scope of the panel is recorded as either "uninsured," "private" or "public." The idea is that if rationality holds, then on average, agents' revealed preferences for health insurance will be consistent with their stated preferences.

Therefore, it is assumed that the 2007 MEPS sample is generated through the optimizing behavior of rational agents or, in other words that rationality is the guiding mechanism for the data generating process. If so the stated attitude toward health insurance cost worthiness variable can be seen as synthesizing the outcome of the margin principle for each respondent in the MEPS. More specifically, an individual, based on her subjective standards (one's preferences, self-knowledge, circumstances, and costs and benefits of being insured), internally compares the marginal benefit of being a health insurance beneficiary to its marginal cost, and then states whether health insurance is worth its cost or not. If the subjectively perceived marginal benefit exceeds the marginal cost, then the respondent expresses health insurance to be "worthy." If the subjectively perceived marginal cost exceeds the marginal benefit then the respondent expresses health insurance to be "not worthy". The respondent may also be indifferent and respond "uncertain."

Given this assumption about the data generating process, which suggests that agents behave rationally, we would expect some degree of consistency between the expressed attitude towards health insurance cost worthiness in early rounds, and the actual revealed choice of health insurance recorded after the last round of the MEPS, if the rationality assumption is appropriate.

The findings in fact reveal that on average, relative to being uncertain, the individuals who think that health insurance is not worth its cost are less likely to be privately or publicly insured, relatively to being uninsured. In addition, relative to being uncertain, the individuals expressing health insurance to be worth its cost are more likely to be insured (privately or publicly) over being uninsured. These results validate the comparative statics predictions from the economic model, and the adequacy of rationality as a guiding mechanism for the 2007 MEPS data generating process. The results are also consistent with the past literature with respect to individuals' skepticism toward health insurance. In fact the findings suggest relatively more skepticism (dissatisfaction) toward public only coverage compared to having some private coverage.

The remaining of the analysis is therefore organized as follows, section 2 formally motivates the hypotheses to be tested, section 3 develops the econometric model which is a new variant of the Mixed-Logit model. Section 4 presents the data and analytical strategy used, section 5 describes the Bayesian Markov Chain Monte Carlo method of estimation implemented. Descriptive findings are presented in section 6, while the econometric results are described in section 7. Section 8 provides discussion and limitations, and finally section 9 concludes the analysis.

3.2 Motivation

Consider the following adaptation from the static model of health investments, and health outcomes by Strauss and Thomas (2007, pp. 3380-3385) with the individual static health production function represented as:[1]

$$H = H(N_1, N; S_H, \mu) \tag{3.1}$$

where H is an array of measured health outcomes that depends on health insurance N_1 and other health inputs N, which are assumed to be under the control of the agent. S_H is a vector of variables affecting the shape of the underlying health production function such as age, socio-demographic characteristics, family background and environmental factors. Measurement errors and the econometrician's limited knowledge about the agent, are captured by μ and represent unobserved characteristics.

Now assume the agent's welfare depends on labor supply, L, and consumption of goods and services, C, which may include health insurance, N_1. Then utility, U, depends on health outcomes, H, as well as a vector of variables, S_U of which S_H is a subset, that affect the shape of the utility function, and unobserved characteristics, ξ. The unobserved characteristics include heterogeneity in tastes that may be related to the unobserved characteristics affecting health production in equation (3.1) since

[1]The notation adopted in this section follows the standard notation presented by Strauss and Thomas (2007, pp. 3380-3385). The initial model is however extended to explicitly account for health insurance as an input in the health production function.

preferences may themselves depend on innate healthiness. Therefore utility can be expressed as:

$$U = U(C, L; H, S_U, \xi). \tag{3.2}$$

Because agents resource allocation decisions are constrained by their budget and time, the budget constraint can be represented as:

$$P_c C^* + P_{n_1} N_1 + P_n N^c = WL + V \tag{3.3}$$

where V and W represent respectively the individual's non-labor income and wage, C^* is the consumption not related to the health production with prices P_c. N_1 is the consumption of health insurance with price P_{n_1}, and N^c is the vector of other purchased health inputs with price vector P_n.

Furthermore assume a person's wage varies with the health output, H, the factors affecting labor productivity, S_w, and unobserved factors, α, such that the wage function is expressed as:

$$W = W(H; S_w, \alpha). \tag{3.4}$$

Finally, assuming the amount of labor supply (L) varies with health output (H) and, denoting the marginal utility of income as λ, then the individual's maximization problem is:

$$\begin{aligned} \max_{C,L} \quad & U(C, L; H, S_U, \xi) \\ \text{subject to} \quad & P_c C^* + P_{n_1} N_1 + P_n N^c = WL + V, \\ & H(N_1, N; S_H, \mu), \\ & W = W(H; S_w, \alpha), \\ & L = L(H). \end{aligned} \tag{3.5}$$

The Lagrangian of this maximization problem is:

$$\begin{aligned} \ell = {} & U\left(C, L; H(N_1, N; S_H, \mu), S_U, \xi\right) \\ & - \lambda \left[P_c C^* + P_{n_1} N_1 + P_n N^c - W(H(N_1, N; S_H, \mu); S_w, \alpha) \cdot L(H(N_1, N; S_H, \mu)) - V\right], \end{aligned} \tag{3.6}$$

Assuming an interior solution, then the first order condition with respect to health insurance consumption is

$$\frac{\partial U}{\partial C} \cdot \frac{\partial C}{\partial N_1} + \frac{\partial U}{\partial H} \cdot \frac{\partial H}{\partial N_1} = \lambda \left(P_{n_1} - L \left[\frac{\partial W}{\partial H} \cdot \frac{\partial H}{\partial N_1} \right] - W \left[\frac{\partial L}{\partial H} \cdot \frac{\partial H}{\partial N_1} \right] \right). \qquad (3.7)$$

The marginal utility of health insurance consumption $\frac{\partial U}{\partial N_1} = \frac{\partial U}{\partial C} \cdot \frac{\partial C}{\partial N_1}$ is zero if N_1 is not valued in consumption, in which case it will not be an element of C in equation (3.2). The right hand side of this first order condition suggests that if health insurance, N_1, raises wages or increases labor supply through improving health outcomes, H, then the nominal cost of health insurance decreases, leading to more health insurance consumption. In this formulation, the rate of change in the shadow cost of health insurance can also depend on the level of health, H. This is in fact evidenced in the biomedical literature by Haas and Brownly(2001), which suggests health outcomes are related nonlinearly to health inputs.

From the first order condition when $\frac{\partial U}{\partial C} \cdot \frac{\partial C}{\partial N_1} = 0$ then we have an inequality, which when rearranged can be written as:

$$\lambda > \frac{\frac{\partial U}{\partial H} \cdot \frac{\partial H}{\partial N_1}}{P_{n_1} - L[\frac{\partial W}{\partial H} \cdot \frac{\partial H}{\partial N_1}] - W[\frac{\partial L}{\partial H} \cdot \frac{\partial H}{\partial N_1}]} \qquad (3.8)$$

where the left hand side, λ, is the shadow cost of health insurance consumption [2], and the right hand side interpreted as the real marginal health benefit of being a health insurance beneficiary. The numerator provides the nominal marginal health

[2]Note the shadow cost of health insurance here is the real value of the resources that must be given up for each unit of health insurance consumption. For private coverage this cost includes direct premiums being paid. For public coverage, shifting of cost sharing burdens the beneficiaries. In 2007 income from social security accounted for roughly half, or more of the annual income for 80 percent of public health insurance beneficiaries, with the average social security benefit being consumed by cost-sharing for Medicare part B and part D premiums. Premiums collected from beneficiaries represented 40 percent of total out-of-pocket spending by beneficiaries, or 10 percent of per capita health care expenses in 2006. This figure represented 65.7 percent of average out-of-pocket spending by beneficiaries in 2009 (Cubanski et al., 2009)

benefit that is adjusted using the price measure in the denominator. Therefore the equation (3.8) suggests that for the representative agent, when the shadow cost of being a health insurance beneficiary is greater than its real marginal health benefit, the agent will forgo insurance, otherwise the agent will purchase a non null amount of health insurance.

Given this comparative static result, the following two hypotheses can be formulated in conjunction with the data generating process assumption made in the introduction for the adult respondents in the 2007 MEPS.

hypothesis1: For the respondents that state health insurance is worth its cost, we would expect on average the shadow cost of health insurance, λ, to be less than the real marginal benefit derived from being a health insurance beneficiary, such that a state of insurance is preferred over that of being uninsured.

hypothesis2: For the respondents that state health insurance is not worth the cost, on average, we would expect the shadow cost λ to be greater than the real marginal benefit derived from health insurance consumption, such that an uninsured state would be preferred over that of being insured. [3]

The above two hypotheses provide a testable link between expressed attitude towards health insurance cost worthiness, and revealed choice of health insurance. Validation of these two hypotheses using the MEPS dataset would imply consistency with the presented economic model's predictions, and hence that the rationality assumption (as discussed in the introduction) in the case of Health insurance choice by

[3]Note that the MEPS sample includes individuals with employer-sponsored health insurance. Respondents might seem to have no choice, however the evidence in the literature suggests that workers sorting across alternative employments reflect their tastes for health insurance, with individuals not valuing health insurance sorting into jobs that do not provide health insurance, and vice versa (Goldstein and Pauly, 1976; Feldman et al., 1997; Monheit and Vistnes, 1999, 2008). For public coverage, eligibility is necessary but not sufficient for coverage. Since enrollment is not automatic, one still has to sign up to be a beneficiary.

decision makers in the MEPS dataset is reasonable.

To test for the consistency of these two hypotheses, the expressed attitude toward health insurance cost worthiness, and the revealed choice of health insurance variables are jointly modeled, conditional on a set of covariates within the panel-like Error Components Mixed Logit framework developed in the next section. For an exposition of mixed logit as currently used in the literature see (Train, 2009, p. 134-150). The two jointly modeled variables are factors each with 3 levels. For the stated attitude variable, the levels are (uncertain, worthy, not worthy), while the insurance choice variable has levels (uninsured, Any private, public Only). The econometric framework is developed in the next section.

3.3 Panel-Like Error Components Mixed-Logit Specification

Mixed logit is a highly flexible behavioral economic model that can approximate any random utility model (McFadden and Train, 2000) [4]. The standard mixed logit model has choice probabilities that can be represented in the form:

$$P_{ij} = \int L_{ij}(\theta) f(\theta) d(\theta) \qquad j = 1, \ldots, m \tag{3.9}$$

where P_{ij} is the probability that individual i chooses alternative j among a set of m alternatives, and $L_{ij}(\theta)$ is the logit probability evaluated at θ:

$$L_{ij}(\theta) = \frac{e^{V_{ij}(\theta)}}{\sum_{k=1}^{m} e^{V_{ik}(\theta)}} \qquad j = 1, \ldots, m \tag{3.10}$$

$f(\theta)$ is a density function and $V_{ij}(\theta)$ represents the observed portion of utility which depends on the parameters θ. When utility is linear in θ, then $V_{ij}(\theta)$ can potentially be represented in two ways according to whether or not regressors vary across alternatives (Cameron and Trevedi, 2005, p. 500). $V_{ij}(\theta) = \theta_{ij}' W_i$ relies on individual specific

[4] The notation adopted in this section follows the standard notation for mixed logit discrete choice modeling, see Train (2009, p. 134-150)

variables W_i , that are constant over alternatives, leading to the multinomial logit representation. When the model relies exclusively on regressors that vary across alternatives, then $V_{ij}(\theta) = \theta_i' W_{ij}$ which leads to the conditional logit representation. The application of mixed logit in the literature has concerned itself exclusively with the later conditional logit representation (Brownstone and Train, 1999; David and Greene, 2003; Dean et al., 2009; Hess et al., 2006; King et al., 2007; Srinivasan and Mahamassani, 2005). This specification allows the analyst to capture decision makers' tastes as they relate to the characteristics of the choice options.

The main interest in the current analysis is to model decision makers' preferences, which under the rational agent paradigm are assumed to be stable across the set of choice alternatives. Preferences depend on the characteristics of the decision maker and her environment, rather than the characteristics of the choice options. This is assumption is necessary in this application since the MEPS dataset does not include information on the characteristics of the choices available to respondents. Therefore, the former representation of $V_{ij}(\theta) = \theta_{ij}' W_i$ is adopted. The multinomial logit representation is adopted here in an attempt to fill the void in the literature, but also because it relies on regressors that are fixed across choice alternatives, which allows one to model agents' stable preferences. With the multinomial logit representation, $V_{ij}(\theta) = \theta_{ij}' W_i$, an identification restriction is needed to ensure the logit formula $\frac{e^{\theta_{ij}' W_i}}{\sum_{k=1}^{m} e^{\theta_{ik}' W_i}}$ sums to one, over all m choice alternatives. The usual restriction of $\theta_{i1}' = \mathbf{0}$ is imposed (Cameron and Trevedi, 2005, p. 500), making the first alternative the base alternative, and the mixed logit probability represented as:

$$P_{ij} = \int \left(\frac{e^{\theta_{ij}' W_i}}{\sum_{k=1}^{m} e^{\theta_{ik}' W_i}} \right) f(\theta) d(\theta) \qquad j = 1, \ldots, m \tag{3.11}$$

which is a weighted average of the standard logit formula evaluated at different values of θ, with weights given by the mixing density function $f(\theta)$. The mixing distribution $f(\theta)$ can be discrete, with θ taking a finite set of distinct values. In most applications of standard mixed logit, the mixing density is specified to be continuous (King et al.,

2007) . Assuming normal mixing density, then $\theta \sim N(\beta, G)$ and the choice probability becomes:

$$P_{ij} = \int \left(\frac{e^{\theta'_{ij} W_i}}{\sum_{k=1}^{m} e^{\theta'_{ik} W_i}} \right) \phi(\theta|\beta, G) d(\theta) \qquad j = 1, \ldots, m \tag{3.12}$$

where $\phi(\theta|\beta, G)$ is the normal density with mean β, and covariance G. The task then is to estimate β, and covariance G.

The mixed logit probability can be derived from individuals' utility maximizing behavior in various ways. The most widely used approach is based on random coefficients (Dean et al., 2009; King et al., 2007), this analysis however relies on the error components approach following (Brownstone and Train, 1999). In this case, utility is specified as:

$$U_{ij} = \beta'_{ij} X_i + b'_i Z_i + \epsilon_{ij}, \tag{3.13}$$

where X_i and Z_i are vectors of observed variables for individual i, β_{ij} are fixed coefficients, b_i is a set of random terms with zero mean and covariance G, and ϵ_{ij} is iid extreme value. The terms in Z_i are error components which create correlations among the utilities for different alternatives, and along with ϵ_{ij}, define the stochastic portion of utility. That is, the random portion of utility is $\xi_{ij} = b'_i Z_i + \epsilon_{ij}$. Utility is correlated across choice alternative because $cov(\xi_{ij}, \xi_{ik}) = E(b'_i Z_i + \epsilon_{ij})(b'_i Z_i + \epsilon_{ik}) = Z'_i G Z_i$, where G is the covariance of b_i.

Letting $\theta'_{ij} = \langle \beta'_{ij}, b'_i \rangle$ and $W'_i = \langle X'_i, Z'_i \rangle$ then the utility function can be written as

$$U_{ij} = \theta'_{ij} W_i + \epsilon_{ij}, \tag{3.14}$$

Note that the identification restriction ($\theta'_{i1} = \mathbf{0}$) imposed on the logit formula translates into a one dimensional reduction of the utility function. This takes care of the fact that only utility differences matter to the decision maker when choosing among a set of alternatives (Train, 2009, p. 19-20). In this case utility differences are considered with respect to the first alternative. The decision maker knows the value of his own θ'_{ij} and ϵ_{ij} for all alternatives and chooses alternative j if and only if $U_{ij} > U_{ik}$ for all $j \neq k$. The researcher on the other hand observes the components of W_i, but not θ_{ij} or the ϵ_{ij}'s. If the elements of θ_{ij} were observed, then the choice probability would be the probability conditional on θ_{ij}:

$$L_{ij}(\theta_{ij}) = \frac{e^{\theta'_{ij}W_i}}{\sum_{k=1}^{m} e^{\theta'_{ik}W_i}} \tag{3.15}$$

because the θ_{ij} are unobserved by the researcher, the above conditional probability cannot be specified. Therefore the unconditional probability which averages over all θ_{ij} is specified as the integral of $L_{ij}(\theta_{ij})$ over all possible θ_{ij} :

$$P_{ij} = \int \left(\frac{e^{\theta'_{ij}W_i}}{\sum_{k=1}^{m} e^{\theta'_{ik}W_i}} \right) \phi(\theta) d(\theta) \tag{3.16}$$

which is the mixed logit probability with normal weighting density as represented in equation (3.12).

To allow for repeated choices by each sampled decision maker, a panel-like error components representation of the utility function is used. This is achieved by defining the utility of individual i in choice situation t choosing alternative j as:

$$U_{ijt} = \beta_{ijt} X_i + b_i Z_i + \epsilon_{ijt} \quad \forall \;\; j = 1, 2, 3; \quad t = 1, 2 \tag{3.17}$$

with ϵ_{ijt} being iid extreme value with zero mean and a given variance over alternatives, choice situations, and decision makers. In the current application, each adult respondent can be seen as making a choice among three mutually exclusive alternatives in each of the two choice situations representing the stated preference and revealed preference for health insurance. The sequence of chosen alternatives, one for each choice situation is $\mathbf{c} = \{c_1, c_2\}$. Conditional on θ the probability that the individual makes this sequence of choices is the product of the logit formulas:

$$\mathbf{L}_{i\mathbf{c}} = \prod_{t=1}^{2} \left(\frac{e^{\theta'_{ic_t}W_i}}{\sum_{k=1}^{3} e^{\theta'_{ik_t}W_i}} \right) \tag{3.18}$$

since the ϵ_{ijt} are iid over choice situations. The unconditional probability of the sequence of choice, with the normal mixing density is then obtained by integrating $\mathbf{L}_{i\mathbf{c}}$ over all values of θ:

$$P_{i\mathbf{c}} = \int \prod_{t=1}^{2} \left(\frac{e^{\theta'_{ic_t} W_i}}{\sum_{k=1}^{3} e^{\theta'_{ik_t} W_i}} \right) \phi(\theta) d(\theta) \tag{3.19}$$

The only distinction introduced by the panel-like representation for repeated choices over the single choice case, is that the integral now involves a product of logit formulas, rather than just one single logit formula. These choice probabilities can be estimated through simulation from a classical stand point, or using Bayesian Markov Chain Monte Carlo methods. Because of the multidimensionality of the choice probabilities, and also the intractability in integrating over the random effects (McCulloch and Searle, 2001), we refer to Bayesian Markov Chain Monte Carlo (MCMC) methods to estimate the model. MCMC methods provide an alternative strategy for marginalizing the random effects that may be more robust than the techniques used to approximate the integrals (Zhao et al., 2006; Brown and Draper, 2006).

The panel-like error components mixed logit model as described here overcomes the independence of irrelevant alternative (IIA) restriction that characterizes the standard multinomial logit.[5] This is because the denominators of the logit formulas are inside the integrals and do not cancel when the ratio of mixed logit probabilities are taken. This ratio $\frac{P_{ij}}{P_{ik}}$ still depends on all the data not just on the information pertaining to the two considered alternatives. The model allows for the correlation of utility across choice situations, choice alternatives, and individuals. For any two individuals (ind_1, ind_2) for example, the correlations of utility across individuals, choice situations (t_1, t_2) and alternatives (j_1, j_2) can be graphically represented as shown in figure(3.1).

The model also captures individuals' preference rankings over various choice al-

[5]see Train (2009, p. 45-50) for a full exposition of the IIA property.

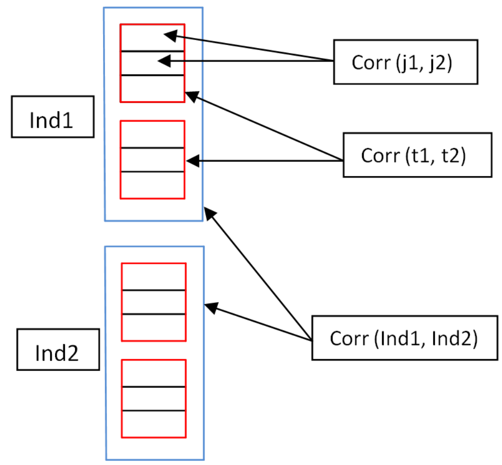

Figure 3.1: Levels of utility correlations in the model. utility correlation across choice alternatives, corr(j_1, j_2) comes from the assumption of stable preference across a set of mutually exclusive alternatives, by rational choice theory. Utility correlation across choice situations, corr(t_1, t_2) is also from rational choice theory, and suggests preference stability (or choice consistency) across choice situations. Utility correlations across individuals, corr(Ind_1, Ind_2) comes from the Multilevel-stratified structure of the MEPS dataset, creating preference interactions across decision makers.

ternatives, since for each individual, a unique fixed effect β is estimated for each alternative j in choice situation t, conditional on each regressor. Variations across choice situations of the estimated fixed effects for a given decision maker and regressor, allow for the possibility of preference reversal, or choice inconsistency across choice situations.

This possibility of choice inconsistency across choice situations is what allow us to use this model to test the above two hypotheses, by investigating the consistency between stated attitude toward health insurance cost worthiness and the revealed choice of health insurance by the adult respondents in the MEPS.

3.4 Data and Analytical Strategy

The empirical analysis is based upon data from the 2007 Medical Expenditure Panel Survey (MEPS) full year population characteristics data. The survey is sponsored by the Agency for Health Care Research and Quality(AHRQ), and designed to overlap two calendar years with a new Panel of sample households selected each year. The household component of the MEPS collects data from a subsample of the National Health Interview Survey and uses stratified and clustered random sampling with weights that produce nationally representative estimates for a wide range of health-related demographic and socioeconomic characteristics for the civilian, non-institutionalized U.S. population. The data from the calendar year 2007 was collected in rounds 1, 2, and 3 for MEPS panel 12 and rounds 3, 4, and 5 for MEPS panel 11.

The survey includes questions on respondents' attitudes toward health insurance and health insurance cost, in a self-administered questionnaire (SAQ) which was administrated in round 2 for panel 12 and round 4 for panel 11. Although the 2007 MEPS includes 30964 individuals, interviewed over the 2-year period, the target population for the SAQ only include adults (person age 18 or older) in the civilian non institutionalized population amounting to 19067 respondents. After accounting for questionnaire non response, the final sample used in this analysis is comprised of 18045 individuals 18 to 85 that were member of the civilian, non-institutionalized portion of the U.S. population in 2007. For more information on the MEPS sampling design, see Ezzati-Rice et al. (2008).

The dependent variables in this study are of two kinds, individuals' attitude towards health insurance cost worthiness "ATTHICW," and individual's health insurance coverage indicator "INSURANCE." The variable ATTHICW is constructed from the variable ADINSB42 provided in the MEPS dataset which is a factor with 5 levels (1. Disagree strongly, 2. Disagree somewhat, 3. Uncertain, 4. Agree somewhat, 5. Agree strongly), relating to the statement: Health insurance is not worth the cost. The variable is recoded into ATTHICW as a factor of 3 levels (1– Uncertain, 2– Agree, 3–Disagree), which is interpreted as (1– Uncertain, 2– Not worthy, 3–worthy) and represents the stated attitude towards health insurance cost. The second dependent variable INSURANCE is constructed from INSCOV07 provided in the MEPS dataset which summarizes health insurance coverage for each respondent.[6]

Because we wish to use "uninsured" as the base category in the estimation, the INSURANCE variable is constructed as a factor with three levels (1. Uninsured, 2. Private, 3.Public). Hence both dependent variables are factors each with three mutually exclusive and exhaustive categories.

Interests here centers on a postulated causal influence from the attributes and environment of individual respondents to their responses. The alternative invariant covariates on which the joint distribution of the two dependent variables is conditioned, include demographic characteristics such as AGE, SEX, EDUCATION, INCOME; health characteristics; and the respondent's marital status. Definitions and summary statistics for the independent variables are given in table (3.1).

To ensure valid inferences about the entire adult population in the U.S., care needs to be taken to account for the stratified design structure of the MEPS dataset. The variables VARSTR and VARPSU provided in the MEPS dataset are used to identify the sample strata and primary sampling units (PSU) required for appropriate variance estimation. For the 2007 MEPS full year file, there are 328 variance strata, with either two to three variance estimation PSUs per stratum. So in addition to the above fixed effects variables, the two variables VARSTR and VARPSU are introduced into the model to capture random stratum and PSU effects, respectively. The corresponding utility representation leading to the observed sequence of choice outcomes $\forall\ i\ =$

[6]INSCOV07 assumes three possible values: 1 = ANY PRIVATE(respondent had any private insurance coverage providing at minimum benefits for hospital and physician services[including TRICARE and MEDIGAP] any time during 2007), 2 = PUBLIC ONLY (respondent reported coverage only under MEDICARE, MEDICAID,or SCHIP, or other public hospital/physician programs during 2007), 3 = UNINSURED (respondents was uninsured during all of 2007). Note that these three categories are mutually exclusive, and respondents with both private insurance/TRICARE and public insurance are coded as "1"

$1, \cdots, N;\ j = 1, \cdots, 3;\ t = 1, 2;\ k = 1, \cdots, 328;\ l = 1, \cdots, 3;$ is given by:

$$U_{klijt} = \mu_{ijt} + S_k + PSU_{lk} + USU_{ilk} + \epsilon_{klijt} \tag{3.20}$$

where U_{klijt} is the utility from alternative j in choice situation t by individual i in primary sampling unit l within strata k. This utility is a linear function of a fixed effect μ_{ijt}, a random strata effect S_k, a random primary sampling unit effect PSU_{lk}, a random individual effect USU_{ilk}, and a disturbance term ϵ_{klijt}. Recalling the one dimensional identification restriction in each choice situation, and abstracting from the nesting indexes l and k, then for each individual i, we have a $((3-1)+(3-1)) = 4$-dimensional vector of relative latent utilities U_i. Note that the identification restriction imposed reduces the number of alternatives in each choice situation by one for each individual. This normalization ensures that the choice probabilities sum to one in each choice situation, and sets the normalizing alternative, which here is the first alternative in each choice situation, as the base alternative. The above equation(3.20) is equivalent to the Panel-like Error Components mixed logit representation of the utility function in equation (3.17) where $\beta_{ijt}X_i = \mu_{ijt},\ b_iZ_i = S_k + PSU_{lk} + USU_{ilk}$ and $\epsilon_{ijt} = \epsilon_{klijt}$.

The random components S_k, PSU_{lk}, USU_{ilk} are distributed with zero means, and variances represented respectively by V_s for the between stratum variations, V_{s_p} for primary sampling units variations within strata, and $V_{s_{p_i}}$ for the between individuals variations within primary sampling units within strata. Unlike V_s and V_{s_p} which are scalar variances, $V_{s_{p_i}}$ is a covariance matrix that represents variations across choice alternatives and choice situations between individuals as shown in equation (3.21) [7]. After accounting for these three sources of variations, all remaining variations are

[7] In the representation of $V_{s_{p_i}}$, $p_r = Private,\ p_u = Public,\ w = worthy,$ and $n_w = notworthy.$ $\sigma_{p_rp_r}$: utility variability between individuals choosing private coverage over being uninsured; σ_{ww}: utility variability between individuals with "worthy" preference over "uncertain"; $\sigma_{p_rp_u}$: utility covariation between individuals with private coverage and individuals with public coverage, both

standard *iid* and captured by the disturbance term ϵ_{klijt}.

$$V_{s_{p_i}} = \begin{bmatrix} \sigma_{p_r p_r} & \sigma_{p_r p_u} & \sigma_{p_r n_w} & \sigma_{p_r w} \\ . & \sigma_{p_u p_u} & \sigma_{p_u n_w} & \sigma_{p_u w} \\ . & . & \sigma_{n_w n_w} & \sigma_{n_w w} \\ . & . & . & \sigma_{ww} \end{bmatrix} \tag{3.21}$$

Using this covariance matrix $V_{s_{p_i}}$, we check for the consistency of the two hypotheses, by observing the signs and significance of the covariance coefficients, with special interest on the coefficients between choice alternatives across choice situations. If the first hypothesis is valid, we will expect a positive and significant covariation between the "worthy" attitude and any of the "private" and "public" insurance options. These two covariance coefficients are represented by $\sigma_{p_r w}$ and $\sigma_{p_u w}$. If the second hypothesis is also valid, we will expect a negative and significant covariance coefficient between the "not worthy" attitude and any of the two "private" and "public" insurance options. The covariance coefficients in this case are represented by $\sigma_{p_r n_w}$ and $\sigma_{p_u n_w}$ respectively.

Because of the multidimensionality of the choice probabilities, and also the intractability in integrating over the random effects (McCulloch and Searle, 2001), we use Bayesian Markov Chain Monte Carlo (MCMC) methods to estimate the model. MCMC methods provide an alternative strategy for marginalizing the random effects that may be more robust than the techniques used to approximate the integrals (Zhao et al., 2006; Brown and Draper, 2006).

3.5 MCMC Sampling Schemes for Model Parameters

As described in equation (3.17) and (3.20), the panel-like error components mixed logit model, is just a spacial case of the more general class of Generalized Linear

relative to being uninsured $\sigma_{p_r w}$: utility covariation between individuals with private coverage and individuals with the "worthy" preference

Mixed Models (GLMMs), with a categorical outcome variable and a specified logit link function. Therefore, following the standard multivariate notation adopted for GLMMs (Hadfield, 2010), the standardized 4-dimensional vector of latent utilities U_i for each individual i in equation (3.17) can be stacked into a single column vector across all N individuals in the sample. In this form, we obtain a $4N$-dimensional latent vector $\mathbf{U}$ of utilities for the whole sample of respondents with

$$\mathbf{U} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\mathbf{u} + \mathbf{e} \tag{3.22}$$

where $\mathbf{X}$ is a design matrix relating the fixed predictors to the data, and $\mathbf{Z}$ is a design matrix relating random predictors to the data. These predictors have associated parameter vector $\boldsymbol{\beta} \sim N(\boldsymbol{\beta}_0, \mathbf{B})$, and $\mathbf{u} \sim N(\mathbf{0}, \mathbf{G})$. The residuals vector is represented by $\mathbf{e} \sim N(\mathbf{0}, \mathbf{R})$. In this formulation $\mathbf{B}$, $\mathbf{G}$ and $\mathbf{R}$ are the expected (co)variance matrices of the fixed effects, random effects and residuals, respectively. They are typically unknown, and must be estimated from the data. Recalling that in a Bayesian analysis no distinction is made between fixed and random effects, as all effects are considered random, we can combine the design matrices ($\mathbf{W} = [\mathbf{X}, \mathbf{Z}]$) and also the parameters ($\boldsymbol{\theta} = [\boldsymbol{\beta}^{'}, \mathbf{u}^{'}]$), and rewrite equation (3.22) as:

$$\mathbf{U} = \mathbf{W}\boldsymbol{\theta} + \mathbf{e} \tag{3.23}$$

The prior distribution for the location effects $\boldsymbol{\theta}$ is multivariate normal, with the zero off-diagonal implying a priory independence between fixed effects and random effects.

$$\boldsymbol{\theta} = \begin{bmatrix} \boldsymbol{\beta} \\ \mathbf{u} \end{bmatrix} \sim N\left(\begin{bmatrix} \boldsymbol{\beta}_0 \\ \mathbf{0} \end{bmatrix}, \begin{bmatrix} \mathbf{B} & \mathbf{0} \\ \mathbf{0} & \mathbf{G} \end{bmatrix}\right) \tag{3.24}$$

The goal of the analysis is to estimate $\boldsymbol{\theta}$. The prior for $\boldsymbol{\theta}$ can be Gibbs sampled in a single block using the method of Gracia-Cortes and Sorensen (2001) as explained below. With conjugate priors, the variance structures ($\mathbf{R}$ and $\mathbf{G}$) follow an inverse-Wishart distribution which can also be Gibbs sampled in a single block. The variance structures ($\mathbf{R}$ and $\mathbf{G}$) for the model in equation(3.22) are represented as:

$$\mathbf{G} = \begin{bmatrix} \mathbf{V}_1 \otimes \mathbf{A}_1 & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{V}_2 \otimes \mathbf{A}_2 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{V}_3 \otimes \mathbf{A}_3 \end{bmatrix} \tag{3.25}$$

where the zeros off-diagonal represent the independence between component terms, and ($\otimes$) is the Kronecker product which allows for dependence between random effects within a component term. $\mathbf{V}_1 \otimes \mathbf{A}_1$ is the expected (co)variance matrix at the stratum level, $\mathbf{V}_2 \otimes \mathbf{A}_2$ the expected (co)variance matrix at the PSU level, and $\mathbf{V}_3 \otimes \mathbf{A}_3$ is the expected (co)variance matrix at the individual level. The (co)variance matrices ($\mathbf{V}$) are low-dimensional and are to be estimated, while the structured matrices ($\mathbf{A}$) are high dimensional and treated as known. Each diagonal element in equation (3.25) corresponds to a component term in the random effect structure of equation (3.20). That is $\mathbf{V}_1 \otimes \mathbf{A}_1 = V_s \cdot 1$, $\mathbf{V}_2 \otimes \mathbf{A}_2 = V_{s_p} \cdot 1$, and $\mathbf{V}_3 \otimes \mathbf{A}_3 = V_{s_{p_i}} \otimes \frac{1}{6}(\mathbf{I}_4 + \mathbf{1}_4)$. The effects of the independent random components are additive, such that equation (3.25) can be equivalently represented as:

$$\mathbf{G} = (\mathbf{V}_1 \otimes \mathbf{A}_1) \oplus (\mathbf{V}_2 \otimes \mathbf{A}_2) \oplus (\mathbf{V}_3 \otimes \mathbf{A}_3) \tag{3.26}$$

In multinomial data each observation is a single sample from a distribution over say m categorical outcomes, therefore the residual variance is not identified by the data. The residual variance covariance matrix must then be set to some arbitrary value which, when proper priors are used, does not pose a problem in a Bayesian analysis. Following recommendations by Hadfield (2010), the residual (co)variance matrix in our choice model is represented as:

$$\mathbf{R} = \frac{1}{m_1 + m_2}(\mathbf{I}_{(m_1+m_2)-2} + \mathbf{1}_{(m_1+m_2)-2}) = \frac{1}{6}(\mathbf{I}_4 + \mathbf{1}_4) \tag{3.27}$$

where $m_1 = 3$ and $m_2 = 3$ represent the number of choice alternatives in each of the two choice situations corresponding respectively to the expressed attitude toward health insurance cost worthiness, and the revealed health insurance choice. $\mathbf{I}_4$ and $\mathbf{1}_4$ are the (4)-dimensional identity matrix, and unit matrix, respectively.

3.5.1 Updating the latent utilities

For a given individual, the conditional density of the (4)-dimensional latent utility vector U_i is given by:

$$P(U_i|\mathbf{y}, \boldsymbol{\theta}, \mathbf{R}, \mathbf{G}) \propto f_i(y_i|U_i) f_N(e_i|\mathbf{r}_i \mathbf{R}_{/i}^{-1} \mathbf{e}_{/i}, r_i - \mathbf{r}_i \mathbf{R}_{/i}^{-1} \mathbf{r}_i^{'}) \tag{3.28}$$

where f_N represents the multivariate normal distribution with specified mean vector and (co)variance matrix. Hence equation (3.28) suggests that the conditional density of the latent vector of utilities for individual i, is proportional to the product of the conditional distribution of the joint outcome y_i, given the vector of latent utilities U_i and the joint probability density of the utility residuals. The (4)-dimensional vector of latent utility residuals e_i for individual i follows a conditional normal distribution, where the conditioning is on the $(4) \times (N - 1)$ residuals associated with the other individuals in the sample. The notation $/i$ denotes vectors or matrices with the i^{th} row and or column removed. This conditioning accounts for residual correlation across individuals.

Since latent utilities are updated in blocks of correlated residuals, correlations occurs across each of the (4) alternative residuals for a given individual. This is achieved through block sampling, where a block is a group of residuals expected to be correlated in equation (3.23). Equation (3.28) can then be rewritten as:

$$P(\mathbf{U}_k|\mathbf{y}, \boldsymbol{\theta}, \mathbf{R}, \mathbf{G}) \propto p_i(y_i|U_k) f_N(\mathbf{e}_k|\mathbf{0}, \mathbf{R}_k) \tag{3.29}$$

where k indexes blocks of latent utilities in equation (3.23), that have non-zero residual covariances. Because residuals are correlated across choice situations, alternatives and individuals, we have a total of $(N) \times (4)$ residual correlations, with $k = 1$. Therefore the conditional density of each latent utility U_{ijt} for all $i = 1 \cdots N$, $j = 1 \cdots m - 1$, and $t = 1, 2$, is obtained by conditioning each e_{ijt} on the remaining $(3)+(N-1)\times(4)$ residuals.

The average posterior (co)variance matrix $\mathbf{M}$ of the single block $(4) \times (N)$ dimensional vector $\mathbf{U}_k$ with $k = 1$ is updated at each iteration of the burn-in period following Haario et al. (2001). An efficient multivariate proposal density with covariance matrix $\nu\mathbf{M}$ is determined using adaptive methods during the burn-in phase. The scalar ν is obtained using the method of Ovaskainen et al. (2008) so that the proportion of successful jumps in the Markov Chain is optimal at a rate of 0.23 for the $(4) \times (N)$ multidimensional vector $\mathbf{U}_k$ with $k = 1$ (Gelman et al., 2004).

3.5.2 Updating the location vector $((\boldsymbol{\theta} = [\boldsymbol{\beta}', \mathbf{u}']))$

The location vector $\boldsymbol{\theta}$ is sampled as a block using a method by Gracia-Cortes and Sorensen (2001) which involves solving the sparse linear system:

$$\tilde{\boldsymbol{\theta}} = \mathbf{C}^{-1}\mathbf{W}'\mathbf{R}^{-1}(\mathbf{U} - \mathbf{W}\boldsymbol{\theta}_\star - \mathbf{e}_\star) \tag{3.30}$$

This system is solved using cholemia factorization from the Sparse library in R by Davis (2006). $\mathbf{C}$ is a sparse matrix (populated primarily with zeros) representing the model coefficient matrix:

$$\mathbf{C} = \mathbf{W}'\mathbf{R}^{-1}\mathbf{W} + \begin{bmatrix} \mathbf{B}^{-1} & \mathbf{0} \\ \mathbf{0} & \mathbf{G}^{-1} \end{bmatrix} \tag{3.31}$$

$\boldsymbol{\theta}_\star = [\boldsymbol{\beta}'_\star, \mathbf{u}'_\star]$ and $\mathbf{e}_\star$ are random draws from the multivariate normal distributions:

$$\begin{bmatrix} \boldsymbol{\beta}_\star \\ \mathbf{u}_\star \end{bmatrix} \sim N\left(\begin{bmatrix} \boldsymbol{\beta}_0 \\ \mathbf{0} \end{bmatrix}, \begin{bmatrix} \mathbf{B} & \mathbf{0} \\ \mathbf{0} & \mathbf{G} \end{bmatrix} \right) \tag{3.32}$$

and

$$\mathbf{e}_\star \sim N(\mathbf{W}\boldsymbol{\theta}_\star, \mathbf{R}) \tag{3.33}$$

A realization from the required probability distribution $P(\boldsymbol{\theta}|\mathbf{U}, \mathbf{W}, \mathbf{R}, \mathbf{G})$ is then obtained as $\tilde{\boldsymbol{\theta}} + \boldsymbol{\theta}_\star$

3.5.3 Updating the variance structure G and R

Since the residual(co)variance matrix $\mathbf{R}$ is not identified, and thus cannot be estimated from the data, its elements are kept fixed as specified in equation (3.27). All information for its estimation comes from the inverse-Wishart prior distribution, following a conditional sampling strategy provided by Korsgaard et al. (1999).

For the $\mathbf{G}$ structure as represented in equation (3.26), the sum of squares matrix associated with each of the three random components has the form:

$$\mathbf{S} = \boldsymbol{\phi}^{'}\mathbf{A}^{-1}\boldsymbol{\phi} \tag{3.34}$$

where $\boldsymbol{\phi}$ is a matrix of random effects with each row indexing the relevant row/column of $\mathbf{A}$, and each column indexing the relevant row/column in $\mathbf{V}$, and also $\mathbf{A}$ and $\mathbf{V}$ defined as in equation (3.25) and equation(3.26). The parameter (co)variance matrix can then be sampled from the inverse-Wishart distribution:

$$\mathbf{V} \sim \mathbf{IW}((\mathbf{S}_p + \mathbf{S})^{-1}, n_p + n) \tag{3.35}$$

where $\mathbf{S}_p$ and n_p are the prior sum of squares and prior degree's of freedom, respectively, and n is the number of rows in the matrix of random effects $\boldsymbol{\phi}$.

3.6 Descriptive Findings

The unconditional joint and marginal distributions of the two dependent variables are presented in table (3.2). Recall that ATTHICW captures a respondent's stated health insurance preference in early rounds of the survey, while INSURANCE captures the health insurance outcome as recorded at the end of the survey panel. The marginal distribution of the insurance preference variable in table (3.2) suggests that for the majority of the respondents, 61.45%, health insurance is worth its cost, for 24.48% health insurance is not worth its cost, and the remaining 14.07% are uncertain. Similarly the marginal distribution of the insurance coverage indicator suggests that, 62.17% of the respondents have some private insurance, 19.54% are only publicly insured, and 18.29% are uninsured.

Now focusing on the joint distribution in table (3.2), for the 62.17% recorded as having some private insured in 2007, 41% responded health insurance is worth its cost, 14.54% stated health insurance is not worth its cost, while the remaining, 6.63% were uncertain. For the 19.54% recorded as only publicly insured, the majority, 12.34%, responded health insurance is worth its cost, 3.75% stated health insurance is not worth its cost, with the remaining 3.44% uncertain. Similarly for the 18.29% recorded as uninsured, 8.10% expressed health insurance to be worth its cost, 6.19% responded health insurance is not worth its cost, while only 4% were uncertain.

The second column of Table (3.3) presents the results of the Chi-squared test for differences in the probabilities of belonging to one of the 3 health insurance categories (uninsured, some private, public only). The Chi-squared statistic of 6752.032 and corresponding p-value $<$ 2.2e-16, strongly suggests the insurance choice categories differ in frequency. The third column of Table (3.3) presents the Chi-squared test for dependence between the insurance choice outcome indicator and the insurance preference indicator. The Chi-squared statistics of 628.0411 and corresponding p-value $<$ 2.2e-16, also strongly suggest health insurance choice is dependent on expressed health insurance preference. The econometric framework is implemented to account for covariates that affect the joint distribution of the two dependent variables. The next section presents the results of the econometric estimation.

3.7 Econometric Results

The described estimation procedure in section 5 for the panel-like error components mixed logit model developed in section 3, is implemented by adapting the R package MCMCglmm by Hadfield (2010) introduced in the quantitative genetics literature. Since the aim of this study is to test the two hypotheses generated in section 2 as explained in section 4, the focus of interpretation of the results is on the estimated covariance matrix representing variations between individuals across choice alternatives and choice situations as shown in equation (3.21). This result is summarized in equation (3.36).

Before proceeding to interpret the results in equation (3.36), it's important to note that the complete set of results for the estimated model is presented in table (3.4), and subsequent trace plots in the appendix. Table (3.4) shows the posterior means of the variances V_s, V_{s_p} and $V_{s_{p_i}}$ for the three random effects in the model, with their corresponding 95% credible intervals. All estimated coefficients have corresponding 95% credible intervals that do not contain zero, suggesting their significance at the 5% level.

$$V_{s_{p_i}} = \begin{bmatrix} 0.5543 & 0.6631 & \mathbf{-0.2328} & \mathbf{1.2172} \\ . & 1.2224 & \mathbf{-0.8153} & \mathbf{0.8168} \\ . & . & 3.8275 & 1.7720 \\ . & . & . & 7.5979 \end{bmatrix} \tag{3.36}$$

With regards to the two hypotheses of interest in this study, the result $\sigma_{p_r w} = 1.2172$ in equation (3.36) suggests that on average, compared to being uncertain, adults expressing health insurance to be worth its cost are more likely to have some private insurance relative to being uninsured. Similarly $\sigma_{p_u w} = 0.8168$ suggests that relative to expressing an uncertain preference, individuals expressing health insurance to be worth its cost are more likely to be publicly only insured compared to being uninsured.

The two results in the above paragraph are consistent with the first hypothesis which stipulates that on average MEPS respondents stating health insurance to be worth its cost, will tend to have a shadow cost of health insurance that is less than the real marginal benefit from health insurance consumption, such that an insured state is preferred (whether private or public) over that of being uninsured.

The strength of likeliness to seek health insurance over being uninsured for respondents that express health insurance to be worth its cost, relative to being uncertain, is stronger for individuals with some private coverage, compared to those with only public coverage (1.2172>0.8168). This observation suggests that on average, for MEPS respondents, the level of optimism toward (or satisfaction from) having some private coverage tends to exceed that of having only public coverage.

The result $\sigma_{p_r n_w} = -0.2328$ in equation (3.36) suggests on average compared to being uncertain, the individuals expressing health insurance to not be worth its cost, are less likely to have some private insurance, relative to being uninsured. Similarly $\sigma_{p_u n_w} = -0.8153$ suggests that on average, relative to being uncertain, adult respondents expressing health insurance to not be worth its cost are less likely to be only publicly insured, relative to being uninsured.

These results are consistent with the second hypothesis which suggests that on average respondents stating health insurance to not be worth its cost will have a shadow cost of health insurance consumption that exceeds the real marginal benefit derived from health insurance consumption, such that an uninsured state is preferred over that of being insured (whether privately or publicly).

The strength of unlikeliness to seek health insurance for those that expressed health insurance to not be worth its cost relative to being uncertain is stronger for public only coverage compared to having some private coverage (0.8153>0.2328). This observation suggest that on average, for MEPS respondents, the level of skepticism of (or dissatisfaction from) public only coverage tend to exceed that of having some private coverage.

3.8 Discussion and Limitations

The evidence from the econometric estimation validates the two hypotheses. This implies consistency with the presented economic model's predictions in section 2, and further suggests that the rationality assumption as a guiding mechanism for the 2007 MEPS data generating process, in the context of health insurance choice, is reasonable.

A potential limitation however of this study comes from the fact that public coverage is only available to individuals of a certain age and income category. Therefore there is a selectivity issue associated with enrollment in this category that can potentially bias inferential results. The inclusion of income and age as covariates in the model accounts however for their effects on the distributional properties of individual respondents across the 3 insurance indicator categories. Furthermore, in this analysis, no conclusions are drawn with respect to the effect of the covariates on the joint distribution of the two dependents variables. Interest lies exclusively in the correlation between the levels of the dependent variables controlling for important covariates defining respondents stable preferences as they relate to the choice process. As such, this limitation has mild significance in the analysis.

In addition to providing an answer to the intended question of interest, the study produces results that are consistent with the past literature, with respect to individuals' satisfaction (or optimism) toward health insurance consumption. The results in fact suggest relatively more skepticism (dissatisfaction) toward public only coverage compared to having some private coverage. This is because the strength of unlikeliness to seek health insurance for those that expressed health insurance to not be worth its cost relative to being uncertain, is stronger for public only coverage compared to having some private coverage. In their review of evidence regarding enrollment into a variety of public programs Remler et al. (2001) note that "it may well be that potential recipients do not value health insurance as strongly as policy analysts do– a possibility worth exploring in depth."(p.15). Also Peterson (2004) notes that among important lessons learned regarding state efforts to expend health insurance coverage is that " because many. . . do not understand or are skeptical about the value of insurance, offering coverage does not translate into people accepting it" (p.174).

A possible explanation for such relative skepticism (dissatisfaction) in public coverage observed for MEPS respondents is that in the U.S., having only public health insurance such as the Medicare fee-for-service program, provide limited coverage to beneficiaries. The plan has substantial cost-sharing requirements and fails to cover for example preventive care or, until not too long ago, prescriptive drugs. On the other hand private coverage provides an alternative with relatively more comprehensive coverage, although with less provider choice. Individuals' skepticism (or dissatisfaction) may be suggestive that respondents value relatively more the depth of coverage (obtained from having some private insurance) than the range of provider (available from public only insurance).

3.9 Conclusion

The rational choice approach by assuming individuals make decisions in a rational, rather than stochastic manner, has allowed economists and other social scientists to model behavior and make predictions about future actions. The whole field of discrete choice modeling in economics has relied on the theoretical vision of rational choice theory by assuming agents behave rationally in the choices they make.

The increasing criticism however, of this theoretical vision in the literature because of conflicting evidence from experimental studies, has suggested the importance of testing for the validity of this assumption in a discrete choice modeling situation prior to proceeding with the modeling exercise. Since no such test has been performed in the literature to the best of my knowledge, this paper has attempted to fill this void by presenting a framework for testing the rationality assumption in discrete choice modeling with application to health insurance choice in the U.S., using the 2007 Medical Expenditure Panel Survey dataset.

For this purpose, a new variant of the mixed logit model was introduced, which relied on the multinomial logit representation of the weighted logit formula, rather than the currently used conditional logit representation in the literature. This representation allowed for the joint modeling of the conditional distribution of respondents' stated preference for health insurance in earlier rounds with revealed health insurance, controlling for a set of alternative invariant covariates capturing respondents'stable preferences as assumed by rational choice theory.

The model was then estimated within the Bayesian framework by adapting the R package MCMCglmm (Hadfield, 2010) introduced in the quantitative genetics literature. Observation of the signs and significance of the estimated covariance coefficients between the levels of the two jointly modeled categorical dependent variables using the variance covariance matrix at the individuals' level, provided the test for the adequacy of the rational choice assumption in this case.

The findings revealed that on average, relative to being uncertain, the individuals expressing health insurance to not be worth its cost were less likely to have some private insurance or be only publicly insured, relative to being uninsured. In addition relative to being uncertain, individuals expressing health insurance to be worth its cost were more likely to be insured (privately or publicly) over being uninsured. Moreover, the findings showed consistency with the past literature with respect to individuals' skepticism toward health insurance in the U.S. (Remler et al., 2001; Peterson, 2004), and suggested relatively more skepticism (dissatisfaction) toward public coverage compared to having some private coverage.

Overall, the results validated the adequacy of rationality as a guiding mechanism for the 2007 MEPS data generating process, and further allowed for the conclusion that the evidence is not enough to reject the idea that adults in the U.S. behave rationally in their choices of health insurance. Therefore one can safely proceed to model adults' health insurance choices in the U.S. within the discrete choice modeling framework.

Table 3.1: Summary statistics for the independent variables in the model

	N = 18035	Mean	SD
Demographic characteristics			
AGE	age of respondent in years	46.16	17.444
SEX	= 1 if respondent is female	0.542	0.498
EDUCYR	years of schooling	12.49	3.217
INCOME	income in $1000	29.998	31.472
Health characteristics			
VERGOOD	= 1 if very good health	0.3342	0.472
GOOD	= 1 if good health	0.3259	0.469
FAIRPOOR	= 1 if fair or poor health	0.1645	0.371
Marital Status			
MARRIED	= 1 if currently married	0.5621	0.496
PMARRIED	= 1 if previously married	0.2022	0.402
Variance estimation Var.			
VARSTR	variance estimation stratum	534.3	489.89
VARPSU	variance estimation PSU	1.664	0.627

Table 3.2: Cross tabulation of attitudes toward (Atthicw) and purchase of (Insurance) insurance. Values are in percentage.

	Attitude (Atthicw)			
Insurance	Uncertain	Not Worth	Worth	Marginal
Uninsured	4.00	6.19	8.10	18.29
Some Private	6.63	14.54	41.00	62.17
Public Only	3.44	3.75	12.35	19.54
Marginal	14.07	24.48	61.45	100

Table 3.3: Significance test results on the joint unconditional distribution of Insurance and Atthicw.

	χ^2 test for differences in the probabilities of enrolling	Pearson χ^2 test for dependence between outcomes
χ^2 statistics	6752.032	628.04
df	2	4
p-value	$< 2.2e-16$	$< 2.2e-16$
5% Significance	yes	yes

Table 3.4: Posterior means and 95% credible intervals for random effects variances

	Post. mean	L-95 Percent CI	U-95 Percent CI
Between stratum variation			
V_s	0.1112	0.02353	0.1758
PSU variation within strata			
V_{s_p}	0.0838	0.02353	0.1413
Variation across individuals within PSU within strata			
$\sigma_{p_r n_w}$	-0.2328	-0.3984	-0.0010
$\sigma_{p_u n_w}$	-0.8153	-1.0815	-0.4237
$\sigma_{p_r w}$	1.2172	0.9331	1.5444
$\sigma_{p_u w}$	0.8168	0.4880	1.1570
$\sigma_{p_r p_r}$	0.5543	0.2996	0.9408
$\sigma_{p_r p_u}$	0.6631	0.2538	1.2533
$\sigma_{p_u p_u}$	1.2224	0.5146	2.2353
$\sigma_{n_w n_w}$	3.8275	1.0911	6.2453
$\sigma_{n_w w}$	1.7720	0.4049	3.6088
σ_{ww}	7.5979	4.8008	10.6073
Iterations	20000		
Thinning interval	10000		
Drawn sample size	2000		
DIC	54695.96		

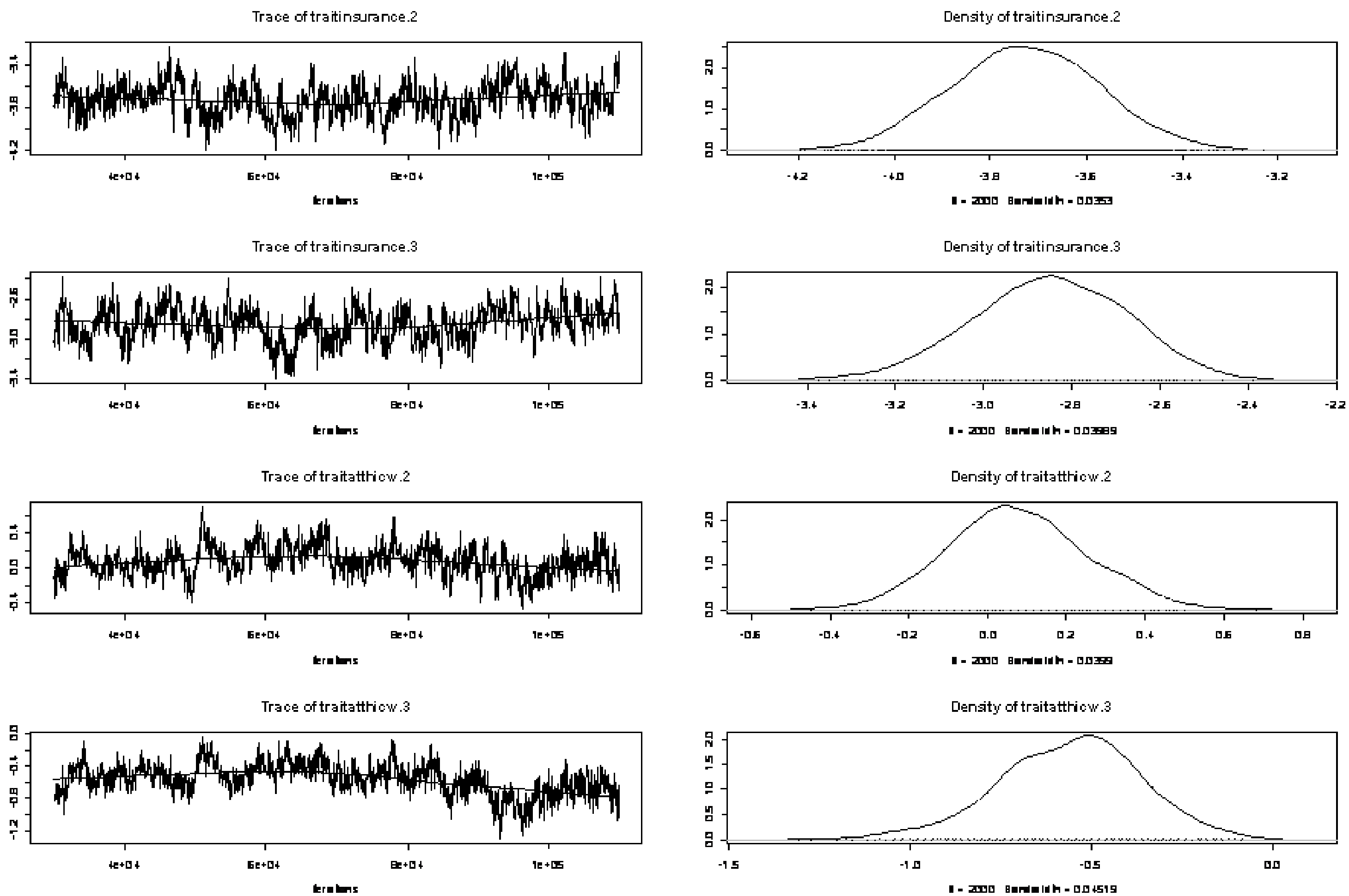

Figure 3.2: Posterior distribution of the estimated INTERCEPT coefficient

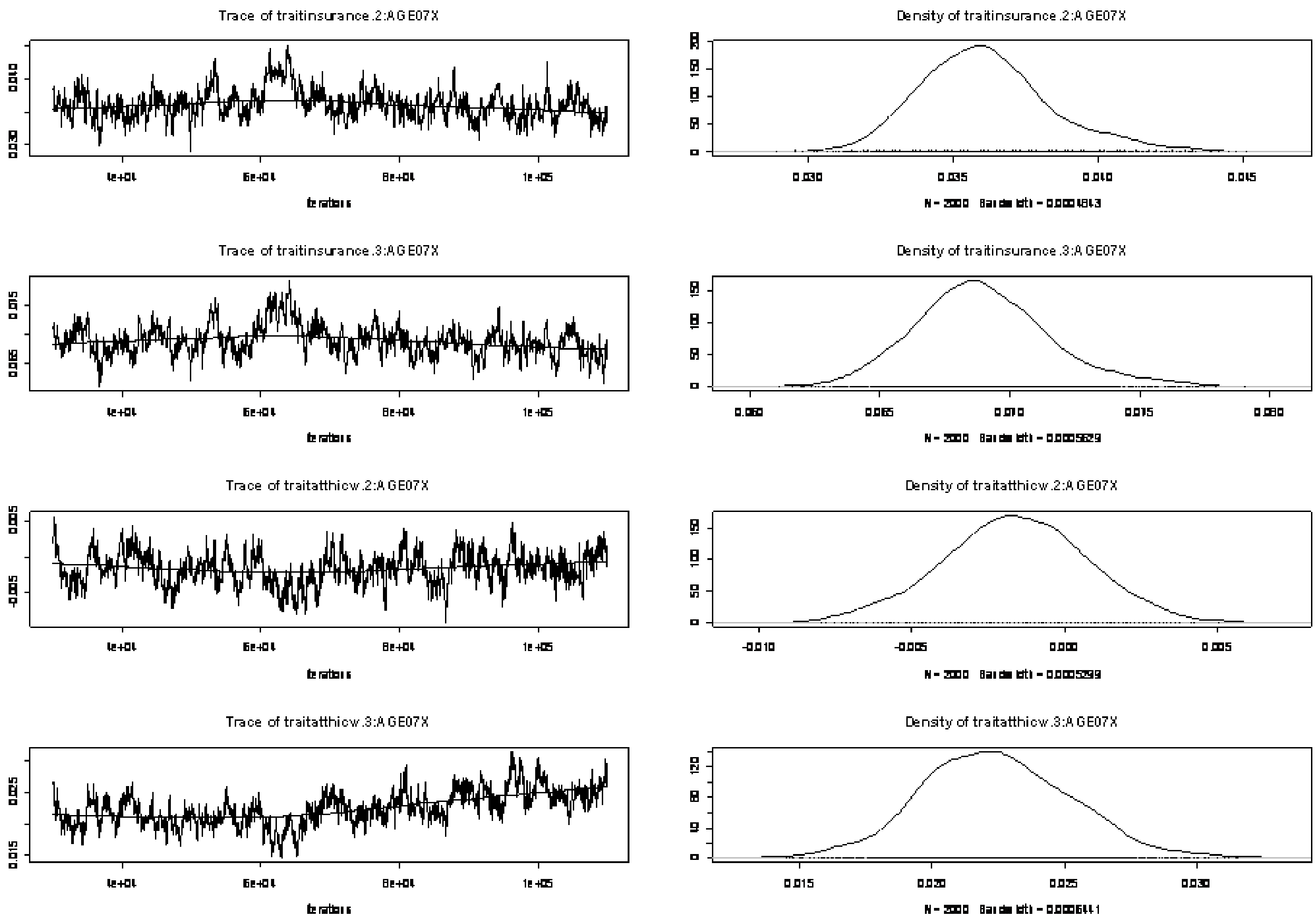

Figure 3.3: Posterior distribution of the estimated AGE coefficient

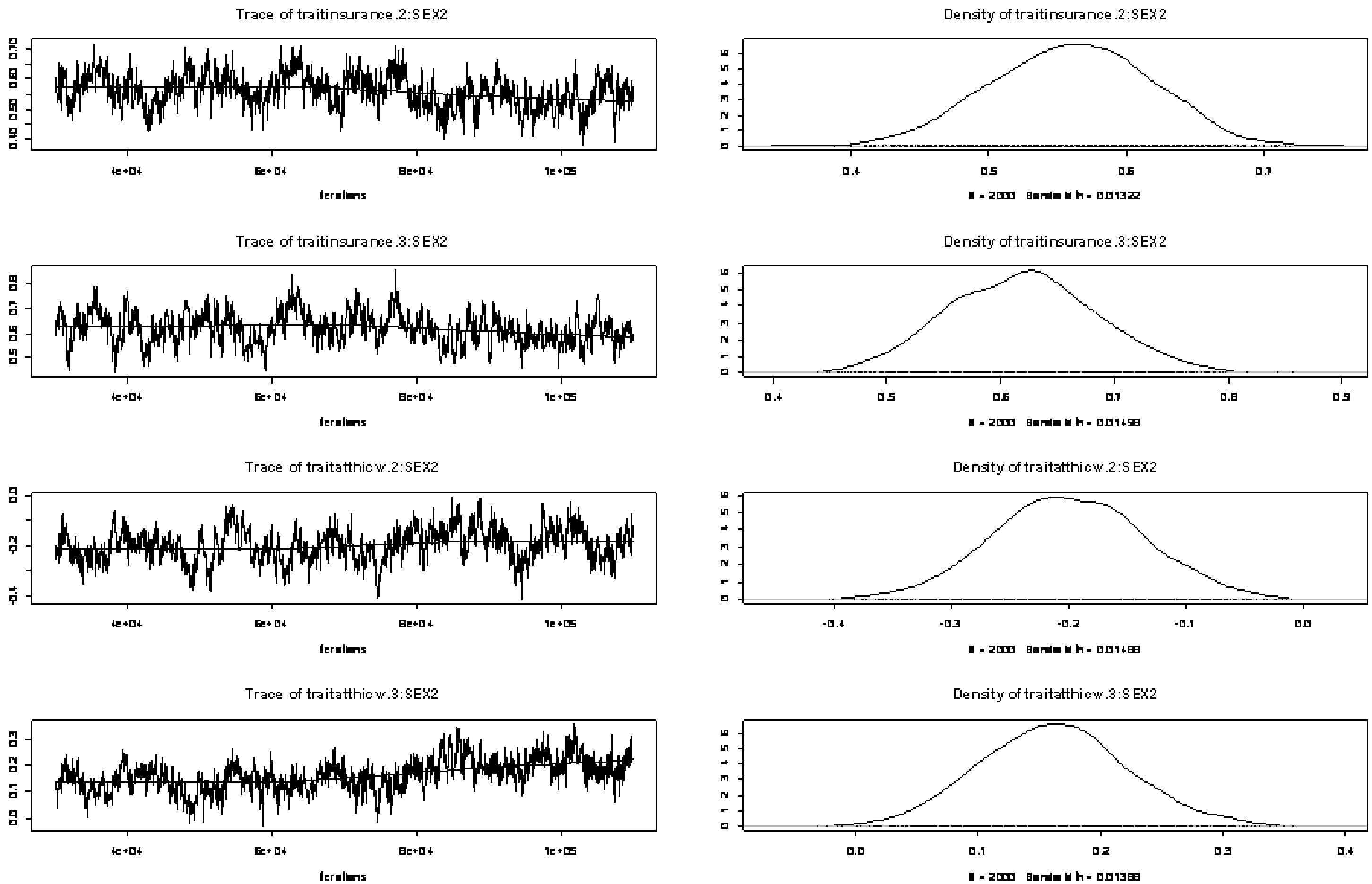

Figure 3.4: Posterior distribution of the estimated gender coefficient

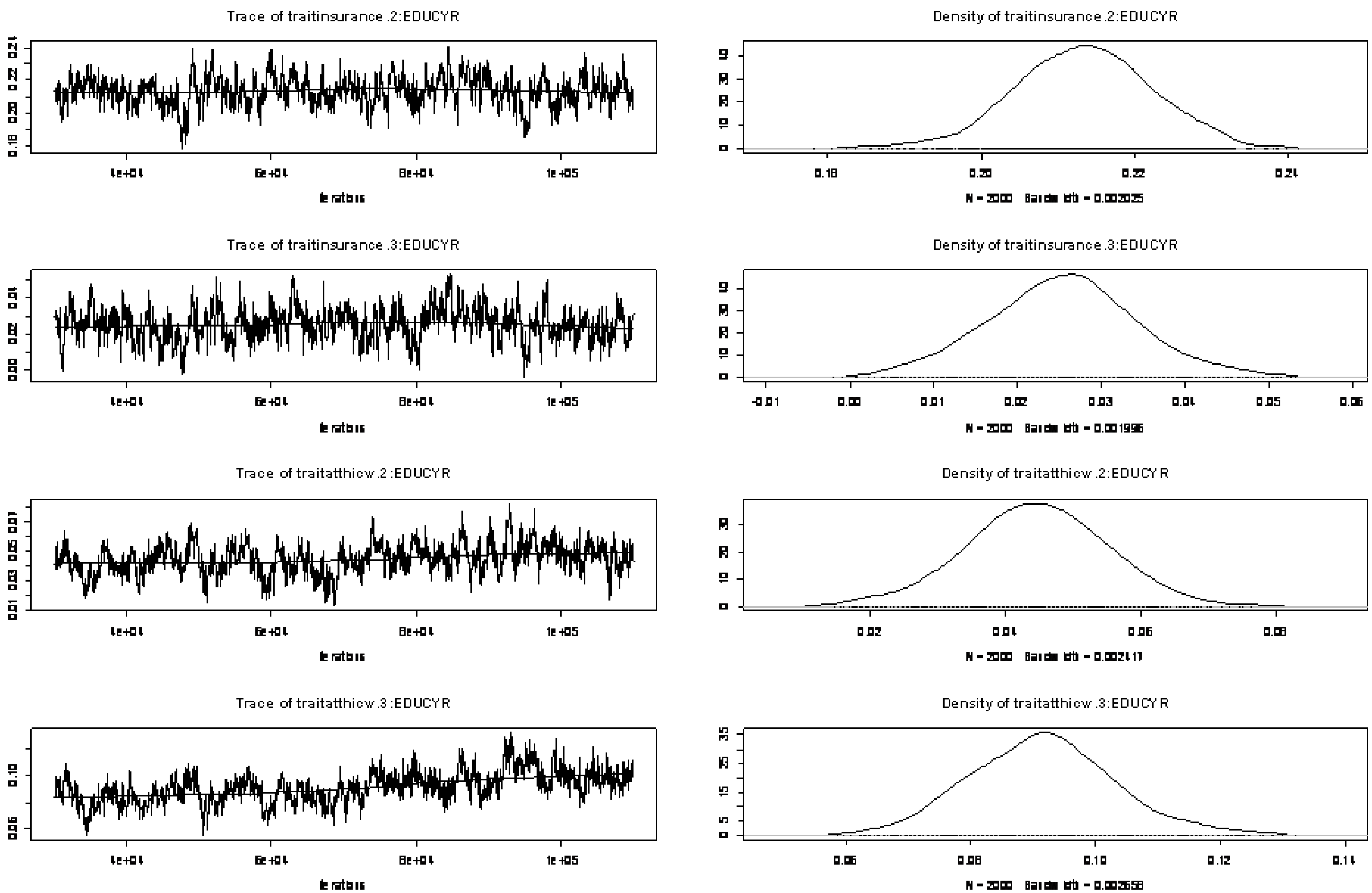

Figure 3.5: Posterior distribution of the estimated coefficients for EDUCYR

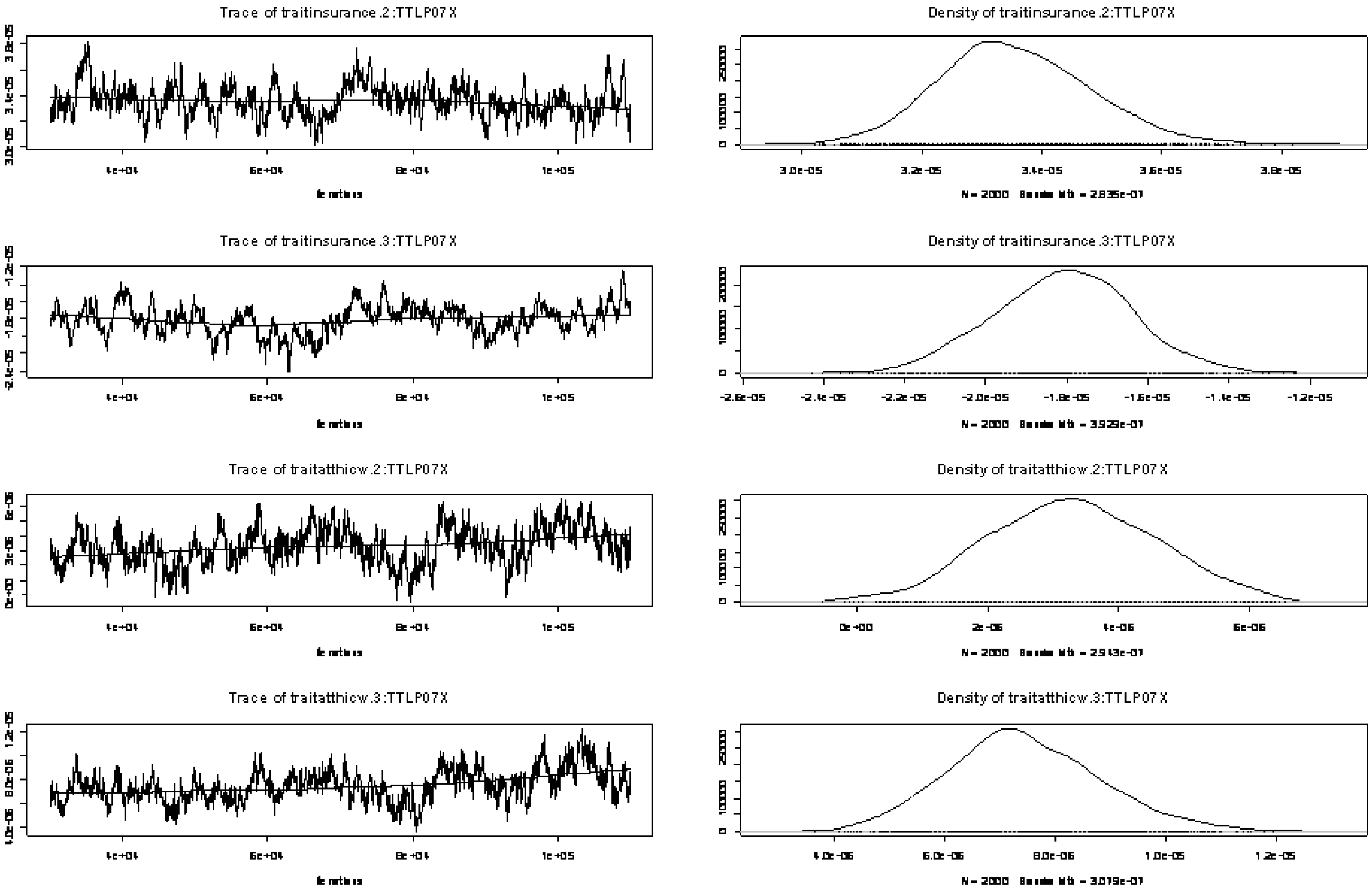

Figure 3.6: Posterior distribution of the estimated coefficients for $INCOME \equiv TTLP07X$

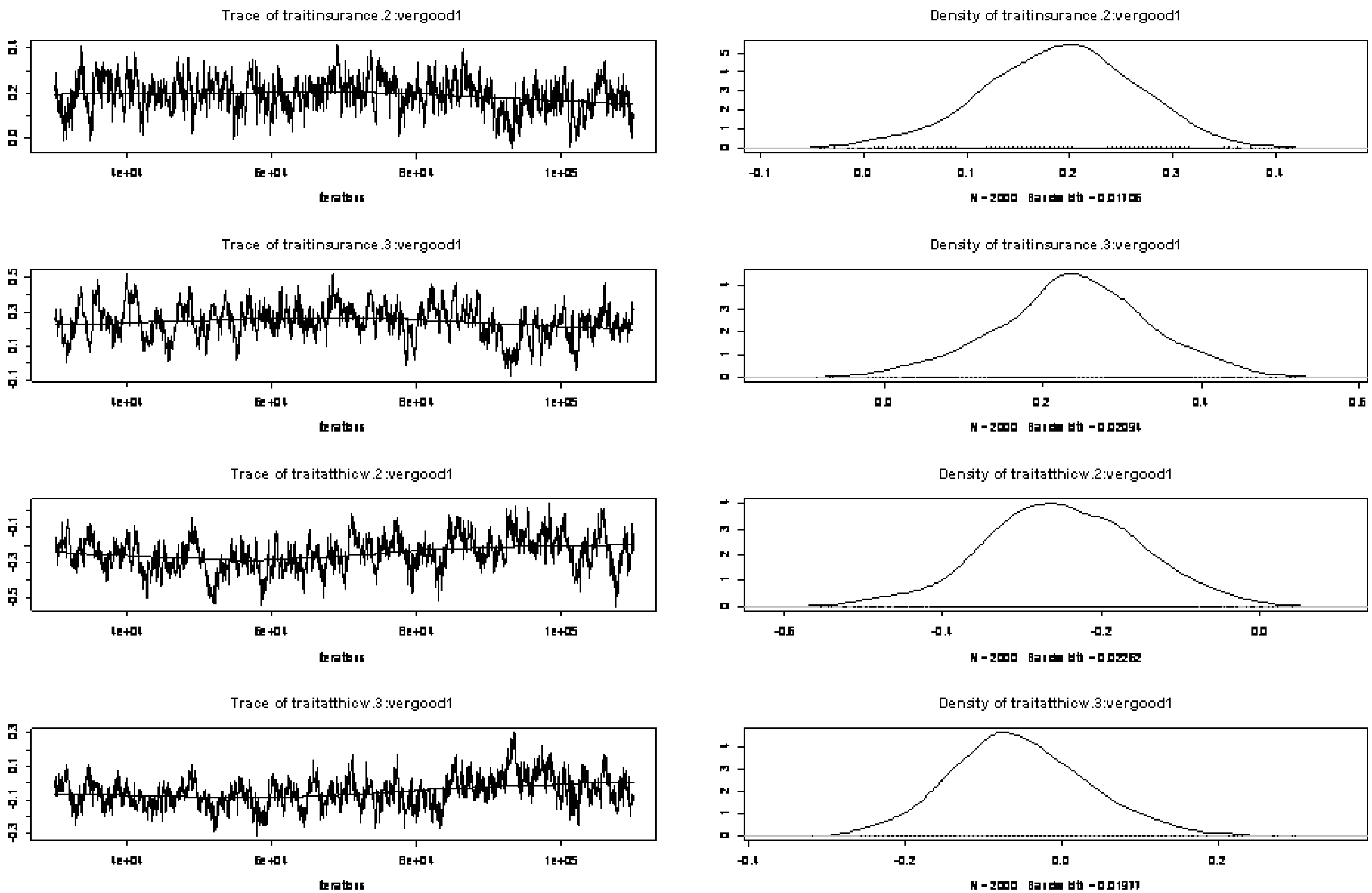

Figure 3.7: Posterior distribution of the estimated coefficients for VERGOOD

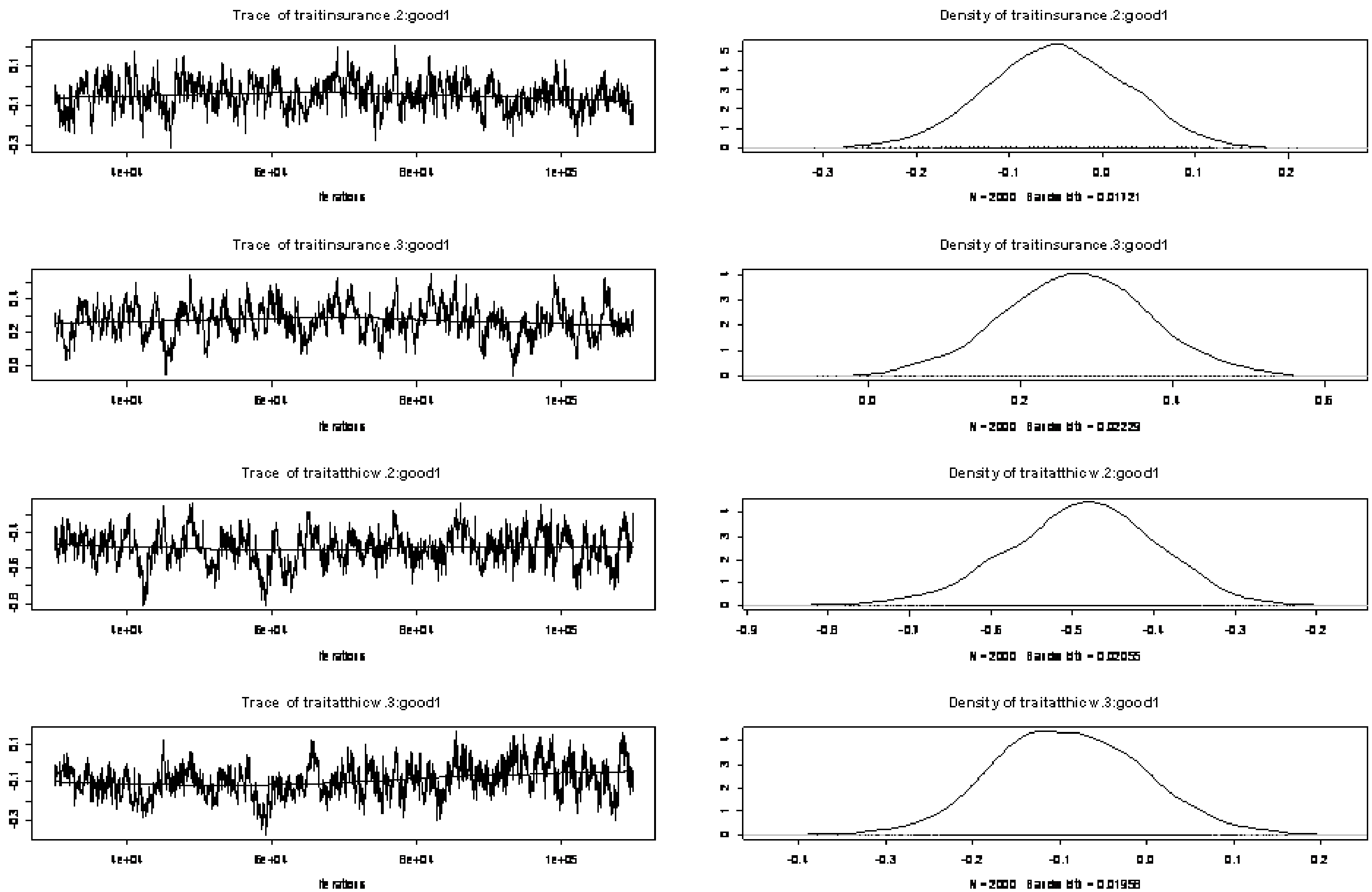

Figure 3.8: Posterior distribution of the estimated coefficients for GOOD

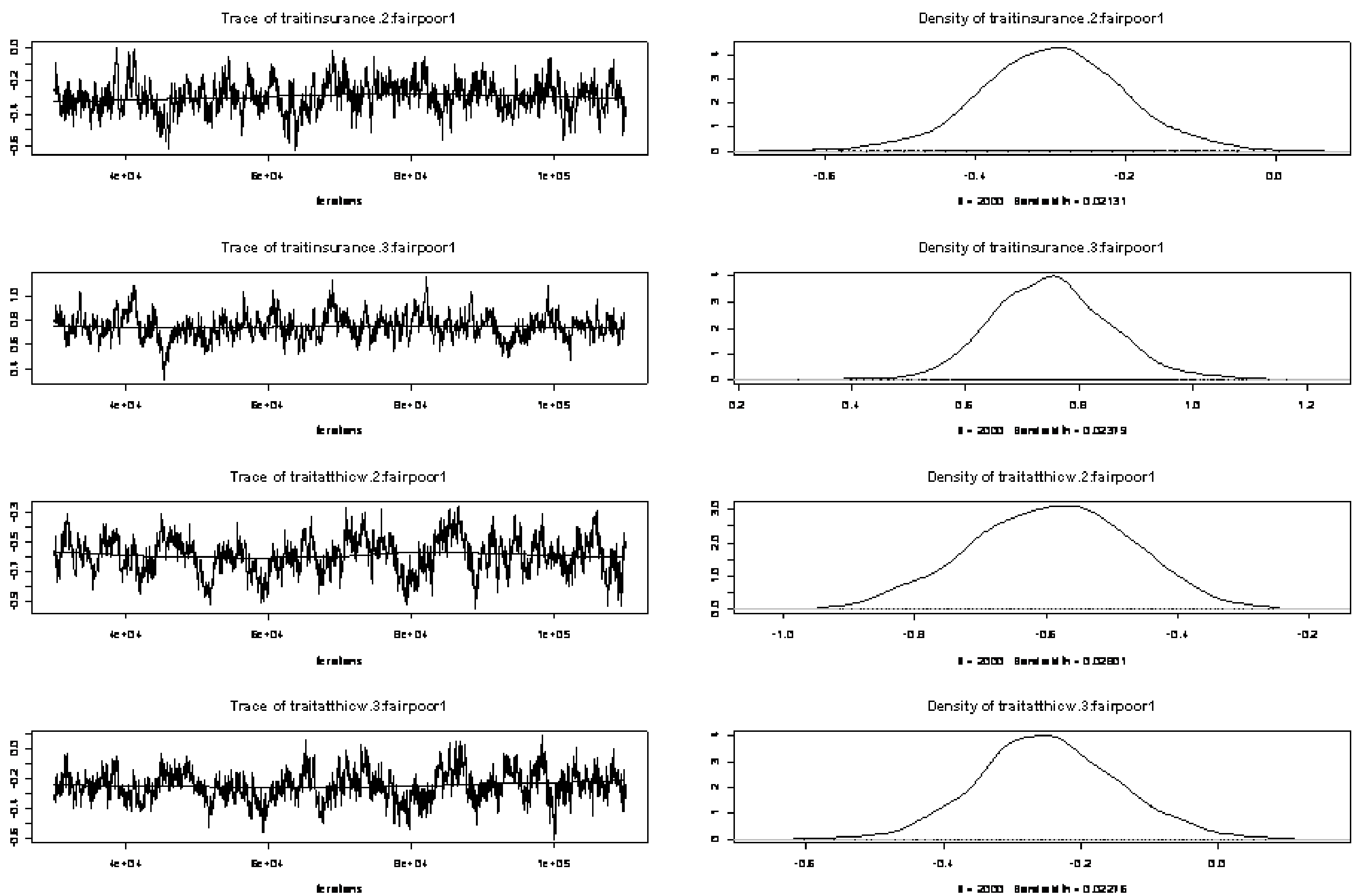

Figure 3.9: Posterior distribution of the estimated coefficients for FAIRPOOR

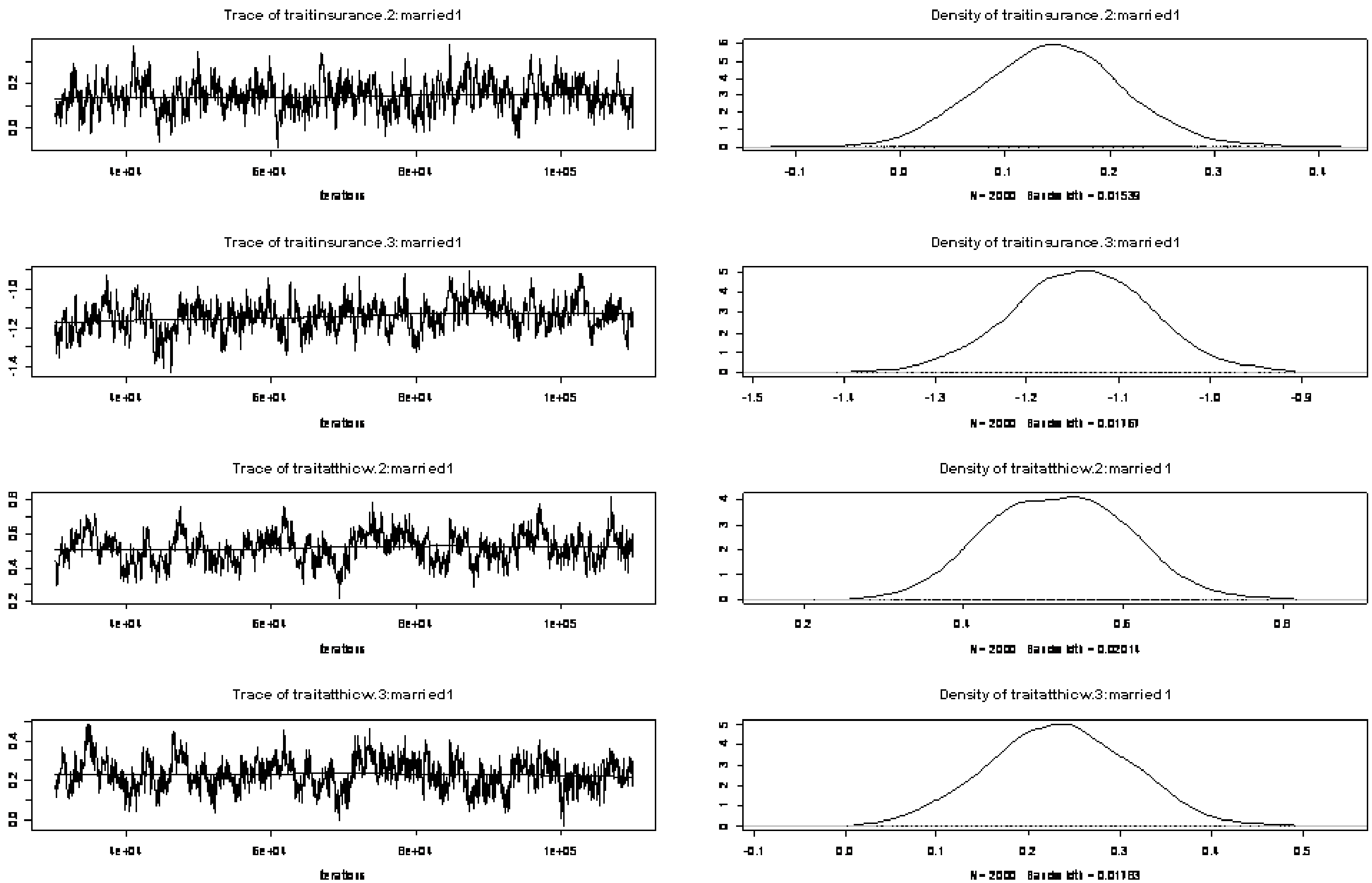

Figure 3.10: Posterior distribution of the estimated coefficients for MARRIED

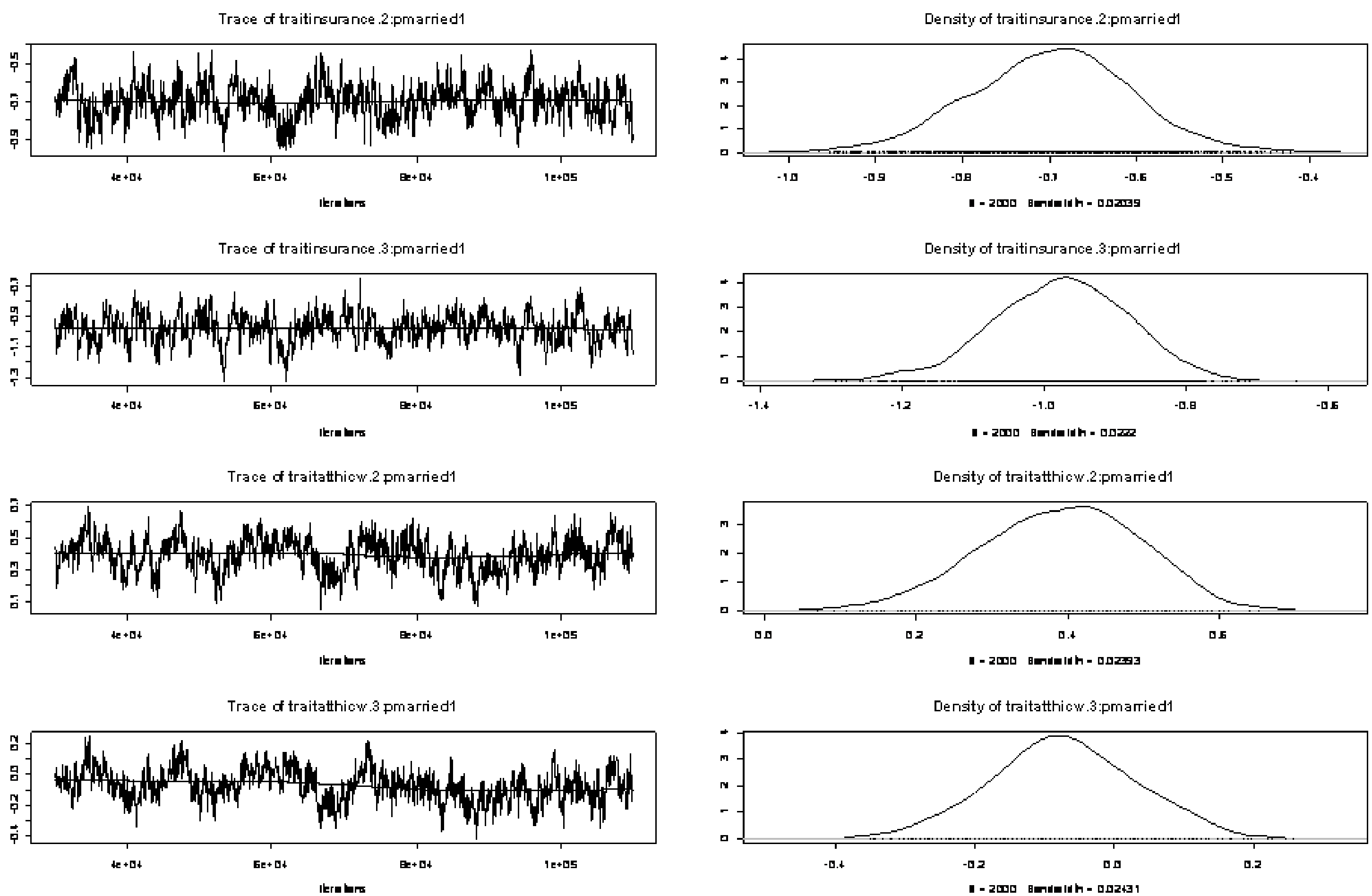

Figure 3.11: Posterior distribution of the estimated coefficients for PMARRIED

CHAPTER 4

MODELING ADULTS HEALTH INSURANCE ENROLLMENT DECISIONS IN THE U.S., UNDER PREFERENCES ENDOGENEITY: A BAYESIAN MULTINOMIAL PROBIT APPROACH

4.1 Introduction

The results of the previous chapter suggest that adult respondents in the 2007 MEPS make their health insurance choices in a way consistent with rational choice theory predictions. It is therefore reasonable to use the discrete choice modeling framework which relies on the assumption that decision makers are rational, to model adults' health insurance choices. Of interest is the effects of stated health insurance preferences, on revealed choices of health insurance. In fact, in economics choice theory is based on the twin concepts of willingness and ability to pay. Within the context of health insurance enrollment decisions, attitudinal questions capturing individuals preferences for health insurance have been shown to have strong predictive power on actual choice behavior (Keane, 2004; Parente et al., 2004).

Early work by Goldstein and Pauly (1976) and Feldman et al. (1997) suggest that workers sorting among employment alternatives reflect their tastes for employment-sponsored health insurance. Monheit and Vistnes (1999), using attitudinal measures, found that weak preferences for health insurance are an important factor in the decision by single wage earners to self-select into jobs without insurance. In a more recent study Monheit and Vistnes (2008), using responses to questions capturing health insurance preferences, conclude that individuals with weak preferences for coverage are more likely to be uninsured than those with strong preferences. The authors also found that single workers and one-wage-earner couples with weak or uncertain preferences are less likely than those with strong preferences to obtain offers of employment-sponsored health insurance and to enroll.

The above referenced literature mainly focuses however on examining the role of health insurance preferences on enrollment decisions into employment-sponsored health insurance(ESI). Furthermore, this literature assumes the effects of health insurance preferences to be exogenous in the statistical sense, which in the case of Monheit and Vistnes (2008) was justified by the fact that responses to the stated preference measures were obtained independently of survey questions regarding health insurance status. Therefore the authors concluded that all concerns of self selection bias were fully mitigated.

The goal is to understand how stated preferences for health insurance by adults in the U.S. affect their choices among the three health insurance enrollment categories (Any private, Public only, Uninsured), and the contribution of this chapter methodological in essence. The exogeneity assumption made by previous authors is relaxed, and a model of health insurance choice is considered, with the stated health insurance preference variable treated as being endogenously determined in the health insurance choice model. The parameters associated with this endogeneity are estimated using the Fully Gibbsian Bayesian Multinomial probit framework by Burgette and Nordheim (2009). The procedure makes use of the data augmentation principle by Albert and Chib (1993), and the partial marginalization principle of van Dyk (2010).

The chapter is organized as follows. Section II describes the empirical model of enrollment decision. Section III provides an exposition of the analytical strategy. Section IV presents the full Gibbs Markov Chain Monte Carlo sampler for the parameters in the model. Section V describes the data, while results are presented in section VI, and section VII concludes the analysis.

4.2 Empirical Model of Enrollment Decisions

The premise underlying the modeling strategy implemented lends itself to the discrete choice framework derived under the assumption of utility maximizing behavior by the agents. Individuals are assumed to be rational and to make health insurance enrollment decisions on the basis of a vector of demographic characteristics (Age, Sex, Marital status, Education level) given their needs/general health conditions captured by the dummies (Excellent, Vergood, Good, and Fairpoor) and enabling factors (Family and personal income, health insurance preference). Although many factors affect this choice process, the contention in this paper is that health insurance preference is a major determinant of the enrollment decision. The general set up of the decision process can be described as follows:

An adult respondent in the MEPS indexed by n faces a choice among m health insurance enrollment alternatives, each providing a given level of utility. The latent utility derived from the choice of alternative j is L_{nj}, for $j = 1, \dots, m$, and is only known to the individual respondent. This utility is decomposed as $L_{nj} = V_{nj} + \epsilon_{nj}$, where ϵ_{nj} captures unobserved factors affecting utility, and not included in the observed part V_{nj} of utility. The individual chooses to enroll in the health insurance category yielding the greatest utility, therefore the behavioral model consists of choosing enrollment alternative i if and only if $L_{ni} > L_{nj}$ for all $j \neq i$. The probability that respondent n chooses alternative i is given as:

$$\begin{aligned} P_{ni} &= Prob(L_{ni} > L_{nj} \ \ \forall\, j \neq i) && (4.1)\\ &= Prob(V_{ni} + \epsilon_{ni} > V_{nj} + \epsilon_{nj} \ \ \forall\, j \neq i) && (4.2)\\ &= Prob(\epsilon_{nj} - \epsilon_{ni} < V_{ni} - V_{nj} \ \ \forall\, j \neq i) && (4.3)\\ &= \int_{\epsilon} I(\epsilon_{nj} - \epsilon_{ni} < V_{ni} - V_{nj} \ j \neq i) f(\boldsymbol{\epsilon}_n) d_{\epsilon_n}, && (4.4) \end{aligned}$$

This choice probability is expressed as a cumulative distribution function of the error differences with an $(m-1)$ dimensional density function $f(\boldsymbol{\epsilon}_n)$. $I(\cdot)$ represents the indicator function, taking a value of 1 when the expression in parentheses is true and 0 otherwise. Probit specification of this choice probability jointly models the unobserved utility components using the normal density, such that $f(\boldsymbol{\epsilon}_n)$, is multivariate normal.

Endogeneity in this probit model is motivated using the following general additive form representation of the utility function for individual n choosing alternative i:

$$L_{ni} = f(y_{ni}, x_n, \beta_n) + \epsilon_{ni}, \tag{4.5}$$

where the systematic portion of the utility contains observed exogenous variables, x_n, relating to person n, the endogenous variable (health insurance preference), y_{ni}, and the parameter vector, β_n. The endogenous variable y_{ni} can be further represented as:

$$y_{ni} = g(z_n, \gamma) + \mu_{ni}, \tag{4.6}$$

where μ_{ni} and ϵ_{ni} are correlated but independent of the exogenous instruments, z_n. This correlation between the errors μ_{ni} and ϵ_{ni} implies that the health insurance preference variable is correlated with unobserved factors affecting utility from enrolling in the various health insurance categories. This characteristic creates the statistical endogeneity of the stated health insurance preference variable, and leads to bias using standard estimation methods, which assumes that the distribution of the outcome variable conditional on the observed regressors has a zero mean. This feature is however addressed by the fully Gibbsian Bayesian Multinomial probit estimation procedure implemented in this paper, which allows for correlations between unobservables.

4.3 Analytical Strategy

The motivation for this empirical analysis is the desire to model health insurance choices by adult respondents in the 2007 Medical Expenditure Panel (MEP) Survey.

In early rounds of the survey respondents state their preferences for health insurance by expressing its worthiness to them. Individuals either agree that health insurance is not worth the cost(NW), or disagree (W), or are uncertain(UC). Then after the last round, we observe the coverage choice made by the respondent over the scope of the panel as either Any Private, Public Only, or Uninsured, conditional on the stated preference for health insurance. *The basic research goal then is to be able to say, for a set of covariates, how the health insurance outcome probabilities vary based on differing attitudes toward health insurance cost worthiness (health insurance preferences).*

In order to address this research question, two interdependent processes are modeled. The first process relates to the insurance preference, and the second process relates to the insurance outcome conditional on the preferences. Because the categories in both preference and outcome variables are unordered the choice of labeling is arbitrary, and can be indexed with $0, 1$ and 2 respectively, with 0 being the base category. These base categories are "uncertain" for the preferences and "uninsured" for the outcomes. Therefore for each individual n, with $n = 1, 2, \ldots N$, we can define $\mathbf{Y}_n$ to be the ordered pair of insurance preference and outcome.

The probit framework is used to model both insurance preference and choice outcome, allowing the errors to be correlated. The model assumes each decision maker constructs latent utilities for each of the choice options, and chooses the option corresponding to the maximum of the utilities. To make things more explicit, in setting up our fully Gibbsian Bayesian Multinomial probit framework, we assume the existence of an 8-dimensional vector $\mathbf{L}_n$ that contains the latent utilities associated with insurance preferences and choice outcomes, relative to the respective base categories. $\mathbf{L}_n$ can be thought of as being blocked into four groups of two, so that $\mathbf{L}_n = (L_n^{p'}; L_n^{0'}, L_n^{1'}, L_n^{2'})$. The first block $L_n^{p'}$ contains utilities for health insurance preferences/selection process relative to the base category (uncertain).

The first element of this block represents the utility associated with choosing "Agree" over choosing "Uncertain;" while the second element in this block represents the utility associated with choosing "Disagree" over choosing "Uncertain." The remaining blocks relate to the outcomes conditional on the preferences 0, 1 and 2 respectively, relative to the base category (uninsured).

Given any preference choice, whether (0–uncertain, 1–Agree, 2–Disagree), with 0, 1 and 2 indexing each of the remaining blocks, each block contains two relative utilities. The first one being the utility associated with choosing "some Private" over being "uninsured," while the second one represents the utility associated with choosing "Any Public" over being "uninsured." If the first two elements of any block in $\mathbf{L}_n$ are both negative, the agent will prefer the base category. Otherwise, the individual will prefer the category that corresponds to the larger of the first two elements in that particular block. More formally, the link between $\mathbf{L}_n$ and $\mathbf{Y}_n = (Y_{n1}, Y_{n2}')$ is given by

$$Y_{n1} = \begin{cases} \mathbf{argmax}_{k\ \in\ 1,2}\ L_k^p & \text{if } \mathbf{max}_{k\ \in\ 1,2}\ L_k^p > 0 \\ 0 & \text{Otherwise.} \end{cases}$$

$$Y_{n2} = \begin{cases} \mathbf{argmax}_{k\ \in\ 1,2}\ L_k^{Y_{n1}} & \text{if } \mathbf{max}_{k\ \in\ 1,2}\ L_k^{Y_{n1}} > 0 \\ 0 & \text{Otherwise.} \end{cases}$$

$\mathbf{L}_n$ is assumed to be linear on observed covariates up to an additive normal disturbance:

$$\mathbf{L}_n = X_n\boldsymbol{\beta} + \boldsymbol{\epsilon}_n \quad n = 1, \ldots N \tag{4.7}$$

with X_n representing a matrix of covariates, $\boldsymbol{\beta}$ a vector of regression parameters, and $\boldsymbol{\epsilon}$ representing the vector of disturbances assumed to be *iid* distributed with mean zero and covariance matrix Σ. The complete data likelihood obtained if the latent utilities were observed is

$$p(\mathbf{L}|\boldsymbol{\beta}, \Sigma) \propto |\Sigma|^{-\frac{N}{2}} exp\left\{-\frac{1}{2}\sum_{n=1}^{N}(\mathbf{L}_n - X_n\boldsymbol{\beta})\Sigma^{-1}(\mathbf{L}_n - X_n\boldsymbol{\beta})\right\} \tag{4.8}$$

Since the latent utilities $\mathbf{L}_n$ are not observed, we have the incomplete data likelihood obtained by forming expectations over all $\mathbf{L}_n$, with the integrals defined over the region implied by $\mathbf{Y}_n$

$$p(\mathbf{Y}|\boldsymbol{\beta}, \Sigma) \propto |\Sigma|^{-\frac{N}{2}} \prod_{n=1}^{N}\int_{Y_n} exp\left\{-\frac{1}{2}\sum_{n=1}^{N}(\mathbf{L}_n - X_n\boldsymbol{\beta})\Sigma^{-1}(\mathbf{L}_n - X_n\boldsymbol{\beta})\right\} d\mathbf{L}_n \tag{4.9}$$

For notational convenience, the parameters in the model are defined as $\boldsymbol{\theta} = (\boldsymbol{\beta}, \Sigma)$. Assuming further X_n to be block-diagonal, then the model in matrix form can be represented as:

$$\mathbf{L}_n = \begin{bmatrix} I_2 \otimes z_n^{'} & \mathbf{0} \\ \mathbf{0} & I_6 \otimes x_n^{'} \end{bmatrix} \boldsymbol{\beta} + \boldsymbol{\epsilon}_n = X_n\boldsymbol{\beta} + \boldsymbol{\epsilon}_n. \tag{4.10}$$

where I_j is the $j \times j$ identity matrix, and $\otimes$ indicates the Kronecker product. z_n is a vector of exogenous covariates relating to the selection process (stated preferences), and x_n relates to the outcome process (revealed choices). In stacked form, equation (4.7) can be expressed as $\mathbf{L} = X\boldsymbol{\beta} + \boldsymbol{\epsilon}$, where $\mathbf{L}$ and $\boldsymbol{\epsilon}$ are $8N \times 1$, and X is $8N \times p$, while $\boldsymbol{\beta}$ is $p \times 1$. In this format, $\boldsymbol{\epsilon}$ is distributed normally with a zero mean, and covariance $I_N \otimes \Sigma$.

Because the scale of the MNP is undefined, it is customary to set the first diagonal element of the covariance matrix Σ to unity in order to achieve identification (Train, 2009, p.100-103). In the presence of endogeneity this identification issue is complicated further, requiring additional diagonal elements to be fixed at unity. In the 3×3 switching model developed here, four choice models are effectively merged, so that Σ is 8×8 and all of the odd-numbered diagonal elements fixed to one, to ensure identification. This gives the following structure for Σ

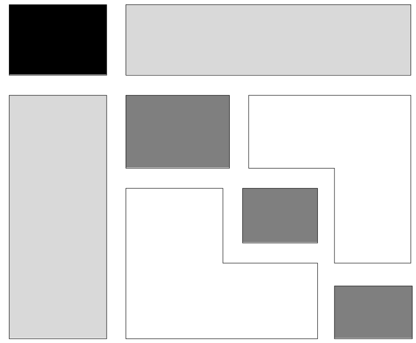

Figure 4.1: Functional partition of the covariance matrix Σ. the dark square corresponds to covariances within stated preferences, and the medium gray squares to the covariances in the choice outcome equations. The light gray rectangles show correlations between stated preference and choice outcome equations, and the white polygons indicate parameters that are not identified by the data.

$$\Sigma = \begin{bmatrix} 1 & \sigma_{12} & \sigma_{13} & \sigma_{14} & \sigma_{15} & \sigma_{16} & \sigma_{17} & \sigma_{18} \\ . & \sigma_{22} & \sigma_{23} & \sigma_{24} & \sigma_{25} & \sigma_{26} & \sigma_{27} & \sigma_{28} \\ . & . & 1 & \sigma_{34} & \sigma_{35} & \sigma_{36} & \sigma_{37} & \sigma_{18} \\ . & . & . & \sigma_{44} & \sigma_{45} & \sigma_{46} & \sigma_{47} & \sigma_{48} \\ . & . & . & . & 1 & \sigma_{56} & \sigma_{57} & \sigma_{58} \\ . & . & . & . & . & \sigma_{66} & \sigma_{67} & \sigma_{68} \\ . & . & . & . & . & . & 1 & \sigma_{78} \\ . & . & . & . & . & . & . & \sigma_{88} \end{bmatrix} \tag{4.11}$$

As illustrated in figure (4.1), which shows the functional partition of the variance covariance matrix Σ, the covariance structure of the selection (stated preference) phase is the darkest square. The covariance structures associated with an outcome choice, conditional on a stated preference, are represented by the medium gray squares. The light gray rectangles show correlation between selection and outcome equations, and are similar to the selection parameters in a standard Heckman selection model (Heckman, 1979)

4.4 The MCMC Sampler for Model Parameters

Application of Bayesian methods to the probit model was first introduced by Albert and Chib (1993), and further described by McCulloch et al. (2000). Endogeneity in the context of probit modeling has also received much attention in the literature, with Chib and Hamilton (2000) describing a model with multinomial probit selection and a binary outcome. Li and Tobias (2005) considered a binary selection model with unordered probit response, Munkin and Trivedi (2008) analyzed an ordered outcome with discrete endogenous covariates, while Burgette and Nordheim (2009) looked at a model where both selection and outcome categories are unordered. The modeling strategy used in this analysis relates closely to the latter, which is an extension of Imai and van Dyk (2005).

To circumvent the problems associated with the lack of closed form solution for the multinomial probit model, data augmentation techniques as described by Albert and Chib (1993) are coupled with MCMC methods. This is accomplished by expanding the parameter space with latent variables, which in a Bayesian context yields a full Gibbs sampler with prior specified on the identified model parameters (Burgette and Nordheim, 2009). In the following the sampler is derived for the case with three selection categories representing the stated preferences for health insurance, and three outcome categories representing the revealed health insurance choices.

4.4.1 Prior distributions

A weakly informative prior is specified following recommendations by McCulloch et al. (2000), who suggest using a weakly informative default prior in the absence of strong prior information. The specification closely follows that of Imai and van Dyk (2005), and Burgette and Nordheim (2009). The inverse of the covariance matrix Σ^{-1} is partitioned into blocks of 2×2, with $\Sigma^{-1} = \left\{\Sigma_{ij}^{-1}\right\}$, for $i, j = 1, 2, 3, 4$. For this model the following independent prior is specified:

$$\boldsymbol{\beta} \sim N(\beta_0, B_0^{-1}) \qquad and \qquad p(\Sigma) \propto |\Sigma|^{-(v+9)/2} \left(\prod_{i=1}^{4} tr(S_{ii}\Sigma_{ii}^{-1}) \right)^{-v} \tag{4.12}$$

where the elements σ_{ii} for $i = 1, 3, 5, 7$ of the covariance matrix Σ are set to one. β_0 and B_0^{-1} represent respectively the prior mean and covariance matrix of $\boldsymbol{\beta}$. v is the prior degrees of freedom for the covariance structure, and S_{ii} are the prior scale of the covariances of the four inherent discrete choice models. The prior specification for Σ is derived from the inverse-Wishart distribution, which is commonly used as a conditional conjugate prior in multivariate normal Bayesian analyses.

To obtain the prior for Σ as specified in equation (4.12), we begin with the unrestricted covariance matrix $\widetilde{\Sigma}$ ~inv-Wishart(v,$\mathbf{S}$) and transform to (a, Σ), with $a = (\widetilde{\sigma}_{11}, \widetilde{\sigma}_{33}, \widetilde{\sigma}_{55}, \widetilde{\sigma}_{77})$, and $\Sigma = A^{-1}\widetilde{\Sigma}A^{-1}$, where $A^{-1} = diag(a_1^{-0.5}, a_2^{-0.5}, a_3^{-0.5}, a_4^{-0.5}, a_5^{-0.5}, a_6^{-0.5}, a_7^{-0.5}, a_8^{-0.5})$. This transformation takes us from an unconstrained positive-definite and symmetric matrix $\widetilde{\Sigma}$, into a constrained covariance structure Σ with all odd positioned diagonal elements set to one to ensure identification, giving the following distribution:

$$p(\Sigma, a) = \frac{|\mathbf{S}|^{v/2}}{2^{4v}\Gamma_8(v/2)} |\Sigma|^{-(v+9)/2} exp\left\{-1/2tr(A^{-1}\mathbf{S}A^{-1}\Sigma^{-1})\right\} \prod_{i=1}^{4} a_i^{-v-1}, \tag{4.13}$$

Integrating the above equation (4.13) with respect to a yields the prior $p(\Sigma)$ as shown in equation (4.12). This prior is made weakly informative and proper by choosing v such that $v \geq 8$ following recommendations by McCulloch et al. (2000).

For the prior of the coefficient vector $\boldsymbol{\beta}$, implementation of the full Gibbs sampling scheme requires that we be able to switch between the scale of the restricted covariance matrix Σ and that of the unrestricted covariance matrix $\widetilde{\Sigma}$.

In the following, parameters with tilde relates to the unrestricted covariance scale, while those without tilde are defined with respect to the restricted covariance scale. Since the matrix A is diagonal, and the X_n is block diagonal with only one non-zero entry per column, pre-multiplying the utility equation by A gives

$$A\mathbf{L}_n = \widetilde{\mathbf{L}}_n = AX_n\boldsymbol{\beta} + A\boldsymbol{\epsilon}_n \tag{4.14}$$

$$= X_n\mathbf{A}\boldsymbol{\beta} + A\boldsymbol{\epsilon}_n \tag{4.15}$$

$$= X_n\widetilde{\boldsymbol{\beta}} + \widetilde{\boldsymbol{\epsilon}}_n, \tag{4.16}$$

where $\mathbf{A}$ is the expansion of A to the dimensions implied by β, that is

$$\mathbf{A} = diag(A_{ii}\mathbf{1}_s^{'}, A_{jj}\mathbf{1}_o^{'})' \quad \text{for } i = 1, 2$$

indexing the diagonal elements of the stated preference (selection) equations, and $j = 3, 4, 5, 6, 7, 8$ indexing the diagonal elements of the revealed choice (outcome) equations. Therefore the A_{ii} represent the $(i, i)^{th}$ entry of $\mathbf{A}$, and $\mathbf{1}_s^{'}$ and $\mathbf{1}_o^{'}$ are vectors of ones with lengths equal to the number of covariates in the selection and outcome processes respectively. From the priori independence of $\widetilde{\Sigma}$ and $\boldsymbol{\beta}$, and the identity provided in equation (4.14), $\widetilde{\boldsymbol{\beta}} = \mathbf{A}\boldsymbol{\beta}$ we get the conditional distribution $\widetilde{\boldsymbol{\beta}}|\widetilde{\Sigma} \sim N(\mathbf{A}\beta_0, \mathbf{A}B_0^{-1}\mathbf{A})$.

4.4.2 Steps of the sampler

Using the equalities

$$\widetilde{\boldsymbol{\beta}} = \mathbf{A}\boldsymbol{\beta} \quad \widetilde{\mathbf{L}}_n = A\mathbf{L}_n, \quad \widetilde{\Sigma} = A\Sigma A, \tag{4.17}$$

sampling can be done on either the unrestricted unidentified scale with "tilde," or on the restricted identified scale with "no tilde." The sampler goes through the following steps, transforming the parameter vector from the unidentified scale, $\widetilde{\boldsymbol{\theta}}$, to the identified scale $\boldsymbol{\theta}$ between steps 2 and 3 and steps 4 and 1.

1. sample $L_{nj}|(\mathbf{L}_{n,-j}, \Sigma, \boldsymbol{\beta}, A)$ for $n = 1, \ldots, N$ and $j = 1, \ldots, 8$, where $\mathbf{L}_{n,-j}$ is the vector of latent utilities, $\mathbf{L}_n$, with the j^{th} element removed.

2. Sample $(A|\Sigma, \boldsymbol{\beta}, \mathbf{L}) = (A|\Sigma)$ which is Gamma distributed

- transform to $\widetilde{\mathbf{L}}_n = A\mathbf{L}_n$

3. sample $(\widetilde{\boldsymbol{\beta}}|\widetilde{\mathbf{L}}, \Sigma, A)$

   - Record $\boldsymbol{\beta} = \mathbf{A}^{-1}\widetilde{\boldsymbol{\beta}}$

4. sample$(\widetilde{\Sigma}|\widetilde{\boldsymbol{\beta}}, \widetilde{\mathbf{L}})$

   - Transform $\widetilde{\Sigma}$ *to* (A, Σ)
   - Transform to $\mathbf{L} = A^{-1}\widetilde{\mathbf{L}}_n$

The sampler as described uses the partial marginalization principle by van Dyk (2010), which samples the working parameters (non-zero elements of A) twice per iteration, while maintaining a stable distribution at each step.

4.4.3 Posterior distribution

Following Burgette and Nordheim (2009), I begin with $p(\widetilde{\boldsymbol{\beta}}|\widetilde{\mathbf{L}}, \widetilde{\Sigma}, \mathbf{Y}) = p(\widetilde{\boldsymbol{\beta}}|\widetilde{\mathbf{L}}, \widetilde{\Sigma})$. Hence, from the density of , $\widetilde{\mathbf{L}}$, we have

$$
\begin{aligned}
p(\widetilde{\boldsymbol{\beta}}|\widetilde{\mathbf{L}}, \widetilde{\Sigma}) \propto \exp & \left\{ \frac{1}{2} \sum_n \left(\widetilde{\mathbf{L}}_n - X_n\widetilde{\boldsymbol{\beta}} \right)' \widetilde{\Sigma}^{-1} \left(\widetilde{\mathbf{L}}_n - X_n\widetilde{\boldsymbol{\beta}} \right) \right\} \\
\exp & \left\{ \frac{1}{2} \left(\widetilde{\boldsymbol{\beta}} - \widetilde{\beta}_0 \right)' \widetilde{B}_0 \left(\widetilde{\boldsymbol{\beta}} - \widetilde{\beta}_0 \right) \right\}
\end{aligned} \tag{4.18}
$$

which is a multivariate normal distribution with mean

$$
\widetilde{\boldsymbol{\beta}} = \left[\widetilde{B}_0 + \sum_n X_n' \widetilde{\Sigma}^{-1} X_n \right]^{-1} \left[\widetilde{B}_0 \widetilde{\beta}_0 + \sum_n X_n \widetilde{\Sigma}^{-1} \widetilde{\mathbf{L}}_n \right] \tag{4.19}
$$

and covariance $\left[\sum_n X_n' \widetilde{\Sigma}^{-1} X_n \right]^{-1}$.

Next, it can be noted that the posterior distribution of the unrestricted covariance matrix $\widetilde{\Sigma}$ conditional on the 8-dimensional latent vector of utility $\widetilde{\mathbf{L}}$ and the unrestricted parameters $\widetilde{\boldsymbol{\beta}}$ is inv-Wishart$(\widetilde{\mathbf{M}}, \mathbf{N}{+}v)$ with $\widetilde{\mathbf{M}} = \sum_n \left(\widetilde{\mathbf{L}}_n - X_n\widetilde{\boldsymbol{\beta}} \right) \left(\widetilde{\mathbf{L}}_n - X_n\widetilde{\boldsymbol{\beta}} \right)' + \mathbf{S}$.

Finally, the latent vector of utilities $\mathbf{L}_n$ is distributed as a truncated multivariate normal, conditional on (Y, θ), with a mean determined by X_n and $\boldsymbol{\beta}$, and a covariance structure Σ.

4.4.4 Identification of Σ

Recall from figure (4.1) that not all elements of the covariance matrix Σ are identifiable. This is because we only observe a single outcome for each respondent, therefore no information exists in the data to estimate the correlation between potential (but unobserved) outcomes. Here these unidentified parameters play the same role as working parameters (Albert and Chib, 1993), as they circumvent direct calculation of the likelihood function, and improve mixing properties of the Markov Chain, without affecting the inferential properties of the posterior distribution in so far that a proper prior is specified over them. To see how, consider the decomposition of the covariance matrix Σ, into its identifiable part $\boldsymbol{\psi}$ and a vector of unidentifiable part $\boldsymbol{\alpha}$, therefore we have $p(\mathbf{Y}|\boldsymbol{\psi}, \boldsymbol{\beta}), \boldsymbol{\alpha}) = p(\mathbf{Y}|\boldsymbol{\psi}, \boldsymbol{\beta})$, and

$$\begin{aligned}
p(\boldsymbol{\psi}, \boldsymbol{\beta}|\mathbf{Y}) &\propto p(\mathbf{Y}|\boldsymbol{\psi}, \boldsymbol{\beta})p(\boldsymbol{\psi}, \boldsymbol{\beta}) && (4.20)\\
&= p(\mathbf{Y}|\boldsymbol{\psi}, \boldsymbol{\beta}) \int p(\boldsymbol{\psi}, \boldsymbol{\beta}), \boldsymbol{\alpha}) d_{\boldsymbol{\alpha}} && (4.21)\\
&= \int p(\mathbf{Y}|\boldsymbol{\psi}, \boldsymbol{\beta})p(\boldsymbol{\psi}, \boldsymbol{\beta}), \boldsymbol{\alpha}) d_{\boldsymbol{\alpha}} && (4.22)\\
&= \int p(\mathbf{Y}|\boldsymbol{\psi}, \boldsymbol{\beta}), \boldsymbol{\alpha})p(\boldsymbol{\psi}, \boldsymbol{\beta}), \boldsymbol{\alpha}) d_{\boldsymbol{\alpha}}, && (4.23)
\end{aligned}$$

Where the last line corresponds to the Markov chain over the expended parameter space including α, and suggests that a proper representation of the prior for the covariance matrix Σ in equation (4.12) will be

$$p(\boldsymbol{\psi}) \propto \int |\Sigma|^{-(v+9)/2} \left(\prod_{i=1}^{4} tr(S_{ii}\Sigma_{ii}^{-1}) \right)^{-v} d_{\boldsymbol{\alpha}}, \qquad (4.24)$$

which is the prior representation in equation (4.13) marginalized over the unidentified parameters α. Because this prior representation is free from the unidentified part of the covariance matrix, the effect of the distributional assumption made about the unidentified parameters is fully contained in this prior.

4.5 Data and Variable Description

The empirical analysis is based upon data from the 2007 Medical Expenditure Panel Survey (MEPS) full year population characteristics data. The survey is sponsored by the Agency for Health Care Research and Quality(AHRQ), and designed to overlap two calendar years with a new Panel of sample households selected each year. The household component of the MEPS collects data from a subsample of the National Health Interview Survey and uses stratified and clustered random sampling with weights that produce nationally representative estimates for a wide range of health-related demographic and socioeconomic characteristics for the civilian, non institutionalized U.S. population. The data from the calendar year 2007 was collected in rounds 1, 2, and 3 for MEPS panel 12 and rounds 3, 4, and 5 for MEPS panel 11.

The survey includes questions on respondents' attitudes toward health insurance and health insurance cost, in a self-administered questionnaire (SAQ) which was administrated in round 2 for panel 12 and round 4 for panel 11. Although the 2007 MEPS includes 30964 individuals, interviewed over the 2-year period, the target population for the SAQ only include adults (person age 18 or older) in the civilian non institutionalized population amounting to 19067 respondents. After accounting for questionnaire non response, the final sample used in this analysis is comprised of 18035 individuals 18 to 85 that were member of the civilian, non-institutionalized portion of the U.S. population in 2007. For more information on the MEPS sampling design, see Ezzati-Rice et al. (2008).

The dependent variable in this study is a factor with three mutually exclusive and exhaustive categories representing the health insurance coverage indicator "INSURANCE." It is constructed from INSCOV07 provided in the MEPS dataset which has 3 levels (1–Some Private, 2–Public Only, 3–Uninsured). Because we wish to use "uninsured" as the base category in the estimation, the INSURANCE variable is constructed as a factor with three levels (0. Uninsured, 1. Some Private, 2.Public Only).

Of main interest is the role played by health insurance preference,"ATTHICW," in this choice process. The variable ,"ATTHICW," is constructed from the variable ADINSB42 provided in the MEPS dataset which is a factor with 5 levels (1. Disagree strongly, 2. Disagree somewhat, 3. Uncertain, 4. Agree somewhat, 5. Agree strongly), relating to the statement: Health insurance is not worth the cost. The variable is recoded into ATTHICW as a factor of 3 levels (0– Uncertain, 1– Agree, 2–Disagree) by combining the first two and last two categories of ADINSB42. This new variable is interpreted as (0– Uncertain, 1– Not worthy, 2–worthy) and represents the stated attitude towards health insurance cost. Since a respondent could interpret Uncertain as something other than indifference, we can consider the choice options to be unordered. The remaining covariates include demographic characteristics such as AGE, SEX, EDUCATION, INCOME; marital status; health characteristics; and regional dummies. Definitions and summary statistics for the covariates are given in table (4.1).

4.6 Results

The selection (health insurance preference) and the observed outcome (health insurance coverage) are guided by two separate but inter-related processes. In fact, it is assumed that personal income (INCOME) influences individuals' attitude toward health insurance (health insurance preference), while family income (FAMINC) influences the likelihood of falling in a given coverage category (health insurance outcome). As such, FAMINC is used here as a genuine exclusion restriction to ensure more robust identification of the estimated parameters (Heckman, 2000). This is motivated

Table 4.1: Summary statistics for the independent variables in the model

	N = 18035	Mean	SD
Demographic characteristics			
AGE	age of respondent in years	46.160	17.445
SEX	= 1 if respondent is female	0.542	0.498
MARRIED	= 1 if Currently married	0.562	0.496
COLLEGE	= 1 if at least one year of college	0.444	0.497
INCOME	Individual's income in 1000	29.998	31.471
FAMINC	Family's income in 1000	59.942	52.721
FAMSIZ	Number of family members	3.045	1.664
Health characteristics			
VERGOOD	= 1 if very good health	0.334	0.472
GOOD	= 1 if good health	0.326	0.469
FAIRPOOR	= 1 if fair or poor health	0.165	0.371
Regional dummies			
MIDWEST	= 1 if respondent is from the Midwest	0.209	0.407
NORTHEAST	= 1 if respondent is from the Northeast	0.150	0.357
WEST	= 1 if respondent is from the West	0.257	0.437
Variance estimation Var.			
VARSTR	Variance estimation stratum	534.3	489.89
VARPSU	Variance estimation PSU	1.664	0.627

by the fact that in expressing health insurance preference, adult respondents take into account subjective/personal information, while the actual observed coverage at the end of the year is affected by other family members and whether they have health insurance coverage that can be extended to the respondent.

The **R** package **endogMNP** is used to fit the model. Three Markov chains of length 22000 iterations were run, with the over dispersed default starting values, a burn-in period of 2000 iterations and a thinning interval of 5 iterations. To assess convergence of the chains, the **coda** package in **R** (Plummer et al., 2006)is used to compute the Gelman-Rubin convergence diagnostic (Gelman and Rubin, 1992). For the results presented, the Gelman statistics had values below 1.2 for all 138 estimated parameters, indicating acceptable convergence of the chains.

Table 4.2 shows the mean and variance of the marginal posterior distribution of the coefficients corresponding to the selection process. Looking at the coefficients on SEX, we see that females are less likely than males to express "not-worthy" compared to "uncertain" as their preference for health insurance (-0.0623), while more likely than males to express "worthy" compared to "uncertain" (0.675).

The coefficient value 0.0720 for MARRIED suggests that compared to unmarried adults, currently married individuals are more likely to express "worthy" compared to "uncertain" as their preference for health insurance. In relation to education, the coefficients on COLLEGE (0.0850 and 0.1655) suggest that adults with at least one year of college experience are more decisive in the expression of their health insurance preference (worthy or Not-worthy), compared to those with no college experience, whom tend to be less decisive (uncertain). The coefficients values (0.0034 and 0.0049) for INCOME suggest that an increase in personal income increases respondents' decisiveness in the expression of their health insurance preference (Not-worthy or Worthy over Uncertain).

The negative coefficients on the health characteristics dummy variables suggest

Table 4.2: Posterior means and standard deviations for the $\boldsymbol{\beta}$ parameters related to the insurance preference or selection process

	Not-Worthy (Atthicw=2)	Worthy (Atthicw=3)
INTERCEPT	0.1906**	0.2018**
	(0.0584)†	(0.0540)
AGE	0.0005	0.0076**
	(0.0008)	(0.0006)
SEX	−0.0623**	0.0675**
	(0.0259)	(0.0227)
MARRIED	0.1035	0.0720**
	(0.0284)	(0.0251)
COLLEGE	0.0850**	0.1655**
	(0.0260)	(0.0204)
INCOME	0.0034**	0.0049**
	(0.0004)	(0.0004)
VERGOOD	−0.1063**	−0.0210
	(0.0373)	(0.0296)
GOOD	−0.2134**	−0.0543
	(0.0401)	(0.0313)
FAIRPOOR	−0.2562**	−0.1139**
	(0.0453)	(0.0350)
MIDWEST	−0.0894**	−0.0235
	(0.0341)	(0.0270)
NORTHEAST	−0.0187	−0.0150
	(0.0396)	(0.0317)
WEST	−0.0987	−0.0695**
	(0.0296)	(0.0239)

† standard deviation of the parameter's posterior distribution in parentheses.

** indicates that zero is excluded from the 95% credible set.

that adults with EXCELLENT health conditions are more likely to find health insurance "Not-worthy" compared to adults with relatively less ideal health conditions (VERGOOD, GOOD, FAIRPOOR). This suggests that individuals with excellent health conditions do not find the need for health insurance as much as do those with very-good, good, or fair and poor health conditions. Finally the coefficients on the regional dummy variables suggest that relative to southerners, adults from the MIDWEST are less likely to express "Not-Worthy" compared to "uncertain," as their health insurance preference (-0.0894), while those from the WEST are less likely to express "Worthy" compared to "uncertain" (-0.0695).

Tables 3.3 and 3.4 provide coefficient estimates related to the outcome conditional on the preference categories. Table 4.3 summarizes estimates for "Public" coverage, while table 4.4 presents estimates for "Private" coverage. If adult respondents express health insurance preference based on the utility derived from such preference, then we should worry about self-selection bias, if we wish to predict health insurance coverage across all preferences for a given adult.

The results will be consistent with the presence of self-selection bias in the following sense. If modeling the distribution of a given coverage category conditional on each of the preference category and a set of covariates provides different estimates of the intercept, then the coverage outcome of interest is partly determined by the type of health insurance preference the respondent chooses to express. This dependence of the coverage outcome on the choice of preference by the adult respondent creates the self-selection bias, when we wish to predict health insurance coverage across all preferences for a given adult.

Looking at the intercept estimates in table 3.3, adults with weak preferences (Uncertain and Not-worthy) are less likely to be publicly covered only (-0.8652 and -0.440) compared to being uninsured, while individuals with strong preference for health insurance (worthy) are more likely (0.3458) to have public coverage only com-

Table 4.3: Posterior means and standard deviations for the $\boldsymbol{\beta}$ parameters related to choosing public coverage over being uninsured conditional on each stated preference

	Public|Uncertain (Insurance=2|Atthicw=1)	Public|Not-Worthy (Insurance=2|Atthicw=2)	Public|Worthy (Insurance=2|Atthicw=3)
CONST	-0.8652^{**}	-0.440^{**}	0.3458^{**}
	(0.1088)†	(0.1342)	(0.1166)
AGE	0.0062^{**}	0.0032	-0.0055
	(0.0016)	(0.0017)	(0.0028)
SEX	0.1306	0.0708^{**}	0.0389
	(0.0418)	(0.0323)	(0.0358)
MARRIED	0.0357	0.2145^{**}	0.3542^{**}
	(0.0540)	(0.0413)	(0.0455)
COLLEGE	0.2234^{**}	0.2218^{**}	0.3681^{**}
	(0.0524)	(0.0443)	(0.0383)
FAMINC	0.039^{**}	0.0064^{**}	0.0130^{**}
	(0.0010)	(0.0008)	(0.009)
FAMSIZ	-0.0681^{**}	-0.0923^{**}	-0.1883^{**}
	(0.168)	(0.0130)	(0.0125)
VERGOOD	0.1250^{**}	0.0779	0.0940^{**}
	(0.0577)	(0.0040)	(0.0423)
GOOD	0.0501	0.0465	-0.0239
	(0.0544)	(0.0386)	(0.0390)
FAIRPOOR	0.0474	-0.1186^{**}	-0.3162^{**}
	(0.0639)	(0.0544)	(0.0547)
MIDWEST	0.2244^{**}	0.2342^{**}	0.1864^{**}
	(0.0561)	(0.0473)	(0.0475)
NORTHEAST	0.2762^{**}	0.2133^{**}	0.0264
	(0.0661)	(0.0485)	(0.0769)
WEST	0.0621	0.0804^{**}	-0.0754
	(0.0465)	(0.0356)	(0.0435)

† standard deviation of the parameter's posterior distribution in parentheses.

** indicates that zero is excluded from the 95% credible set.

pared to being uninsured. These results are consistent with the existence of self-selection bias as described above. The effect of AGE on public coverage varies by health insurance preference and is significant only for individuals with "uncertain" preference. The coefficient value of (0.0062) suggests that an increase in age leads to increased likelihood of coverage through public insurance only over being uninsured, for individuals with "uncertain" health insurance preference. Also females with "Not-worthy" preference are more likely (0.0708) than their male counterparts to have public coverage only, compared to being uninsured. In addition currently married adults are more likely than those not currently married to be covered through public insurance only, over being uninsured. This observation is true for both individuals with "worthy" and "Not-worthy" preferences (0.3542 and 0.2145 respectively).

The effect of college education on coverage through public only, is significant across all health insurance preference categories. Although this effect is almost similar for adults with weak preference (uncertain, and Not-worthy) which are 0.2234 and 0.2218 respectively, it is relatively larger, 0.3681 for individuals with strong preference for health insurance (worthy). These coefficient values suggest that regardless of insurance preference, compared to adults with no college experience, those with at least one year of college experience are more likely to be publicly covered only, over being uninsured. The positive coefficient estimates for family income (FAMINC) across all health insurance preference suggests that an increase in family income increases the likelihood of being only publicly insured over being uninsured. The negative coefficient estimates however, across all insurance preferences for family size (FAMSIZ) suggests that an increase in family size decreases the likelihood of coverage through public only, over being uninsured. This less intuitive result may be explained by the fact that while increased family size may affect uninsured or private coverage status, depending on whether or not other family members have coverage that can be extended to the respondent, family size has no effect on public coverage status which is based solely on age and income requirement that must be met by the respondent.

With respect to health characteristics, relative to having an excellent health condition, adults with very good health conditions are more likely to be only publicly covered over being uninsured, when their preference for health insurance is either "uncertain" or "Worthy." On the other hand, relative to having an excellent health condition, adults with fair or poor health conditions are less likely to be only publicly covered over being uninsured, when their preference for health insurance is either "Not-worthy" or "worthy."

Looking at the regional dummy variables, the positive and significant coefficient values for MIDWEST across all preference categories suggest that relative to southerners, adults from the Midwest are more likely to choose public coverage over being uninsured irrespective of insurance preference. Similarly, relative to southerners, adults from the Northeast are more likely to choose public coverage over being uninsured, however they only do so when they have weak preference for health insurance (Uncertain, or Not-Worthy).

Now turning to the estimates for "Private" coverage conditional on all preference categories as summarized in table 3.4, interpretation is done as in table 3.3. The intercept values of -1.3010 and -1.9300 for adults with weak health insurance preference (Uncertain and Not-Worthy) suggest that individuals with such preferences are less likely to have any private coverage relative to being uninsured. On the other hand, the intercept value of (1.5640) for individuals with strong preference (Worthy) suggest that adults with such preferences are more likely to have some private coverage relative to being uninsured.

The effect of AGE on private coverage is positive and significant across all health insurance preference categories. This suggests that an increase in age increases the probability of having some private coverage over being uninsured, irrespective of health insurance preference. The positive and significant coefficient values for SEX

Table 4.4: Posterior means and standard deviations for the $\boldsymbol{\beta}$ parameters related to choosing private coverage over being uninsured conditional on all preference levels

	Private|Uncertain (Insurance=3|Atthicw=1)	Private|Not-Worthy (Insurance=3|Atthicw=2)	Private|Worthy (Insurance=3|Atthicw=3)
INTERCEPT	-1.3010^{**}	-1.9300^{**}	1.5640^{**}
	(0.1795)†	(0.1760)	(0.1571)
AGE	0.0291^{**}	0.0322^{**}	0.0216^{**}
	(0.0024)	(0.0025)	(0.0031)
SEX	0.2760^{**}	0.3710^{**}	0.0905^{**}
	(0.0610)	(0.0569)	(0.0537)
MARRIED	-0.2770^{**}	-0.4103^{**}	-0.3317^{**}
	(0.0806)	(0.0740)	(0.0334)
COLLEGE	0.2028^{**}	-0.1665^{**}	0.3444^{**}
	(0.0855)	(0.0613)	(0.0312)
FAMINC	0.0095^{**}	0.0094^{**}	0.0109^{**}
	(0.0014)	(0.0010)	(0.0006)
FAMSIZ	0.0519^{**}	0.1061^{**}	0.1090^{**}
	(0.0220)	(0.0183)	(0.0102)
VERGOOD	0.0035	0.2132^{**}	-0.0072
	(0.1147)	(0.0754)	(0.0551)
GOOD	-0.0543	0.4077^{**}	0.0479
	(0.1000)	(0.0948)	(0.0492)
FAIRPOOR	0.1528	0.7317^{**}	0.3793^{**}
	(0.1202)	(0.0964)	(0.0601)
MIDWEST	0.2040^{**}	0.3301^{**}	0.0139
	(0.0951)	(0.0829)	(0.0643)
NORTHEAST	0.5254^{**}	0.4031^{**}	0.3224^{**}
	(0.094)	(0.0877)	(0.1026)
WEST	0.2060^{**}	0.3005^{**}	0.2234^{**}
	(0.0803)	(0.0810)	(0.0617)

† standard deviation of the parameter's posterior distribution in parentheses.

** indicates that zero is excluded from the 95% credible set.

across all insurance preference categories suggests that females are more likely than males to have some private coverage over being uninsured, regardless of health insurance preference. On the other hand, the negative coefficients values for MARRIED across all preference categories suggest that currently married individuals are less likely than their unmarried counterpart to have some private coverage compared to being uninsured.

The direction of the effect of COLLEGE on private coverage varies across insurance preferences. Adults with at least one year of college experience are less likely than those with none to have some private coverage relative to being uninsured, when their preference for health insurance is "Not-worthy," but are more likely when their health insurance preference is "Uninsured" or "Worthy." An increase in family income (FAMINC) increases the likelihood of having some private coverage over being uninsured, irrespective of health insurance preference. This effect is relatively stronger however for individuals with the "worthy" preference. Similarly, an increase in family size (FAMSIZ) increases the likelihood of private coverage over being uninsured, across all insurance preference categories.

Looking at coefficient estimates for the health characteristics variables, we can say that relative to having an EXCELLENT health condition, adults with GOOD or VERY-GOOD health conditions are more likely to have some private coverage, but only when their preference for health insurance is "Not-Worthy." However, individuals with FAIR or POOR health condition, relative to adults with EXCELLENT health condition, are more likely to have some private coverage when they are more decisive in the expression of their health insurance preference (worthy or Not-worthy).

The coefficient estimates on the regional dummy variables suggest that relative to southerners, irrespective of health insurance preference, adults from NORTHEAST and WEST are more likely to have some private coverage over being uninsured. However, for adult respondents from the MIDWEST this is only true when they have a weak preference for health insurance (Uninsured or Not-worthy).

Although all identifiable elements of the variance covariance matrix are estimated, table 4.5 only summarizes the diagonal elements which correspond to the variances. All the estimated variance coefficients have related 95 percent posterior intervals not containing zero, suggesting their significance at the 5 percent level. Estimation of the variance covariance matrix although not of primary interest, allows the standard error of the estimated coefficients on the covariates, to reflect the correct variability, which also includes variability associated with the selection process.

4.7 Conclusion

This essay has concerned itself with modeling health insurance choices by adults in the U.S., using the 2007 MEPS dataset. More specifically, the basic research goal was to be able to say, for a set of covariates, how the insurance choice probabilities vary based on differing attitudes towards health insurance cost worthiness (Health insurance preference).

The results in fact suggested the existence of self-selection bias as coverage outcome was found to depend on the choice of health insurance preference made by the individual respondents. Overall the analysis extended the past literature by capturing the endogeneity of health insurance preference in the revealed coverage outcome process, while providing results consistent with the existing literature. In fact a major result in the literature is that individuals with weak preference (uncertain or Not-worthy), are less likely to be insured compared to uninsured, while individuals with strong preferences (worthy) are more likely to be insured compared to being uninsured (Monheit and Vistnes, 2008). The Fully Gibbsian Bayesian Multinomial Probit framework (Burgette and Nordheim, 2009) implemented in this essay produced similar findings, but also provided a more accurate measure of the effects of the covariates by accounting for the self- selection bias.

Table 4.5: Posterior means, standard deviation and 95 percent credible intervals for the diagonal elements of the covariance matrix

	Estimates	L-95 Percent CI	U-95 Percent CI
σ_{nw}	1.2735 (0.0330)†	1.2036	1.3430
σ_{w}	0.7265 (0.0330)	0.6566	0.7960
$\sigma_{p_u\|u}$	0.5017 (0.0901)	0.3793	0.8000
$\sigma_{p_r\|u}$	1.4983 (0.0901)	1.1997	1.6210
$\sigma_{p_u\|nw}$	0.5028 (0.1061)	0.3450	0.7880
$\sigma_{p_r\|nw}$	1.4973 (0.1061)	1.2120	1.6550
$\sigma_{p_u\|w}$	1.0939 (0.1107)	0.7670	1.2330
$\sigma_{p_r\|w}$	0.9062 (0.1107)	0.7671	1.2330

† standard deviation of the parameter's posterior distribution in parentheses.

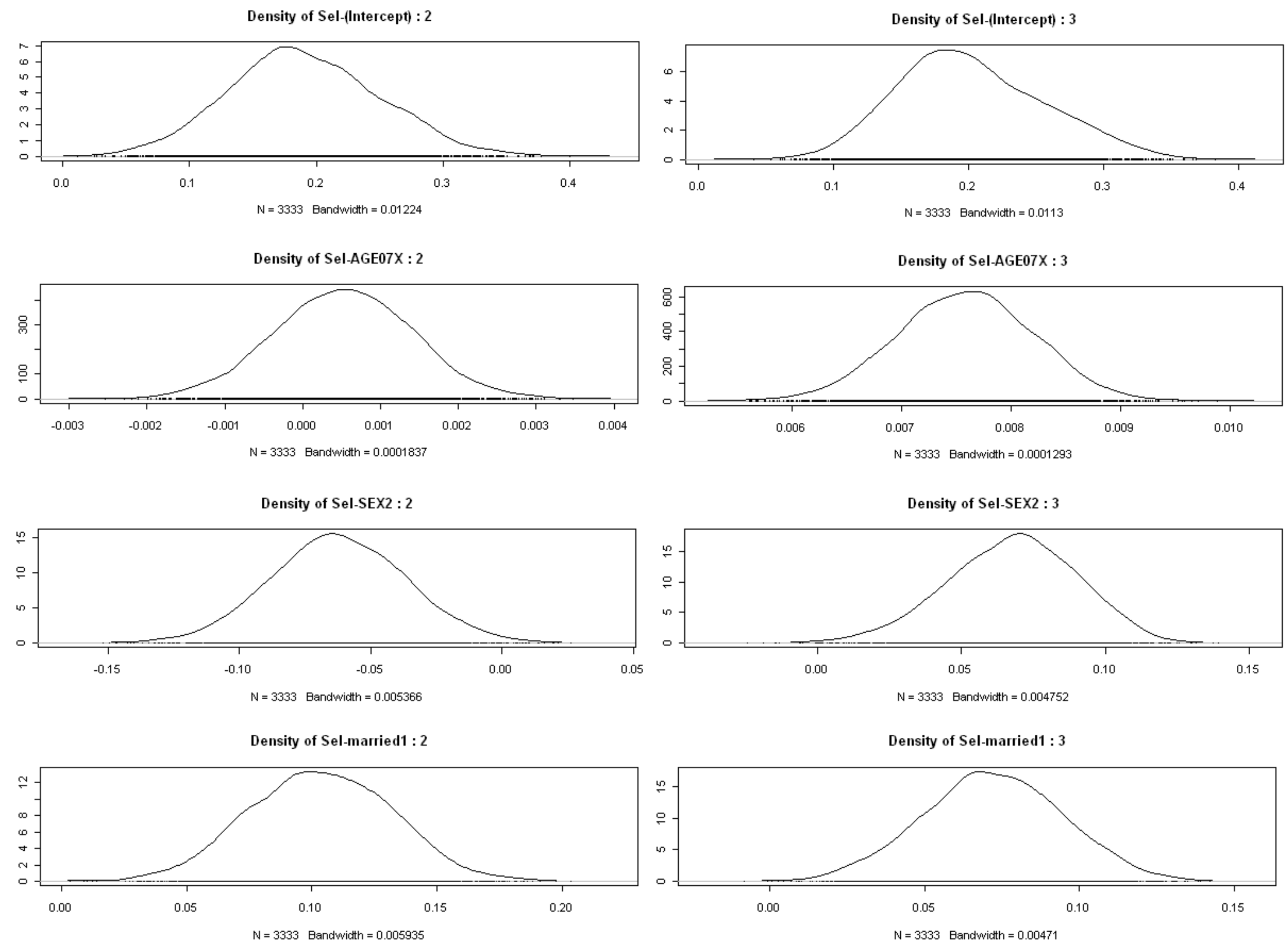

Figure 4.2: Density plots of posterior samples for the $\boldsymbol{\beta}$ parameters related to the selection process. On the left are densities of parameters that model "Not-worthy" preference compared to "uncertain". To the right are densities of parameters that model "Worthy" preference compared to "uncertain"

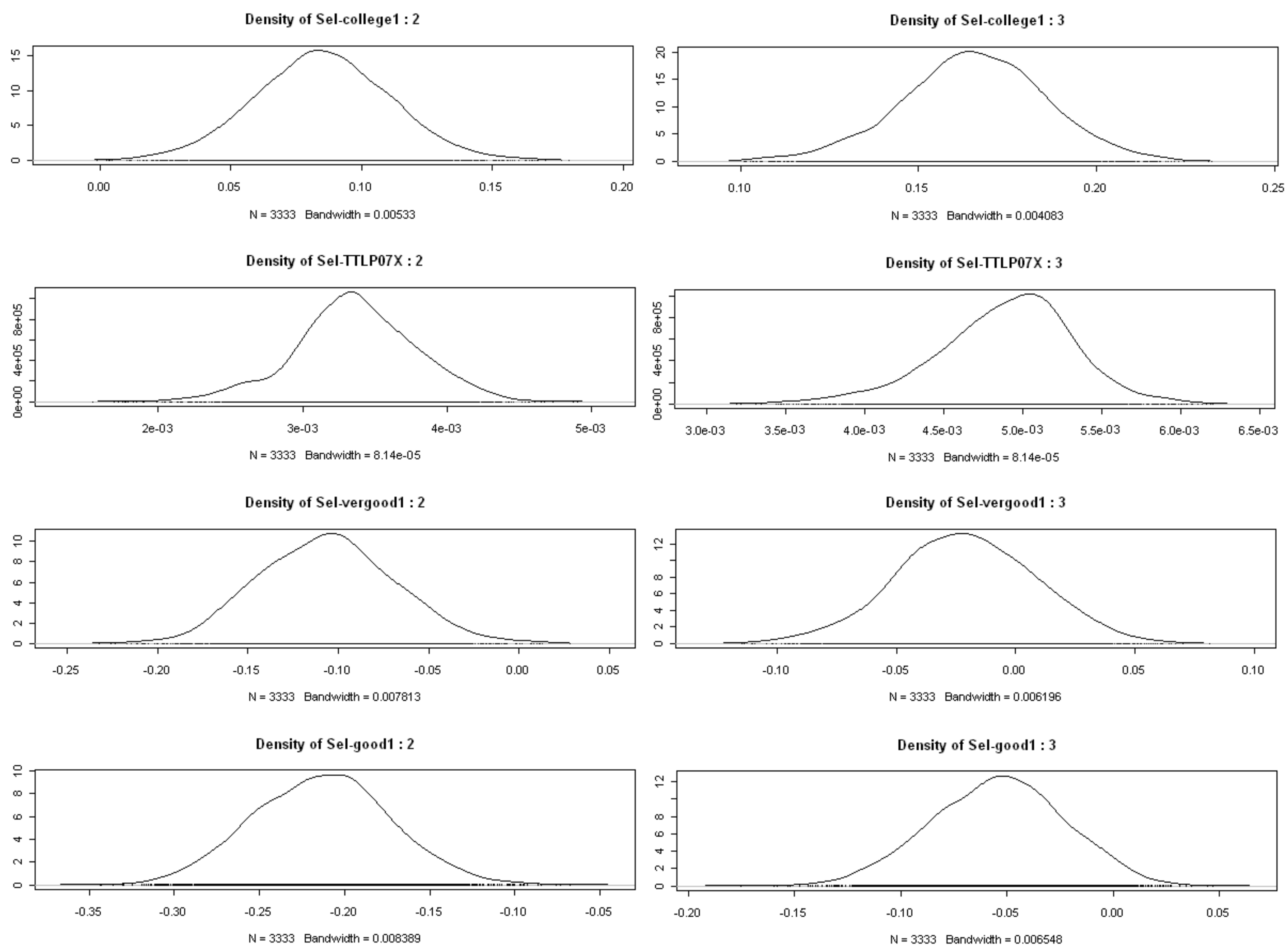

Figure 4.3: Density plots of posterior samples for the $\boldsymbol{\beta}$ parameters related to the selection process. On the left are densities of parameters that model "Not-worthy" preference compared to "uncertain". To the right are densities of parameters that model "Worthy" preference compared to "uncertain"

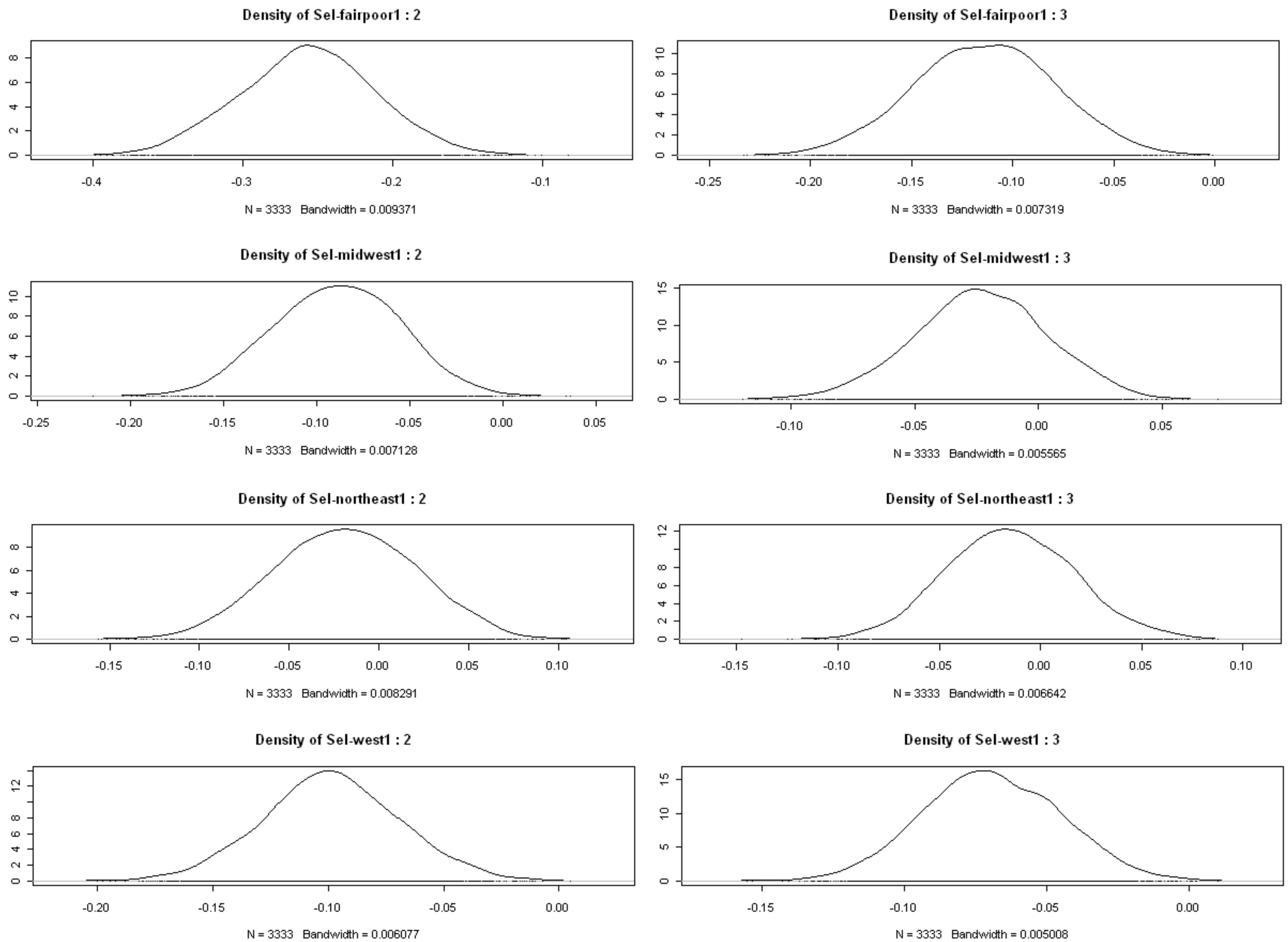

Figure 4.4: Density plots of posterior samples for the $\boldsymbol{\beta}$ parameters related to the selection process. On the left are densities of parameters that model "Not-worthy" preference compared to "uncertain". To the right are densities of parameters that model "Worthy" preference compared to "uncertain"

CHAPTER 5

Conclusion

The three essays in this dissertation considered issues associated with risk and decision-making under uncertainty. In the first essay, uncertainty was modeled analytically using mathematical tools from microeconomics based on the pioneering work of (Von-Neumann and Morgenstern, 1944). The other essays deal with empirical uncertainty that is modeled within the Bayesian econometric framework.

Specifically the first essay models the risks as experienced by untenured research faculty. Untenured tenure-track faculty members are given contracts, with up-or-out rules requiring some quantity and quality quota of research to be met, in order to gain tenure. The faculty member faces both uncertainty in monetary value of publication and output uncertainty. This essay extends the analysis of Chen and Lee (2009) by looking at ex-ante incentive properties of tenure from the perspective of a junior faculty member, rather than a department as accustomed in the principal-agent framework. The results suggest that risk significantly affects faculty output.

A risk averse researcher in the presence of price uncertainty, publishes at a point where the expected value of publication exceeds its marginal cost. Also a faculty member's scientific productivity is shown to be stimulated by increases in base salary when decreasing absolute risk aversion (DARA) described the faculty member's preference, while increased uncertainty in the value of publication provided a lesser incentive for scientific research output production. Moreover, when publication decision can be targeted at different journals, risk averse faculty had the economic incentive to target journals with negatively correlated per-unit returns.

In the second essay, rationality assumption is put to an empirical test. A rational agent expresses an attitude and then takes action that is consistent with that attitude. The proposition is tested within the context of health insurance choice by adults in the U.S., using the 2007 MEPS dataset. The evidence from the econometric estimation supports the notion that agents do act in a way that is consistent with stated preferences.

The results in this essay are consistent with those found by others, in particular with respect to individuals' satisfaction (or optimism) toward health insurance consumption. The results in fact suggest relatively more skepticism (dissatisfaction) toward public only coverage compared to having some private coverage. This is because the strength of unlikeliness to seek health insurance for those that expressed health insurance to not be worth its cost relative to being uncertain, is stronger for public only coverage compared to having some private coverage.

Finally, the third essay addressed the issue of health insurance preference endogeneity in adults' health insurance enrollment decision in the U.S. within a Bayesian multinomial probit framework. The research goal here was to be able to say for a set of covariates, how the health insurance outcome probabilities vary based on differing health insurance preferences. The **R** package `endogMNP` was used to fit the model, with the Gelman diagnostic statistics of less than 1.2 for all 138 estimated parameters, suggesting acceptable convergence of the chains. The sampler used the marginalization principle of van Dyk (2010), producing marginal posterior distributions that were easy to interpret. For example it was found that females were less likely than males to express "not-worthy" compared to "uncertain" as their preference for health insurance, while more likely than males to express "worthy" compared to "uncertain." In addition, the effect of college education on coverage through public only, is significant across all health insurance preference categories. Although this effect was almost similar for adults with uncertain, and Not-worthy preferences, 0.2234 and 0.2218 respectively, it was relatively larger, 0.3681 for individuals with strong preference for health insurance (worthy). These coefficient values suggested that regardless of insurance preference, compared to adults with no college experience, those with at least one year of college experience were more likely to be publicly covered only, over being uninsured. Also, looking at the regional dummy variables, positive and significant coefficient values across all preference categories suggested that relative to southerners, adults from the Midwest were more likely to choose public coverage over being uninsured irrespective of insurance preference. Similarly, relative to southerners, adults from the Northeast were more likely to choose public coverage over being uninsured, however they only do so when they have Uncertain, or Not-Worthy preferences. Finally, the coefficient estimates for the health characteristics variables suggested that relative to having an EXCELLENT health condition, adults with GOOD or VERY-GOOD health conditions were more likely to have some private coverage, but only when their preference for health insurance is "Not-Worthy." However, individuals with FAIR or POOR health conditions, relative to adults with EXCELLENT health condition, were more likely to have some private coverage when they are more decisive in the expression of their health insurance preference (worthy or Not-worthy).

BIBLIOGRAPHY

Albert, J.H. and S. Chib (1993), 'Bayesian analysis of binary and polychotomous response data', *Journal of the American Statistical Association* **88**, 669–679.

American-Diabetes-Association (2010), 'Standards of medical care in diabetes'. available at http://care.diabetesjournals.org/content/33/Supplement_1/S11.full.pdf+html.

Antle, John M. (1983), 'Testing the stochastic structure of production: A flexible moment-based approach', *Journal of Business and Economics Statistics* **1**, 192–201.

Arrow, J. Kenneth (1963), 'Uncertainty and welfare economics of medical care', *American Economic Review* **53**, 941–973.

Bech, M., D. Gurd-Hansen, T. Kjaer, J. Lauridsen and J. Sorensen (2006), 'Graded pairs comparison: Does strength of preference matter?', *Health Economics* **16**, 513–529.

Broder, Josef M. and Rod F. Ziemer (1982), 'Determinants of agricultural economics faculty salaries', *American Journal of Agricultural Economics* **64**, 301–303.

Brown, W. and D. Draper (2006), 'A comparison of bayesian and likelihood-based methods for fitting multilevel methods', *Bayesian Analysis* **1**(3), 473–514.

Brownstone, D. and Kenneth Train (1999), 'Forecasting new product penetration with flexible substitution patterns', *Journal of Econometrics* **89**, 109–129.

Burge, P., N. Devlin, J. Appleby, C. Rohr and J. Grant (2004), 'Do patients always prefer quicker treatment? a discrete choice analysis of patients' stated preferences in the London patient choice project', *Applied Health Economics and Health Policy* **3**(4), 183–194.

Burgette, Lane F. (2010), *endogMNP: R Package for Fitting Multinomial Probit Models with Endogenous Selection*. R package version 0.2-1.
URL: *http://CRAN.R-project.org/package=endogMNP*

Burgette, Lane F. and Erik V. Nordheim (2009), 'A full Gibbs sampler for a multinomial probit model with endogeneity'. http://www.stat.duke.edu/~lb131/switchBurgetteNordheim.pdf.

Burgette, Lane F. and Erik V. Nordheim (2010), 'A full Gibbs sampler for a multinomial probit model with endogeneity'. http://www.stat.duke.edu/~lb131/switchBurgetteNordheim.pdf.

Cameron, Colin A. and Parvin K. Trevedi (2005), *Microeconometrics Methods and Applications*, Cambridge University Press, New York.

Carmichael, H. Lorne (1988), 'Incentives in academics: Why is there tenure?', *Journal of Political Economy* **96**(3), 453–472.

Chavas, Jean-Paul (2004), *Risk Analysis in Theory and Practice*, Elsevier, Inc.

Chen, Zhao and Sang-Ho Lee (2009), 'Incentives in academic tenure under asymmetric information', *Economic Modelling* **26**(2), 300–308.

Chib, S. and B.H. Hamilton (2000), 'Bayesian analysis of cross-section and clustered data treatment models', *Journal of Econometrics* **97**(1), 25–50.

Congdon, Peter (2005), *Bayesian Models for Categorical Data*, John Wiley & Sons Ltd, NJ.

Cubanski, Juliette, Anthony Damico and Tricia Neuman (2009), 'Health care on a budget: An analysis of spending by medicare households', http://www.kff.org/medicare/7859.cfm.

David, Hensher A. and William H. Greene (2003), 'The mixed logit model: The state of practice', *Transportation* **30**, 133–176.

Davis, A. Timothy (2006), *Direct Methods for Sparse Linear Systems*, Fundamentals of Algorithms. SIAM, Philadelphia.

Dean, Regier A., Mandy Ryan, Euan Phimister and Carlo A. Marra (2009), 'Bayesian and classical estimation of mixed logit: An application to genetic testing', *Journal of Health Economics* **28**, 598–610.

DeGroot, Morris H. (1970), *Optimal Statistical Decisions*, McGraw-Hill, New York.

Ezzati-Rice, T.M., F. Rhode and J. Greenblatt (2008), 'Sample design of medical expenditure panel survey household component, 1998-2007', Methodology Report No. 22, Agency for Health care Research and Quality, Rockville, MD. Available at http://www.meps.ahrq.gov/mepsweb/data_files/publications/mr22/mr22.pdf.

Fahrmeir, Ludwig and Gerhard T. Tutz (2001), *Multivariate Statistical Modeling Based on Generalized Linear Models (2 ed.)*, Springer-Verlag, New York.

Faraway, Julian J. (2006), *Extending the Linear Model with R: Generalized Linear, Mixed Effects and Nonparametric Regression Models*, Chapman and Hall/CRC, NW.

Feldman, Roger, Bryan Dowd, Scott Leitz and Lynn Blewett (1997), 'The effect of premiums on the small firm's decision to offer health insurance', *Journal of Human Resources* **32**(4), 635–636.

Fernandez-Huerga (2008), 'The economic behavior of human beings: The institutionalist/postkeynesian model', *Journal of Economic Issues* **42**(3), 709–726.

Freeman, Smith (1977), 'Wage trends as performance displays productive potential: A model and application to academic early retirement', *The Bell Journal of Economics* **8**(2), 419–443.

Gelman, A and D.B. Rubin (1992), 'Inference from iterative simulations using multiple sequences (with discussion)', *Statistical Science* **7**, 457–472.

Gelman, A., J. B. Carlin, H. H. Stern and D. B. Rubin (2004), *Bayesian Data Analysis*, Chapman & Hall, 2nd edition.

Goldstein, Gerald S. and Mark V. Pauly (1976), Group health insurance as a local public good, *in* 'The Role of Health Insurance in the Health Services Sector', National Bureau of Economic Research, Inc., pp. 73–114.

Gracia-Cortes, L. A. and D. Sorensen (2001), 'Alternative implementations of monte carlo em algorithms for likelihood inferences', *Genetics Selection Evolution* **33**(4), 443–452.

Guttmann, Rochelle, Ryan Castle and Denzil G. Fiebig (2009), Use of discrete choice experiments in health economics: An update of the literature, Working Papers 2009/2, CHERE, University of Technology, Sydney.
URL: *http://ideas.repec.org/p/her/chewps/2009-2.html*

Haario, H., E. Saksman and J. Tamminen (2001), 'An adaptive metropolis algorithm', *Bernoulli* **7**(2), 223–242.

Hadfield, Jarrod D. (2010), 'Mcmc methods for multi-response generalized linear mixed models: The mcmcglmm r package', *Journal of Statistical Software* **33**(2), 1–22.

Hanson, K., B. McPake, P. Nakamba and L. Archard (2005), 'Preferences for hospital quality in zambia: Results from a discrete choice experiment', *Health Economics* **14**, 687–701.

Heckman, James J. (1979), 'Sample selection bias as specification error', *Econometrica* **47**(1), 153–161.

Heckman, James J. (2000), 'Causal parameters and policy analysis in economics: a twentieth century retrospective', *Quarterly Journal of Economics* **115**(1), 45–97.

Hess, S., Kenneth E. Train and J.W. Polak (2006), 'On the use of a modified latin hypercube sampling(mlhs) approach in the estimation of a mixed logit model for vehicle choice', *Transportation research Part B: Methodological* **40**(2), 147–163.

Hirsch, J.E. (2005), 'An index to quantify an individual's scientific research output', *Proc Natl Acad Sci USA* **102**(56), 16569–16572.

Imai, K. and D.A. van Dyk (2005), 'A bayesian analysis of the multinomial probit model using marginal data augmentation', *Journal of Econometrics* **124**(2), 311–334.

Jing, Song (2008), 'Tenure and asymmetric information: An analysis of an incentive institution for faculty development in research universities', *Front. Educ. China* **3**(2), 310–319.

Just, Richard E. and Rulon D. Pope (1978), 'Stochastic specification of production function and econometric specifications', *Journal of Econometrics* **7**, 67–86.

Just, Richard E. and Rulon D. Pope (1979), 'Production function estimation and related risk considerations', *American Journal of Agricultural Economics* **61**, 276–84.

Keane, Michael P. (2004), 'Modeling health insurance choice using the heterogeneous logit model', Manuscript, Yale University.

Kerssens, Jan J. and Peter P. Groenewegen (2005), 'Consumer preferences in social health insurance', *The European Journal of Health Economics* **50**(8), 8–15.

King, M.T., J. Hall, E. Lancsar, D. Fiebig, I. Hossain, J. Louviere, H.K. Reddel and C.R. Jenkins (2007), 'Patient preferences for managing asthma: results from a discrete choice experiment', *Health Econ* **16**(7), 703–717.

Korsgaard, I. R., A. H. Andersen and D. Soresen (1999), 'A useful reparameterization to obtain samples from conditional inverse wishart distributions', *Genetics Selection Evolution* **31**(2), 177–181.

Kou, Zonglai and Min Zhou (2009), Multi-tasking vs. screening: A model of academic tenure. Available at SSRN: http://ssrn.com/abstract=1483049.

Lamport, L. (1994), *LaTeXUser's Guide and Reference Manual*, 2nd edn, Addison Wesley.

Li, M. and J. Tobias (2005), 'Bayesian analysis of treatment effects in an ordered potential outcomes model'. Technical report, SUNY-Buffalo.

Manski, Charles F. and Daniel L. McFadden (1981), *Structural Analysis of Discrete Data and Econometric Applications*, The MIT Press, Cambridge.

Markowitz, Harry (1952), 'Portfolio selection', *Journal of Finance* **7**, 77–91.

Marshak, Jacob (1959), 'Binary choice constraints on random utility indications', Stanford Symposium on Mathematical Methods in the Social Sciences. available at http://www.cowles.econ.yale.edu/p/cp/p01b/p0155.pdf.

McCullagh, Peter and John Nelder (1989), *Generalized Linear Models (2 ed.)*, Chapman & Hall, London.

McCulloch, Charles E. and Shayle R. Searle (2001), *Generalized Linear, and Mixed Models*, New York.

McCulloch, R.E., N.G. Polson and P.E. Rossi (2000), 'A bayesian analysis of the multinomial probit model with fully identified parameters', *Journal of Econometrics* **99**(1), 173–193.

McFadden, Daniel and Kenneth Train (2000), 'Mixed mnl models of discrete response', *Journal of Applied Econometrics* **15**, 447–470.

McKenzie, Rishard B. (1996), 'In defense of academic tenure', *Journal of Institutional and Theoretical Economics* **152**, 325–341.

McPherson, Michael S. and Gordon C. Winston (1983), 'The economics of academic tenure: a relational perspective', *Journal of Economic Behavior and Organization* **4**(2), 163–184.

Mohamed, Gad-el-Hak (2004), 'Publish or perish—an ailing enterprise?', Physics Today. available at http://www.people.vcu.edu/ gadelhak/Opinion.pdf.

Monheit, Alan C. and P. Jessica Primoff Vistnes (1999), 'Health insurance availability at the workplace: How important are worker preferences?', *Journal of Human Resources* **34**(4), 770–785.

Monheit, Alan C. and P. Jessica Primoff Vistnes (2008), 'Health insurance enrollment decisions: Preferences for coverage, worker sorting, and insurance take-up.', *Inquiry* **45**(2), 153–167.

Munkin, R.E. and P.K. Trivedi (2008), 'Bayesian analysis of the ordered probit model with endogenous selection', *Journal of Econometrics* **143**(2), 334–348.

National-Diabetes-Statistics (2008), '[article online]', NIH Publication No. 083892. available at http://diabetes.niddk.nih.gov/dm/pubs/statistics/DM_Statistics.pdf.

Nelder, John and Robert Wedderburn (1972), 'Generalized linear models', *Journal of the Royal Statistical Society, Series A* **132**, 370–384.

Ovaskainen, O., H. Rekola, E. Meyke and E. Arjas (2008), 'Bayesian methods for analyzing movements in heterogeneous landscapes from mark-recapture data', *Ecology* **89**(2), 542–554.

Parente, Stephen T., Roger Feldman and Jon B. Christianson (2004), 'Employee choice of consumer-driven health insurance in a multiplan, multiproduct setting', *Heath services research* **39**(4), 1091–1111.

Peterson, A. Mark (2004), 'Editors note: Insurance coverage and the states.', *Journal of Health Politics, Policy and Law* **29**(2), 173–175.

Plummer, M., N. Best, K. Cowles and K. Vines (2005), 'Coda: Output analysis and diagnostics for mcmc', R package version 0.14.2. available at http://cran.r-project.org/web/packages/coda/index.html.

Plummer, Martyn, Nicky Best, Kate Cowles and Karen Vines (2006), 'Coda: Convergence diagnosis and output analysis for mcmc', *R News* **6**(1), 7–11.
URL: *http://CRAN.R-project.org/doc/Rnews/*

Raudenbush, Stephen and Anthony S. Bryk (2002), *Hierarchical Linear Models: Applications and Data Analysis Methods (2 ed.)*, Thousand Oaks, CA:Sage.

Remler, K. Dahlia, Jason E. Rachlin and Sherry A. Glied (2001), 'What can the take-up of other programs teach us about how to improve health insurance programs ?'. NBER Working Paper # 8185. March.

Roy, A. D. (1951), 'Some thoughts on the distribution of earnings', *Oxford Economic Papers Series* **3**, 135–146.

Ryan, M., D. Netten, D. Skatun and P. Smith (2006), 'Using discrete choice experiments to estimate a preference-based measure of outcome: An application to social care for older people', *Journal of Health Economics* **25**, 927–944.

Ryan, M. and J. Hughes (1997), 'Using conjoint analysis to assess women's preferences for miscarriage management', *Health Econ* **6**, 261–273.

Ryan, M. and K. Gerard (2003), 'Using discrete choice experiments to value health care', *Appl Health Econ Policy Analysis* **6**, 55–64.

Ryan, Mandy (2004), 'Using conjoint analysis to assess women's preferences for miscarriage management', *BMJ* **328**(7436), 360–361.

Sandmo, Agnar (1971), 'On the theory of competitive firm under price uncertainty', *American Economic Review* **61**, 65–73.

Schram, Sanford F. and Brian Caterino (2006), *Making Political Science Matter: Debating Knowledge, Research, and Method*, New York and London: New York University Press.

Srinivasan, K. and H. Mahamassani (2005), 'Dynamic kernel logit model for the analysis of longitude discrete choice data: Properties and computational assessment', *Transportation Science* **39**, 160–181.

Strauss, John and Duncan Thomas (2007), Health over the life course, *in* 'Handbook of Development Economics', Vol. 4, Elsevier, pp. 3375 – 3474.

Thurstone, Louis L. (1927), 'A law of comparative judgement', *Psychological Review* **34**, 273–286.

Train, Kenneth E. (2009), *Discrete Choice Methods with Simulation*, Cambridge University Press, New York.

van Dyk, David A. (2010), 'Marginal markov chain monte carlo methods', *Statistica Sinica* **20**, 1423–1454.

VonNeumann, John and Oscar Morgenstern (1944), *Theory of Games and Economic Behavior*, Princeton University Press, Princeton.

Zhao, Y., J. Staudenmayer, B. A. Coull and M. P. Wand (2006), 'General design bayesian generalized linear mixed models', *Statistical Science* **21**(1), 35–51.

VITA

Ibrahim Landolo Cheick Oumar Niankara

Candidate for the Degree of

Doctor of Philosophy

Dissertation: ESSAYS IN RISK AND APPLIED BAYESIAN ECONOMETRIC MODELING

Major Field: Economics

Biographical:

Personal Data: Born in Ouagadougou, Province of Kadiogo, Burkina Faso on April 12, 1982.

Education:
Received the B.S. degree from Oklahoma State University, Stillwater, Oklahoma, USA, 2007, in Economics with quantitative studies option.
Received the M.S. degree from Oklahoma State University, Stillwater, Oklahoma, USA, 2009, with a major in Economics
Completed the requirements for the degree of Doctor of Philosophy with a major Economics from Oklahoma State University in August, 2011.

Work Experience:
Fall 2008 - Spring 2011: Graduate Teaching Associates (OSU)

Publications:
Ibrahim L. Niankara.(2009). "The Role of Bandwagon and Snob Effects in Students Demand for Binge Drinking in College", *Journal of Global Business and Trade* 5(2), 65-78.

Conference Presentation:
2009: International Academy for Global Business and Trade, 6^{th} annual conference and Research Symposium (OSU)
2011: Eastern Economic Association (EEA), 37^{th} annual conference, Sheraton Hotel, New York City, NY

Name: Ibrahim L. C. O. Niankara — Date of Degree: December, 2011

Institution: Oklahoma State University — Location: Stillwater, Oklahoma

Title of Study: ESSAYS IN RISK AND APPLIED BAYESIAN ECONOMETRIC MODELING

Pages in Study: 123 — Candidate for the Degree of Doctor of Philosophy

Major Field: Economics

My dissertation consists of three essays modeling economic agents' optimizing behaviors under various conditions, with special focus on the role of risk and Bayesian prior updating. Each essay combines in a multidisciplinary fashion tools from various fields to uniquely model agent's behaviors.

The first essay is an analytical investigation of the impact of risk on publication decisions by untenured faculty members. It is motivated by the fact that an untenured member of a university faculty often faces pressure to publish research or lose his position in the institution. In the absence of a formal risk market, the member must manage exposure to this uncertainty privately. Combining tools from the physics literature with modeling tools from probability theory and the economics of risk this essay analyzes ex-ante incentive properties of tenure from the perspective of an untenured faculty member. The modeling principle used in this analysis is the first of its kind on the topic, and provides useful insights on the implicit cost of risk and its role in publication decisions.

The second essay investigates the consistency of revealed health insurance choices, with expressed attitudes towards health insurance cost, in an attempt to test for the rationality principle in the context of health insurance enrollment decisions by Medical Expenditure Panel Survey (MEPS) respondents. A new variant of the mixed logit model is used that relies on the multinomial logit formulation of the weighted logit formula. Current models in the literature are based on conditional logit representation.

The third essay models the effect of an individual's stated preferences for health insurance on their revealed choices of health insurance, using a framework in which stated preferences are assumed to be endogenous in the statistical sense. A discrete choice analysis is implemented where both stated preferences and choice outcomes are modeled through the multinomial probit and estimated in the Bayesian paradigm using a full Gibbs sampler. This estimation strategy is the first of its kind in the health economics literature, and improves computational efficiency and overall simplicity compared to related work using full information maximum likelihood. The analysis uses data from the 2007 MEPS.

ADVISOR'S APPROVAL: ______________________________