

2006

Essays on private information and choice

Dzmitry Asinski
Iowa State University

Follow this and additional works at: <https://lib.dr.iastate.edu/rtd>

 Part of the [Labor Economics Commons](#)

Recommended Citation

Asinski, Dzmitry, "Essays on private information and choice " (2006). *Retrospective Theses and Dissertations*. 1489.
<https://lib.dr.iastate.edu/rtd/1489>

This Dissertation is brought to you for free and open access by the Iowa State University Capstones, Theses and Dissertations at Iowa State University Digital Repository. It has been accepted for inclusion in Retrospective Theses and Dissertations by an authorized administrator of Iowa State University Digital Repository. For more information, please contact digirep@iastate.edu.

Essays on private information and choice

by

Dzmitry Asinski

A dissertation submitted to the graduate faculty
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Major: Economics

Program of Study Committee:
Brent Kreider, Major Professor

Philip Dixon
Joseph Herriges
Peter Orazem
Justin Tobias
Oscar Volij

Iowa State University

Ames, Iowa

2006

Copyright © Dzmitry Asinski, 2006. All rights reserved.

UMI Number: 3229050

INFORMATION TO USERS

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleed-through, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

UMI[®]

UMI Microform 3229050

Copyright 2006 by ProQuest Information and Learning Company.

All rights reserved. This microform edition is protected against unauthorized copying under Title 17, United States Code.

ProQuest Information and Learning Company
300 North Zeeb Road
P.O. Box 1346
Ann Arbor, MI 48106-1346

Graduate College
Iowa State University

This is to certify that the doctoral dissertation of
Dzmitry Asinski
has met the dissertation requirements of Iowa State University

Signature was redacted for privacy.

~~Committee~~ Member

Signature was redacted for privacy.

~~Committee~~ Member

Signature was redacted for privacy.

Committee Member

Signature was redacted for privacy.

Committee Member

Signature was redacted for privacy.

~~Committee~~ Member

Signature was redacted for privacy.

~~Major Professor~~

Signature was redacted for privacy.

For the Major Program

TABLE OF CONTENTS

LIST OF TABLES	v
LIST OF FIGURES	vi
CHAPTER 1. GENERAL INTRODUCTION	1
1.1 Introduction	1
1.2 Thesis Organization	2
CHAPTER 2. SELECTION AND INCENTIVE EFFECTS OF HEALTH INSURANCE IN THE PRESENCE OF RISK AVERSION	4
2.1 Introduction	5
2.2 The Model	9
2.2.1 General Outline	9
2.2.2 Estimation Details	13
2.3 Data	21
2.4 Results	24
2.5 Concluding Remarks	28
CHAPTER 3. CAPITAL STRUCTURE AND PERFORMANCE OF BUSI- NESS START-UPS: THE ROLE OF UNOBSERVED INFORMATION AND INCENTIVES	29
3.1 Introduction	30
3.2 Literature Review	32
3.2.1 Capital Structure Theories	32
3.2.2 Capital Structure and Performance	35

3.2.3	Models with Both Types of Informational Frictions	36
3.2.4	Selection and Moral Hazard in Lending and Insurance Markets	38
3.3	The Model	38
3.3.1	General Outline	38
3.3.2	Estimation Details	42
3.3.3	Identification	49
3.4	Data	51
3.5	Results	54
3.6	Concluding Remarks	58
CHAPTER 4. EMPLOYER-PROVIDED HEALTH INSURANCE: PREFERENCES AND CHOICE		59
4.1	Introduction	59
4.2	The Model	62
4.3	Simulation	65
4.4	Homogeneous Population Case	68
4.4.1	Maximum-Likelihood Estimation	68
4.4.2	Monte-Carlo Experiment	71
4.5	Directions for Future Research	72
CHAPTER 5. GENERAL CONCLUSION		74
APPENDIX A. Posterior distributions for chapter 2		76
APPENDIX B. Posterior distributions and posterior output for chapter 3		78
APPENDIX C. First order conditions for chapter 4		91
BIBLIOGRAPHY		93
ACKNOWLEDGEMENTS		100

LIST OF TABLES

Table 2.1	Variable definitions and descriptive statistics	23
Table 2.2	Variable definitions and descriptive statistics	24
Table 2.3	Results	25
Table 3.1	Variable definitions and descriptive statistics	55
Table 3.2	Variable definitions and descriptive statistics (cont'd)	56
Table 3.3	Results: Survival	57
Table 3.4	Results: Employment Growth	57
Table B.1	Results	82
Table B.2	Results	83
Table B.3	Results	84
Table B.4	Results	85
Table B.5	Results	86
Table B.6	Results	87
Table B.7	Results	88
Table B.8	Results	89
Table B.9	Results	90

LIST OF FIGURES

Figure 4.1	The trade-off between preferences and choice	66
Figure 4.2	The trade-off between preferences and choice	67
Figure 4.3	MLE, Homogeneous Population	72

CHAPTER 1. GENERAL INTRODUCTION

1.1 Introduction

The main goal of this dissertation is to explore the role played by private information in different markets. There are two main classes of imperfect information models that provide testable predictions for empirical work. The first, adverse selection, assumes that economic agents possess information about their characteristics not observed by other market participants. These characteristics directly affect both agents' own decisions and relevant market outcomes. For example, high risk people tend to choose insurance policies with low deductibles and coinsurance rates. At the same time, these people get in accidents more often. The second model, moral hazard, assumes that economic agents have incentives to behave differently when their actions are unobservable. For example, insured individuals tend to devote less effort to various preventive activities and by doing so they increase their own risk.

Both adverse selection and moral hazard predict that there should be positive correlation between some choice variable (insurance coverage) and some outcome variable (risk). Although the prediction of both of these models is the same, the mechanisms that generate it are different. In adverse selection models, agents know their risk and choose insurance coverage. In moral hazard models, agents change their own risk depending on the coverage they have. The welfare implications of these two asymmetric information problems and the ways to solve them are different. For example, mandatory insurance would solve the adverse selection problem but it would exacerbate the moral hazard problem. It is, therefore, important to empirically separate their effects.

1.2 Thesis Organization

In this dissertation, I concentrate on two broad areas: health insurance and business financing. The dissertation consists of three essays. The first two papers are empirical and use similar endogenous treatment models. They are estimated using Markov Chain Monte Carlo methods. In the third essay, I develop and simulate a structural job search model to illustrate the role played by unobserved preferences for health insurance and by availability of employer-provided insurance coverage.

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

The second essay (joint with Olena Chyruk) is the first paper to examine the relationship between capital structure and performance of business start-ups in the presence of imperfect information. Economic theory identifies two potential informational problems that affect the

trade-off between inside and outside financing. On the one hand, firms may self select into certain capital structures. For example, they may choose higher leverage based on favorable expected future prospects to avoid sharing the returns with outside equity holders. On the other hand, capital structure may affect the performance because of moral hazard. For example, outside equity dilutes ownership, and this decreases the entrepreneur's incentives to work. The goal of this essay is to investigate which effect (selection or incentive) dominates in the data. Capital structure and performance of business start-ups are estimated jointly using a unique data set collected by the National Federation of Independent Business (NFIB) Foundation. The results suggest that debt does not have a significant incentive effect on performance in business start-ups after controlling for self-selection. In contrast, both selection and incentive effects are present in the case of outside equity indicating that outside investors are able to both overcome informational opaqueness of business start-ups and provide better incentives for performance.

The third essay examines the role of unobserved preferences for health insurance in the context of job search. The majority of Americans get health insurance through their employers. One of the potential concerns related to employer-provided health insurance is whether some categories of workers (typically with low educational attainment) actually have access to jobs with health insurance offers. This challenges an assumption imposed in some of the previous literature that workers have equal access to both jobs with and without health insurance offers and sort strategically according to their preferences. A worker can be uninsured either because of lack of choice or because of preferences for a job without a health insurance offer. To be able to separate the effect of lack of choice from the effect of preferences, I develop a structural job search model with two job characteristics - the offered wage and the availability of insurance. I simulate the model to demonstrate the interaction between preferences and availability of choice.

CHAPTER 2. SELECTION AND INCENTIVE EFFECTS OF HEALTH INSURANCE IN THE PRESENCE OF RISK AVERSION

A paper to be submitted to *The Journal of Applied Econometrics*

Dzmitry Asinski

Abstract

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

and office-based health care utilization but not in inpatient or emergency room utilization.

2.1 Introduction

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

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

The empirical literature addressing informational asymmetries in insurance markets has traditionally been focused on testing for positive correlation, conditional on observables, between coverage and risk predicted by standard adverse selection and moral hazard models. Adverse selection, on the one hand, predicts that (unobservably) riskier individuals should be more likely to self-select into better coverage. Moral hazard, on the other hand, predicts that individuals with better coverage take fewer precautions because of the reduced penalties for risky behavior thereby increasing their risk level. Both types of informational asymmetries can cause positive correlation between coverage and risk in the data, yet the direction of causality is different. The empirical evidence on the existence of such correlation in health care and health insurance markets, especially coming from adverse selection, is not overwhelming.¹ A

¹The sizeable empirical literature attempting to test for asymmetric information and to separate effects of moral hazard and adverse selection produced diverse results with two important messages emerging from it

possible explanation for the absence of positive correlation between risk and coverage in some studies is provided by recent theoretical advances in contract theory. In particular, more complex asymmetric information models that include unobserved risk-aversion do not generally produce the positive correlation prediction generated by simpler theoretical models.²

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

– the choice of control variables and the population under consideration matter. For example, Cardon and Hendel (2001), Coulson and Stuart (1995), Dionne et al. (1998), Dionne et al. (2003), Dowd et al. (1991) find no evidence of adverse selection in different insurance markets after properly accounting for risk classification. Bradley (2002), Deb and Trivedi (2004), Ettner (1997), Holly et al. (1998), Hurd and McGarry (1997) do find such evidence.

²Indeed, Chiappori et al. (2002) demonstrate that positive correlation between risk and coverage persists when the basic adverse selection model of Rothschild and Stiglitz (1976) is extended to more general environments. In particular, a model with *both* adverse selection and moral hazard has this property. There are, however, two important dimensions in which the positive correlation property generally does not extend: non-competitive environments and unobserved risk-aversion.

selection,³ applied work has generally been hindered by the lack of data allowing researchers to separate different types of private information. Finkelstein and McGarry's study of long-term care insurance markets is the only study I am aware of that tests and finds evidence of different types of asymmetric information. They use an individual's subjective evaluation of the probability of using a nursing home within the next five years to show that, controlling for observables, it is a significant predictor of future nursing home use. In addition, the preventive care measures are found to be negatively correlated with subsequent nursing home use and positively correlated with the propensity to have long-term care insurance. The authors argue that these two types of private information cancel each other to produce an insignificant estimate of the overall selection effect.⁴

I follow Gardiol et al. (2003) in calling the two main effects of interest the *selection effect* and the *incentive effect*. The selection effect occurs if individuals select into different insurance states (plans) based on private information. The presence of the selection effect is a necessary but not sufficient condition for adverse selection in the market.⁵ The incentive effect is comprised of *ex-ante moral hazard* and *ex-post moral hazard*. The former effect is present if the probability of experiencing an adverse health event (sickness) is increased because of the reduced benefits from engaging in preventive activities. The latter effect appears because the insured individuals face a greatly reduced price of health care.⁶

In order to test for the presence of multiple sources of private information, I extend the standard two-equation endogenous treatment model along the lines of the general discrete choice model of Walker and Ben-Akiva (2002) by explicitly modeling indicators of latent risk-

³See, for example, Araujo and Moreira (2001), de Meza and Webb (2001).

⁴It is also interesting to note that Meer and Rosen (2004) estimate the demand for health care using instrumental variables (IV) techniques to control for potential endogeneity of health insurance status. Their main contribution to the literature is to introduce a new instrument – self-employment status. They find that after accounting for endogeneity, health insurance status is still an important predictor of health care utilization. What is more interesting is that compared to un-instrumented estimates, the IV estimates are larger. One possible explanation is that self-employed individuals are more likely to be more risk-loving than the general populace. It suggests that this particular instrument may control for some part of risk-aversion.

⁵The adverse selection is an equilibrium concept, and I only model individual behavior in this paper. Nevertheless, the term is often used in the applied literature to denote selection.

⁶It is also worth noting that from a theoretical standpoint, the classical moral hazard model requires some unobservable action by an agent which directly influences the outcome of interest. For example, an agent's engagement in preventive activities is unobserved, and these preventive activities can directly influence health status and health care expenditures.

aversion. The resulting hybrid model includes the endogenous treatment system and a latent variable measurement equation. The endogenous treatment model has been used extensively in the literature to separate the effects of moral hazard and adverse selection system, and it consists of an outcome equation (health care utilization, which serves as a proxy for risk) and a treatment (insurance coverage) equation. I estimate the resulting system of equations jointly using Markov Chain Monte Carlo (MCMC) methods. The choice of methodology is strongly motivated by the computational advantages of Bayesian methods over classical procedures in related models (see Munkin and Trivedi (2003)).

In this paper I use a unique set of variables available in the Medical Expenditure Panel Survey (MEPS) that measure individual responses to a series of questions about risk-taking behavior in general and attitudes toward health insurance in particular.⁷ I assume that these self-reported attitudes measure one dimension of private information available to the individuals – latent risk-aversion. One of the goals of this paper is to provide additional insights into what responses to these attitudinal questions actually convey in the context of medical care utilization and health insurance choice. The importance of such insights for future research is underlined by the resurgence of interest in the role played by individual preferences in making decisions related to health insurance.⁸

I use binary indicators for both medical care utilization and health insurance status. The most important advantage of the binary utilization and insurance variables is that they lend themselves naturally to the analysis of disparities in access to care by the insured and uninsured.⁹ A multinomial variable reflecting the choice of a contract from an available menu is an often used alternative. Unfortunately, the MEPS dataset I use does not contain reliable information on the actual menus that people face.¹⁰ Furthermore, studies analyzing the choice of a contract from an available menu typically restrict their samples to employed individuals

⁷The MEPS variables used in this paper first appeared in 2000. The precursor to the MEPS – the 1987 National Medical Expenditure Survey (NMES) – also contained these variables.

⁸For example, Monheit and Vistnes (2004) show that self-reported attitudes toward health insurance play an important role in job search.

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

¹⁰In fact, there is a restricted-use file in the MEPS containing information on the actual menu of contracts facing each employed individual. Unfortunately, a very high non-response rate renders these data unusable for the purposes of this paper.

with offers of multiple contracts, which affects the generality of their results.

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

The paper proceeds as follows. Section 2.2 outlines the econometric model and estimation details, Section 2.3 discusses the data, Section 2.4 provides results, and Section 2.5 concludes.

2.2 The Model

2.2.1 General Outline

Individuals make two choices in the model: whether to be insured and whether to utilize medical care. The two decisions can be thought of as corresponding to two distinct stages of decision making. In the first stage, individuals observe a private signal about their expected future health care needs and choose an insurance option according to this expectation. In the second stage, health care needs are realized and, conditional on chosen insurance status, the medical care decision is made. The important feature of the decision making process is the fact that the propensity to consume health care, which would obviously influence both the health insurance decision and the medical care consumption decision, is unobserved.

Both the selection and the incentive effects are expected to generate positive correlation between the presence of health insurance and health care utilization conditional on observables. The direction of causality is very different, however, for the two effects. The selection effect

induces people who expect to be relatively heavy users of medical care to select into the insured status. The incentive effect induces insured individuals to consume more health care, conditional on their health status, because they face a reduced price.

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

$$\begin{aligned} U_i &= 1[U_i^* > 0] \\ HI_i &= 1[HI_i^* > 0], \end{aligned} \tag{2.1}$$

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

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

$$RL_i = 1[RL_i^* > 0]. \tag{2.2}$$

The latent utilities governing the two decisions in the model are assumed to be equal to:¹²

$$\begin{aligned} U_i^* &= x_i\theta_1 + \gamma HI_i + \alpha_1 R_i^* + \zeta_1 RL_i^* + \epsilon_{1i} \\ HI_i^* &= x_i\theta_2 + z_i\nu + \alpha_2 R_i^* + \zeta_2 RL_i^* + \epsilon_{2i}, \end{aligned} \tag{2.3}$$

where x_i is a vector of exogenous variables; z_i is a vector of instruments;¹³ θ_1 and θ_2 are vectors

¹¹This reflects the way risk-aversion is measured in my data.

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

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

of parameters, α_1 , α_2 , ζ_1 , ζ_2 , ν and γ are scalar parameters; the error terms ϵ_{1i} , ϵ_{2i} are assumed to be independently and identically distributed $\sim N(0, \sigma_\epsilon)$; the latent risk type R_i^* is assumed to have standard normal distribution. Finally, I assume that the latent risk aversion RL_i^* is determined as follows

$$RL_i^* = x_i\delta + \eta_i, \quad (2.4)$$

where x_i is the vector of exogenous variables; δ is a scalar parameter; and the error term η_i is assumed to have a standard normal distribution. Note that although the errors ϵ_{1i} , ϵ_{2i} are uncorrelated, the problem of endogeneity of insurance choice still remains because of the presence of latent variables R_i^* and RL_i^* . Substituting the expression for the latent risk aversion RL_i^* into (2.3) I obtain the three equation system:

$$\begin{aligned} U_i^* &= x_i\beta_1 + \gamma HI_i + \alpha_1 R_i^* + \zeta_1 \eta_i + \epsilon_{1i} \\ HI_i^* &= x_i\beta_2 + z_i\nu + \alpha_2 R_i^* + \zeta_2 \eta_i + \epsilon_{2i} \\ RL_i^* &= x_i\delta + \eta_i, \end{aligned} \quad (2.5)$$

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

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

$$\Sigma = \begin{pmatrix} \alpha_1^2 + \zeta_1^2 + \sigma_\epsilon^2 & \alpha_1\alpha_2 + \zeta_1\zeta_2 & \zeta_1 \\ \alpha_1\alpha_2 + \zeta_1\zeta_2 & \alpha_2^2 + \zeta_2^2 + \sigma_\epsilon^2 & \zeta_2 \\ \zeta_1 & \zeta_2 & 1 \end{pmatrix} = \begin{pmatrix} 1 & \rho_{12} & \rho_{13} \\ \rho_{12} & 1 & \rho_{23} \\ \rho_{13} & \rho_{23} & 1 \end{pmatrix}, \quad (2.6)$$

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

model in more detail later.

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

If latent risk-aversion is an important factor in decision making, the two-equation endogenous treatment model can lead to incorrect conclusions about the existence of imperfect information. Without loss of generality, I assume that higher values of η_i reflect lower risk-aversion (higher RL_i^*). Then the theory predicts that $\zeta_1 > 0$ because more risk-loving individuals (higher η_i) will be less cautious and will be less actively engaged in prevention activities, thereby worsening health status and increasing the probability of using health care.¹⁴ At the same time, more risk-loving individuals are less likely to be insured, i.e. $\zeta_2 < 0$. It then follows that $\zeta_1\zeta_2 < 0$, which in turn implies that $sign(\rho_{12}) = sign(\alpha_1\alpha_2 + \zeta_1\zeta_2)$ is ambiguous. More specifically, the estimate of $\rho_{12} = \alpha_1\alpha_2 + \zeta_1\zeta_2$ can be very close to zero, implying that there is no private information and no selection effects. However, a conclusion based on $sign(\rho_{12})$ alone could clearly be erroneous as there is, in fact, private information ($\alpha_1\alpha_2 > 0$ and $\zeta_1\zeta_2 < 0$) of two different kinds that induce selection in the opposite directions.

The inclusion of an observed index of latent risk-aversion allows me to distinguish between two very different scenarios: (1) *no private information* and (2) *private information of multiple types*.¹⁵ This can be achieved by combining the information on the estimate of ρ_{12} with information on the estimates of ρ_{13} and ρ_{23} . Given the theoretical prediction discussed above,

¹⁴It is also conceivable that more risk averse individuals will be more likely to use health care because much of what we call prevention consists of frequent visits to a doctor to perform various check-ups. In this paper I consider various measures of health care utilization to try to distinguish between different possible effects of risk aversion on health care utilization.

¹⁵As a matter of practical implementation, it is not clear that the attitudinal questions used in this paper actually do measure risk aversion type and not, say, risk type or some combination of risk type and risk aversion type. Still, the important message of my argument is that the attitudinal questions do measure some facet of a potentially multifaceted private information set. By comparing correlations it is possible to reach certain conclusions about the presence of imperfect information of multiple types and also gain insights into what these questions are asking.

I should obtain $sign(\rho_{13}) = sign(\zeta_1) = +$ and $sign(\rho_{23}) = sign(\zeta_2) = -$. If the estimate of ρ_{12} is insignificant but the estimates of *both* ρ_{13} and ρ_{23} are significantly different from zero, I will conclude that there is private information instead of reporting no informational asymmetries.

The next subsection will outline estimation of the identified system of reduced-form equations given by¹⁶

$$\begin{aligned} U_i^* &= x_i\beta_1 + \gamma HI_i + \varepsilon_{1i} \\ HI_i^* &= x_i\beta_2 + z_i\nu + \varepsilon_{2i} \\ RL_i^* &= x_i\delta + \varepsilon_{3i}. \end{aligned} \tag{2.7}$$

The parameter γ measures the incentive effect of having health insurance on the probability of utilizing any health care. The correlations between the error terms of the three equations are designed to capture the unobserved private information about the latent risk type and risk-aversion type. The theoretical predictions of the signs of the error correlation matrix elements are as follows:

$$\Sigma = \begin{pmatrix} 1 & +? & + \\ +? & 1 & - \\ + & - & 1 \end{pmatrix}. \tag{2.8}$$

2.2.2 Estimation Details

2.2.2.1 Derivation of the Posterior Distribution

I use Bayesian methods to fit the model defined by 2.1, 2.2, and 2.7. The joint posterior distribution of the parameters of the model has to be simulated because it does not have a convenient analytical form. The parameter set is split into blocks and a variant of Gibbs sampling algorithm (see, for example, Tierney (1994)) is used to iteratively draw values from posterior distribution of each block of parameters conditional on other parameters of the model.

¹⁶An alternative approach would be to model the choice of insurance status and the choice of health care utilization with added explanatory variable RL_i . The coefficients on the latter would tell how observed indicators of latent factors affect the insurance status and the health care utilization decisions. However, such an approach ignores the potential endogeneity of self-reported indicators to the actual decisions made. It also ignores the potentially continuous nature of the latent risk-aversion index.

The posterior output of this Markov Chain is used to make inferences about the parameters of the interest. Because the conditional posterior distribution of error correlations is not of a standard form, a tailored Metropolis-Hastings algorithm is used to sample from it (Metropolis et al. (1953), Hastings (1970)). Lastly, I follow Albert and Chib (1993) by augmenting the parameter set with the latent data. The three equations for each individual are stacked in the following manner

$$y_i^* = \begin{pmatrix} U_i^* \\ HI_i^* \\ RL_i^* \end{pmatrix}_{3 \times 1}, y_i = \begin{pmatrix} U_i \\ HI_i \\ RL_i \end{pmatrix}_{3 \times 1}, \varepsilon_i = \begin{pmatrix} \varepsilon_{1i} \\ \varepsilon_{2i} \\ \varepsilon_{3i} \end{pmatrix}_{3 \times 1},$$

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

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

$$y_i^* = X_i \beta + \varepsilon_i$$

$$\varepsilon_i \sim N(0, \Sigma).$$

The observations are then stacked together as

$$y^* = \begin{pmatrix} y_1^* \\ y_2^* \\ \vdots \\ y_n^* \end{pmatrix}_{3n \times 1}, y = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}_{3n \times 1}, X = \begin{pmatrix} X_1 \\ X_2 \\ \vdots \\ X_n \end{pmatrix}_{3n \times k}, \varepsilon = \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{pmatrix}_{3n \times 1}$$

to produce

$$\begin{aligned} y^* &= X\beta + \varepsilon \\ &\sim N(X\beta, I_n \otimes \Sigma). \end{aligned} \quad (2.9)$$

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

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

where the second line follows from the assumed independence across individuals. Conditional on the parameters of the model, the augmented likelihood can be expressed as¹⁷

$$\begin{aligned} p(y^*|\beta, \Sigma) &= (2\pi)^{-\frac{3n}{2}} |I_n \otimes \Sigma|^{-\frac{1}{2}} \exp\left(-\frac{1}{2}(y^* - X\beta)'(I_n \otimes \Sigma)^{-1}(y^* - X\beta)\right) \\ &\propto |\Sigma|^{-\frac{n}{2}} \exp\left(-\frac{1}{2} \sum_{i=1}^n (y_i^* - X_i\beta)' \Sigma^{-1} (y_i^* - X_i\beta)\right). \end{aligned} \quad (2.11)$$

I place the following independent Normal prior distribution on β ¹⁸

$$\beta \sim N(\mu_{\beta 0}, V_{\beta 0}), \quad (2.12)$$

where $\mu_{\beta 0}$ and $V_{\beta 0}$ denote the prior mean and covariance matrix of β . Because identification requires the use of the normalized covariance matrix Σ , priors are placed directly on the

¹⁷Please refer to appendix for the details of derivations.

¹⁸I use proper priors for all parameters. The exact hyperparameters used in estimation are given in the next subsection.

identified correlation parameters ρ_{12} , ρ_{13} , and ρ_{23} , and are chosen to ensure that Σ is positive definite with probability 1. Specifically,

$$\begin{aligned}\rho_{12} &\sim UNIF(-1, 1) \\ \rho_{13} &\sim UNIF(-1, 1) \\ \rho_{23}|\rho_{12}, \rho_{13} &\sim UNIF\left(\rho_{12}\rho_{13} - \sqrt{(1 - \rho_{12}^2)(1 - \rho_{13}^2)}, \rho_{12}\rho_{13} + \sqrt{(1 - \rho_{12}^2)(1 - \rho_{13}^2)}\right).\end{aligned}\tag{2.13}$$

2.2.2.2 Posterior Simulation

The conditional posteriors of both β and Σ are proportional to the product of likelihood and the respective prior distribution. It turns out that the conditional posterior for β is also Normal:¹⁹

$$\begin{aligned}p(\beta|y^*, \Sigma) &\sim N(\mu_{\beta 1}, V_{\beta 1}) \\ V_{\beta 1} &= \left(\sum_{i=1}^n X_i' \Sigma^{-1} X_i + V_{\beta 0}^{-1}\right)^{-1} \\ \mu_{\beta 1} &= V_{\beta 1} \left(\sum_{i=1}^n X_i' \Sigma^{-1} y_i^* + V_{\beta 0}^{-1} \mu_{\beta 0}\right).\end{aligned}\tag{2.14}$$

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

$$\begin{aligned}p(\rho_{12}, \rho_{13}, \rho_{23}|y^*, \beta) &\propto |\Sigma|^{-\frac{n}{2}} \exp\left(-\frac{1}{2} \sum_{i=1}^n (y_i^* - X_i \beta)' \Sigma^{-1} (y_i^* - X_i \beta)\right) \times \\ &\frac{1}{\sqrt{(1 - \rho_{12}^2)(1 - \rho_{13}^2)}}.\end{aligned}\tag{2.15}$$

This distribution does not take any convenient standard form. To sample from (2.15), I employ a Metropolis-Hastings (MH) step within Gibbs algorithm. I choose the proposal density to be Normal centered at the sample correlation of errors, which are “known” given β and y^* . The variance of the proposal density is chosen so that the acceptance rate is approximately 35%.²⁰ Finally, the data augmentation step draws the values of latent variables U_i^* , HI_i^* , and RL_i^*

¹⁹The reader is referred to appendix for details.

²⁰The general rule of thumb for acceptance rate is about 44% when drawing one parameter and is about 23% when drawing a large number of parameters (see Koop (2003) and Train (2003)).

conditional on the observed data y_i and parameters of the model β, Σ . The distribution of latent data is truncated normal:

$$y_i^* | \beta, \Sigma, y \sim TN_{C(y_i)}(X_i \beta, \Sigma), \quad (2.16)$$

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

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

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

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

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

The details of the algorithm are as follows:

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

Step 1: Draw β^1 from the distribution given in (2.14) conditional on $(y_i^*)^0$ and Σ^0 (I use the following hyperparameters for the prior distributions: $\mu_{\beta 0} = 0$, $V_{\beta 0} = 1000 * I_k$).

Step 2: Draw the elements of the covariance matrix Σ using the Metropolis-Hastings algorithm:

- a. Let $\widetilde{\Sigma}^0 = \Sigma^0$;

- b. Compute the errors ε_i given the realized β^1 from *Step 1* and latent data $(y_i^*)^0$;
- c. Compute the sample correlations ρ_{ij}^{smp} between errors in equations i and j ;
- d. Draw a candidate value of $(\rho_{12}^{cand}, \rho_{13}^{cand}, \rho_{23}^{cand})$ from the proposal density $p^{prop}(\cdot|y^*, \beta)$, given by:²¹

$$\begin{aligned}\rho_{12} &\sim TN_{[-1,1]}(\rho_{12}^{smp}, \sigma_\rho^2) \\ \rho_{13} &\sim TN_{[-1,1]}(\rho_{13}^{smp}, \sigma_\rho^2) \\ \rho_{23} &\sim TN_{[\rho_{12}\rho_{13}-\sqrt{(1-\rho_{12}^2)(1-\rho_{13}^2)}, \rho_{12}\rho_{13}+\sqrt{(1-\rho_{12}^2)(1-\rho_{13}^2)}]}(\rho_{23}^{smp}, \sigma_\rho^2).\end{aligned}$$

Then construct the candidate covariance matrix Σ^{cand} as

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

- e. Accept the candidate covariance matrix Σ^{cand} and assign $\widetilde{\Sigma}^1 = \Sigma^{cand}$ with probability equal to

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

where $p(\cdot|y^*, \beta)$ is the conditional posterior density given in (2.15), otherwise assign $\widetilde{\Sigma}^1 = \widetilde{\Sigma}^0$;

- f. Repeat steps [d] and [e] T times and assign $\Sigma^1 = \widetilde{\Sigma}^T$.²²

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

- a. Compute the errors ε_{2i} and ε_{3i} given β^1 from *Step 1* and latent data $(HI_i^*)^0$ and

²¹The variance of the proposal density was chosen so that the acceptance rate was roughly 35%. To determine the average acceptance rate the algorithm was initially run with 100 MH step replications. Subsequently, the number of MH steps was reduced to 10.

²²The Metropolis-Hastings is repeated to achieve faster convergence.

$(RL_i^*)^0$;

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

$$TN_{[0,\infty)}(x_i\beta_1^1 + \gamma^1 HI_i + \begin{pmatrix} \rho_{12}^1 \\ \rho_{13}^1 \end{pmatrix} \begin{pmatrix} 1 & \rho_{23}^1 \\ \rho_{23}^1 & 1 \end{pmatrix}^{-1} \begin{pmatrix} \varepsilon_{2i} \\ \varepsilon_{3i} \end{pmatrix}, \\ 1 - \begin{pmatrix} \rho_{12}^1 \\ \rho_{13}^1 \end{pmatrix} \begin{pmatrix} 1 & \rho_{23}^1 \\ \rho_{23}^1 & 1 \end{pmatrix}^{-1} \begin{pmatrix} \rho_{12}^1 \\ \rho_{13}^1 \end{pmatrix}), \quad \text{if } U_i \geq 0.$$

$$TN_{(-\infty,0]}(x_i\beta_1^1 + \gamma^1 HI_i + \begin{pmatrix} \rho_{12}^1 \\ \rho_{13}^1 \end{pmatrix} \begin{pmatrix} 1 & \rho_{23}^1 \\ \rho_{23}^1 & 1 \end{pmatrix}^{-1} \begin{pmatrix} \varepsilon_{2i} \\ \varepsilon_{3i} \end{pmatrix}, \\ 1 - \begin{pmatrix} \rho_{12}^1 \\ \rho_{13}^1 \end{pmatrix} \begin{pmatrix} 1 & \rho_{23}^1 \\ \rho_{23}^1 & 1 \end{pmatrix}^{-1} \begin{pmatrix} \rho_{12}^1 \\ \rho_{13}^1 \end{pmatrix}), \quad \text{if } U_i \leq 0.$$

c. Compute the errors ε_{1i} given β^1 from *Step 1* and latent data $(U_i^*)^1$;

d. Draw $(HI_i^*)^1$ from

$$TN_{[0,\infty)}(x_i\beta_2^1 + z_i\nu^1 + \begin{pmatrix} \rho_{12}^1 \\ \rho_{23}^1 \end{pmatrix} \begin{pmatrix} 1 & \rho_{13}^1 \\ \rho_{13}^1 & 1 \end{pmatrix}^{-1} \begin{pmatrix} \varepsilon_{1i} \\ \varepsilon_{3i} \end{pmatrix}, \\ 1 - \begin{pmatrix} \rho_{12}^1 \\ \rho_{23}^1 \end{pmatrix} \begin{pmatrix} 1 & \rho_{13}^1 \\ \rho_{13}^1 & 1 \end{pmatrix}^{-1} \begin{pmatrix} \rho_{12}^1 \\ \rho_{23}^1 \end{pmatrix}), \quad \text{if } HI_i \geq 0.$$

$$TN_{(-\infty,0]}(x_i\beta_2^1 + z_i\nu^1 + \begin{pmatrix} \rho_{12}^1 \\ \rho_{23}^1 \end{pmatrix} \begin{pmatrix} 1 & \rho_{13}^1 \\ \rho_{13}^1 & 1 \end{pmatrix}^{-1} \begin{pmatrix} \varepsilon_{1i} \\ \varepsilon_{3i} \end{pmatrix}, \\ 1 - \begin{pmatrix} \rho_{12}^1 \\ \rho_{23}^1 \end{pmatrix} \begin{pmatrix} 1 & \rho_{13}^1 \\ \rho_{13}^1 & 1 \end{pmatrix}^{-1} \begin{pmatrix} \rho_{12}^1 \\ \rho_{23}^1 \end{pmatrix}), \quad \text{if } HI_i \leq 0.$$

- e. Compute the errors ε_{2i} given β^1 from *Step 1* and latent data $(HI_i^*)^1$;
- f. Draw $(RL_i^*)^1$ from

$$TN_{[0,\infty)}(x_i\delta^1 + \begin{pmatrix} \rho_{13}^1 \\ \rho_{23}^1 \end{pmatrix} \begin{pmatrix} 1 & \rho_{12}^1 \\ \rho_{12}^1 & 1 \end{pmatrix}^{-1} \begin{pmatrix} \varepsilon_{1i} \\ \varepsilon_{2i} \end{pmatrix}, \\ 1 - \begin{pmatrix} \rho_{13}^1 \\ \rho_{23}^1 \end{pmatrix} \begin{pmatrix} 1 & \rho_{12}^1 \\ \rho_{12}^1 & 1 \end{pmatrix}^{-1} \begin{pmatrix} \rho_{13}^1 \\ \rho_{23}^1 \end{pmatrix}), \quad \text{if } RL_i \geq 0.$$

$$TN_{(-\infty,0]}(x_i\delta^1 + \begin{pmatrix} \rho_{13}^1 \\ \rho_{23}^1 \end{pmatrix} \begin{pmatrix} 1 & \rho_{12}^1 \\ \rho_{12}^1 & 1 \end{pmatrix}^{-1} \begin{pmatrix} \varepsilon_{1i} \\ \varepsilon_{2i} \end{pmatrix}, \\ 1 - \begin{pmatrix} \rho_{13}^1 \\ \rho_{23}^1 \end{pmatrix} \begin{pmatrix} 1 & \rho_{12}^1 \\ \rho_{12}^1 & 1 \end{pmatrix}^{-1} \begin{pmatrix} \rho_{13}^1 \\ \rho_{23}^1 \end{pmatrix}), \quad \text{if } RL_i \leq 0.$$

step 5: repeat steps 1-4 S times.

The Gibbs algorithm generates a sample of size S from *conditional* posterior distribution of each of the parameters of the model. The first S_0 draws are discarded as burn-in because the Markov Chain has to converge to the *joint* posterior distribution of the parameters of the model. The remaining S_1 draws constitute the sample from the joint posterior distribution used for the analysis. I obtained 25,000 draws from the posterior distribution. The first 5,000 were discarded as burn-in, and the remaining 20,000 were used for analysis.

I performed some diagnostics of the Markov Chains used in this paper. The first-order autocorrelation coefficient for ρ_{12} was about 0.93 (it takes about 60 iterations for the effect of a shock to disappear). The first-order autocorrelation coefficients for the remaining parameters of interest (ρ_{12} , ρ_{12} , and γ) were substantially lower at around 0.74, 0.73, and 0.77, respectively. Although the ρ_{12} was mixing slowly, the large number of post-burn-in replications (20,000) insures that the sampler covers the posterior distribution sufficiently well. The values of

Gelman-Rubin R statistic were lower than 1.015 for all parameters of interest.²³ These values are within the usual 1.2 cut-off value suggesting that the samplers converged well. The Gelman-Rubin statistics were obtained by running the Gibbs sampler 50 times with overdispersed starting values of all the parameters of the model (10 values were chosen manually to insure that extremes were covered, the remaining 40 runs were started at the values drawn from a normal distribution centered at zero with a large variance for β and drawn from the uniform distributions for the correlation parameters).

2.3 Data

I use the data from the Household Component (HC) of the Medical Expenditure Panel Survey (MEPS), a national survey of the US civilian non-institutionalized population administered by the Agency for Healthcare Research and Quality (AHRQ). The HC contains detailed information about individuals' demographics, employment, income, health, health insurance status, and health care utilization. Although the MEPS is a panel survey, its longitudinal dimension is quite short. Respondents are interviewed five times over the course of two and a half years. It is an ongoing survey that began in 1996, with new panels introduced each year, resulting in annual files containing overlapping panels. I use the 2000 MEPS, the first year to contain a series of attitudinal questions that were asked in the paper-and-pencil Self-Administered Questionnaire (SAQ). There are four agree-disagree (ranging from 1-disagree strongly to 5-agree strongly) questions:²⁴

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

The sample is restricted to the adult non-elderly population aged 18 to 64.²⁵ Following

²³The values of Gelman-Rubin statistic were lower for ρ_{13} and ρ_{23} . This is likely due to the fact that third equation does not have endogeneity links to the first two equations.

²⁴In the econometric implementation, I group individuals whose answers were 'strongly disagree' and 'disagree somewhat' into one class and the rest of responses into the other.

²⁵Age is computed as of July 1, 2000, the appropriate age for any analysis involving variables from SAQ.

Monheit and Vistnes (2004) who used the same variables, I also exclude full-time students. Full-time students, children, and the elderly are likely to face substantially different choice sets regarding the insurance status. I also restrict the sample to individuals who were either insured the whole year or uninsured the whole year. The temporarily uninsured have been found to differ from both the full-year insured and the full-year uninsured in their health care utilization patterns (see Li and Trivedi (2004)). The group of part-year uninsured is highly heterogeneous and, therefore, probably deserves special attention, which is beyond the scope of this paper. Following Monheit and Vistnes (2004), I also exclude individuals whose answers to the SAQ questionnaire were provided by a proxy (most often a spouse). After deleting observations with missing values, I arrived at a sample of 7,967 individuals. Definitions of all variables along with summary statistics are given in Tables 2.1 and 2.2.

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

The set of controls used in all three equations consists of demographic variables, education, family income, employment status, regional and MSA dummies, and a variety of measures of health status. It is well-known that in endogenous treatment models the treatment parameter γ is not non-parametrically identified. Identification in my model can be achieved via the nonlinear functional form of the model, but for robustness I also impose a set of exclusion restrictions suggested by Li and Trivedi (2004). In particular, in the insurance equation I use a set of variables related to the demographic characteristics and health condition of dependents like children and spouse: spouse's age (SPAGE), spouse's education (SPEDU), a dummy for

Table 2.1 Variable definitions and descriptive statistics

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

Table 2.2 Variable definitions and descriptive statistics

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

spouse's any priority conditions (SPRICOND), and a dummy for presence of a child who easily gets sick (SICCHILD).²⁶

2.4 Results

Posterior means and standard deviations of the parameters of interest (correlation coefficients and the endogenous treatment parameter γ) are given in Table 2.3.²⁷

Two robust results emerge from Table 2.3. First, the posterior means and standard deviations of the treatment parameter γ suggest that the probability of utilizing office-based and outpatient services is strongly affected by insurance status. At the same time, consistent with previous literature, I find no evidence of an incentive effect of health insurance in the consumption of inpatient or emergency room services. Second, the estimates of the covariance matrix Σ do indicate the presence of private information. Although the overall selection effect, as measured by the correlation between unobservables in the utilization and insurance

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

²⁷The posterior statistics for all the remaining parameters of the model are not reported here to save space but are available upon request.

Table 2.3 Results

dependent variables	γ	ρ_{12}	ρ_{13}	ρ_{23}
OBVISIT, INSURED, NONEEDHI	0.6811*** (0.1811)	-0.0263 (0.1030)	-0.1507*** (0.0282)	-0.2571*** (0.0256)
OBVISIT, INSURED, HINOTWRTH	0.6290*** (0.1635)	0.0044 (0.0925)	-0.0635*** (0.0238)	-0.2721*** (0.0220)
OBVISIT, INSURED, TAKERISK	0.7024*** (0.1714)	-0.0393 (0.0975)	-0.0601*** (0.0218)	-0.1416*** (0.0228)
OBVISIT, INSURED, OVRCMILL	0.7252*** (0.1832)	-0.0514 (0.1042)	-0.1062*** (0.0218)	-0.1285*** (0.0239)
OPVISIT, INSURED, NONEEDHI	0.5838*** (0.2224)	-0.1805 (0.1294)	-0.0851** (0.0366)	-0.2573*** (0.0253)
OPVISIT, INSURED, HINOTWRTH	0.4820** (0.2173)	-0.1189 (0.1244)	-0.0592** (0.0283)	-0.2726*** (0.0223)
OPVISIT, INSURED, TAKERISK	0.5253** (0.2372)	-0.1444 (0.1355)	-0.0401 (0.0250)	-0.1420*** (0.0232)
OPVISIT, INSURED, OVRCMILL	0.5630*** (0.2132)	-0.1673 (0.1230)	-0.0746*** (0.0250)	-0.1277*** (0.0239)
ERVISIT, INSURED, NONEEDHI	-0.0358 (0.2397)	0.0485 (0.1356)	-0.0857** (0.0389)	-0.2556*** (0.0254)
ERVISIT, INSURED, HINOTWRTH	-0.0619 (0.2302)	0.0632 (0.1296)	-0.0620** (0.0304)	-0.2722*** (0.0220)
ERVISIT, INSURED, TAKERISK	-0.0296 (0.2348)	0.0442 (0.1330)	0.0075 (0.0272)	-0.1414*** (0.0230)
ERVISIT, INSURED, OVRCMILL	-0.0257 (0.2400)	0.0422 (0.1353)	-0.0680*** (0.0281)	-0.1280*** (0.0237)
IPDIS, INSURED, NONEEDHI	0.1721 (0.2948)	0.1158 (0.1627)	-0.1027** (0.0445)	-0.2561*** (0.0256)
IPDIS, INSURED, HINOTWRTH	0.1393 (0.2615)	0.1368 (0.1444)	-0.1112*** (0.0334)	-0.2720*** (0.0217)
IPDIS, INSURED, TAKERISK	0.0539 (0.3030)	0.1802 (0.1665)	-0.0972*** (0.0312)	-0.1414*** (0.0232)
IPDIS, INSURED, OVRCMILL	0.1297 (0.2560)	0.1412 (0.1423)	-0.1020*** (0.0311)	-0.1277*** (0.0235)

Notes: Standard errors are displayed in parentheses below coefficients; *** - significant at 1%,

** - significant at 5%.

equations, ρ_{12} , is statistically not distinguishable from zero, the estimates of both ρ_{13} and ρ_{23} are negative and significant for most of the specifications. The presence of selection on private information would not have been detected in the simpler endogenous treatment model without incorporating the indicators of latent information.

The summary of results reported in Table 2.3 along with the theoretical predictions from Section 2.2 are given below:

$$\Sigma(\text{hypothesis}) = \begin{pmatrix} 1 & +? & + \\ +? & 1 & - \\ + & - & 1 \end{pmatrix}, \quad \Sigma(\text{estimated}) = \begin{pmatrix} 1 & 0 & - \\ 0 & 1 & - \\ - & - & 1 \end{pmatrix}. \quad (2.17)$$

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

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

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

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

TAKERISK provides the most direct measure of risk-aversion (not contaminated by attitudes toward other things).

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

2.5 Concluding Remarks

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

CHAPTER 3. CAPITAL STRUCTURE AND PERFORMANCE OF
BUSINESS START-UPS: THE ROLE OF UNOBSERVED
INFORMATION AND INCENTIVES

A paper to be submitted to *The Journal of Finance*

Dzmitry Asinski¹, Olena Chyruk²

Abstract

This essay is the first paper to examine the relationship between capital structure and performance of business start-ups in the presence of imperfect information. Economic theory identifies two potential effects caused by imperfect information that determine the relationship between capital structure and performance. On the one hand, capital structure may affect performance because of moral hazard (incentive effect). For example, outside equity dilutes ownership and thus decreases an entrepreneur's incentives to exert effort. On the other hand, firms may self select into certain capital structures based on private information about their future expected success (selection effect). For example, they may choose higher leverage based on favorable future prospects to avoid sharing the returns with outside equity holders. In this paper we investigate which effect dominates in the data by specifying an econometric model to jointly estimate capital structure and performance of business start-ups. We estimate the model using a Gibbs sampling algorithm with data augmentation. We use a unique data set collected by the National Federation of Independent Business (NFIB) Foundation. Our results suggest that debt does not have a significant incentive effect on performance in business start-ups after controlling for self-selection. In contrast, both selection and incentive effects are present in the

¹Primary author.

²Graduate Student, The University of Iowa.

case of outside equity indicating that outside investors are able to both overcome informational opaqueness of business start-ups and provide better incentives for performance.

3.1 Introduction

The relationship between business financing (capital structure) decisions and performance of companies is determined by the presence of two potential effects – selection and incentive effects. Both of them are caused by imperfect information. The selection effect happens if entrepreneurs possess private information about characteristics of their own business ideas that are relevant both to the expected success of their enterprises and to their business financing decisions.³ For example, if an entrepreneur expects her business to be very successful in the future, she would prefer to engage debt as opposed to external equity to retain all the residual profits after repaying the loan. The incentive effect (moral hazard) happens if an entrepreneur's actions (for example, effort) are not observed. Economic theory suggests that different capital structures create different incentives for entrepreneurs. For example, the dilution of an entrepreneur's ownership through outside equity may decrease her incentives to exert effort because she has to share the returns with other stakeholders. Lower effort, in turn, translates into a weaker performance of her firm. Thus, both of these effects can induce correlation between capital structure and performance in the data. The mechanisms that generate this correlation are completely different for each effect.

The main goal of this paper is to test whether firms self-select into certain capital structures based on private information, and whether the observed capital structures create incentive effects for performance of business start-ups. Our research contributes to the literature in several ways. To our knowledge, this is the first paper to test for both selection and incentive effects of capital structure of business start-ups. The issues that we study are especially relevant to small businesses, which are much more informationally opaque than established

³The examples of such characteristics may be indicators of quality, productivity, or risk properties of a business project that are known only to the entrepreneur (Darrough and Stoughton, 1986). For example, Thompson (2005) uses a data set of firms in the US shipbuilding industry in 19th century to show that after controlling for firms' quality, survival does not depend on a firm's age. In other words, success of firms measured by survival over time is driven by initial unobservable quality characteristics rather than by learning-by-doing or other competing explanation.

corporations. We also contribute to the empirical literature that evaluates existing capital structure and agency theories. While these theories were extensively tested for established publicly traded firms, there are only a few papers that study capital structure of start-up firms that are not plagued by survivorship bias or limited to certain regions and industries (see, for example, Cassar (2004)).⁴ In addition, we also contribute to the empirical literature by proposing new instruments for the effect of capital structure on performance that are consistent with theories of entrepreneurship and lending. Lastly, the results we obtain in this paper suggest possible directions for future theoretical research on financing and performance of entrepreneurial firms.

It is important to explicitly address the issue of potential selection on unobservables (selection effect) to obtain unbiased estimates of the effect of capital structure on performance (incentive effect).⁵ We deal with the issue of potential endogeneity of capital structure by using a type of the endogenous treatment model. To approximate initial capital structure we employ three broad measures of outside business financing: bank and government loans, funds supplied by individual investors, and any outside financing whether provided by banks or outside investors. To measure performance we use growth in number of employees and survival after three years.

Our results indicate that the effect of leverage on the performance of business start-ups is insignificant contrary to previous findings on well-established firms. In contrast, both selection and incentive effects are present in the case of outside equity suggesting that outside investors might be able to both overcome informational opaqueness of business start-ups and provide better incentives for performance. This might explain why many new firms first acquire outside equity and later shift into debt financing.

This paper is organized as follows. Section 3.2 provides an extensive literature review on capital structure and business performance theories. Section 3.3 describes our econometric model, identification, and the algorithm used for estimation. Sections 3.4 discusses the data.

⁴Survivorship bias results from the fact that we typically have a non-random sample of all business start-ups because we do not observe the failed ones.

⁵The bias results from the fact that a capital structure is correlated with the error term in the performance regression (both are affected by the private information).

Section 3.5 provides the results and Section 3.6 concludes.

3.2 Literature Review

To evaluate the effect of capital structure on firm survival and growth while controlling for selection, we concentrate on two large theoretical areas – optimal capital structure theories and theories of incentives implied by a given capital structure. Although there are several theories that incorporate both selection and incentive effects, we are not aware of a single unifying theory that explains the problem at hand. While there are many competing finance theories of optimal capital structure, most were developed for large companies with dispersed ownership where the influence of a single manager or owner is likely to be insignificant. In case of start-up firms, an entrepreneur herself and her incentives are major forces behind the choices and performance of her firm. Furthermore, contract theory suggests several possible effects of capital structure on the entrepreneur's effort and/or her choice of risk-return characteristics of a given project. This section reviews the most significant implications of the relevant theoretical literature.

In our empirical analysis, we are primarily interested in whether selection and moral hazard are present. Given the various predictions provided by the literature on capital structure and performance, we do not have prior expectations about the way unobservable characteristics and capital structure affect performance. Instead, we summarize the literature into three main hypotheses and later discuss the consistency of our results with these hypotheses.

3.2.1 Capital Structure Theories

There are four leading theories of optimal capital structure developed in the finance literature (see Myers (2003) for a comprehensive summary). Here we briefly describe them and concentrate on the ones that seem most relevant for a business start-up financing:

1. *The Modigliani-Miller value-irrelevance theory* states that sources of financing do not matter for the value of the firm as long as the capital markets are “perfect”. Perfect markets imply that markets for capital are not only competitive and frictionless, they

are also complete and it is possible to insure against any possible contingency that may arise. This theory is mainly used as a starting point to identify the situations when markets are incomplete and capital structure may matter.

2. *The static trade-off theory* implies that firms choose optimal debt-to-equity ratios such that the tax benefits associated with debt are equal to the distress costs associated with extra debt at the margin. Interest tax benefits of debt increase the value of the firm, while too much debt increases risk of bankruptcy or exacerbates agency costs like conflicts between creditors and shareholders. However, there appears to be no definitive research that shows that tax incentives play any significant role in debt policy decisions. Furthermore, empirical research shows that companies do not try to achieve their optimal debt ratios.
3. *The pecking-order theory* incorporates asymmetric information about the firm's assets and growth opportunities. As a result, the value of shares may not reflect the true value of the firm and the issuance of new shares may signal that either those opportunities are good or managers are trying to sell overvalued shares. In equilibrium, this leads to the following capital structure: firms prefer internal to external finance; if external financing is needed, firms first issue less risky debt and only then outside equity. In other words, if the manager of a firm has favorable information about assets and growth opportunities, he will avoid external equity financing (Myers and Majluf, (1984)).
4. *Agency and asymmetric information theories of capital structure* pioneered by Jensen and Meckling (1976) state that some unobservable characteristics or actions influence the choice of capital structure. Agency theory recognizes that interests of managers and owners in the firm are not aligned. This implies a similar capital structure to pecking-order theory although the underlining principles are different: pecking-order theory assumes that interests of managers and outside owners coincide. If the firm starts as being fully owned by the manager, the incentives are aligned properly. If external financing is required then the firm should turn to debt to maintain proper incentives.

Once debt becomes too risky and there is a chance of default, then the firm turns to the last source of additional capital, outside equity.

It is worth emphasizing that these theories were developed to explain financing decisions of big corporations. Among them, the last two theories are most suitable for small and start-up firms because these firms usually are most opaque and do not have an established record. This may explain why the majority of start-up firms do not have any outside financing at all and all the initial investment is financed by the entrepreneur's savings.⁶

We concentrate on several relevant theories that describe the relationship between unobservable characteristics of the entrepreneur (or a firm) and capital structure. Leland and Pyle (1977) and Ross (1977) develop models in which "good" entrepreneurs signal their quality. The paper of Leland and Pyle (1977) assumes that quality of the project is known only to the risk-averse entrepreneur who wants to diversify his risks in the project. They show that if the level of self-financing is observable, then good-quality entrepreneurs would signal their quality by partially financing their projects at the expense of diversification. In Ross (1977) "good" entrepreneurs choose debt because the risk of bankruptcy is lower for them than for "bad" entrepreneurs. Myers and Majluf (1984) argue that if the entrepreneur tries to raise outside equity to finance the project then the value of the project may not be perceived as very high by outsiders since the entrepreneur wants to share its proceeds. Therefore, financial investors would demand a high price for external finance. In that case, debt is preferable to outside equity even though bankruptcy becomes an issue. This leads to the pecking order theory with internal financing being the cheapest way to finance the business.

All these theories predict that "better" firms would choose debt to signal their quality and that they would prefer debt to outside equity. Therefore, if there were no incentive effects of capital structure then we would expect to see that debt is positively related to performance of start-ups because intrinsically good firms choose debt and perform well.⁷ We summarize the

⁶Berger and Udell (2003) show that among small firms sampled in the 1993 National Survey of Small Business Finance personal and close relatives funds contribute around 45% of all capital employed. In our data set this number is close to 64% (See Table 1 below).

⁷However, Berger and Udell (2002) show that if debt is too large then there would be an adverse effect of debt on performance as the risk of default increases, implying a negative relationship between performance and debt.

existing capital structure theories in the following hypothesis:

Hypothesis 1: Firms that expect to be more successful would prefer debt to outside equity and we would expect a positive relationship (measured by correlation) between debt and performance measures; and negative relationship between outside equity and performance measures.

3.2.2 Capital Structure and Performance

In this subsection we describe main theoretical and empirical research on the effect of capital structure on performance. In general, these theories are referred to as agency theories, and they assume that there are either effort incentives or risk incentives of capital structure. For example, outside equity – which usually results in lower ownership share of the entrepreneur-manager – may induce the entrepreneur to exert less effort. Moreover, the entrepreneur may choose either riskier projects, or perquisites, or simply withdraw assets from the firm. This problem of providing the correct incentives to insiders has received a lot of consideration since it was first raised by Jensen and Meckling (1976). Harris and Raviv (1991) and Myers (2001) provide surveys of this literature. The testable implication of this literature is that debt could be used to alleviate such problems and we should observe a positive relationship between debt and performance. Moreover, the risk of default may discipline entrepreneurs because in case of bankruptcy they risk losing their firms (see, for example, Grossman and Hart (1982)). The following hypothesis summarizes the agency theories of effort incentives of capital structure:

Hypothesis 2: Based on agency theories about effort incentives created by debt, we expect to see a positive relationship between debt and performance measures, and negative relationship between outside equity and performance.

While negative effects of outside equity on performance are well understood, there may also be negative incentive effects of debt. For example, an entrepreneur who acquired debt financing may shift into riskier projects and increase riskiness of her firm. Therefore, for risky firms outside equity may be a good monitoring device and result in a positive relationship between outside equity and performance (Berger and Udell (2003)).⁸ In addition, Dybvig and

⁸This is one of the rationales for venture capital financing of fast-growing high-tech companies.

Wang (2002) argue that debt financing may induce an entrepreneur to keep the revenues and default on her debt. Such problems would result in a negative relationship between debt and performance. The following hypothesis summarizes the agency theories of risk incentives of capital structure:

Hypothesis 3: Based on agency theories about risk incentives created by debt, we expect to see a negative relationship between debt and performance measures, and positive relationship between outside equity and performance.

3.2.3 Models with Both Types of Informational Frictions

There is also a series of papers that incorporate both adverse selection and moral hazard into financing decision. Darrough and Stoughton (1986) develop a model of optimal capital structure under both adverse selection and moral hazard which shows that inside equity decreases with volatility of returns and the entrepreneur's effort decreases with higher expected marginal productivity when marginal productivity and riskiness of the project are unobservable characteristics of the entrepreneur. This model predicts that more efficient entrepreneurs would choose more debt while entrepreneurs with riskier projects would prefer less debt and more outside equity, other things equal. However, due to assumptions of their model it is not possible to determine the effect of effort on performance. Wahrenburg (1996) constructs a principal-agent model of project investment with both adverse selection and so called "false" moral hazard. He shows that the optimal contract can be implemented in terms of debt and equity payoffs with higher ability agents keeping 100% stake in the project and repaying debt to the principal. Bajaj, Chan and Dasgupta (1998) modify Leland and Pyle's (1977) model to allow for both adverse selection and moral hazard in business financing with the degree of moral hazard measured as control rights of the entrepreneur. This model explains capital structure and firm performance in terms of exogenous ownership structure. They show that debt increases with ownership because ownership works as a signal of quality with higher ownership meaning better quality, and that the performance also increases with ownership. Jullien, Salaniè and Salaniè (2005) construct a model of moral hazard with agents differing in

unobservable risk-aversion. They find that more risk-averse entrepreneurs would prefer debt to equity. Moreover, the probability of success of the project increases with inside equity because stimulates the entrepreneur's effort. The testable prediction of these models is that debt is positively related to performance and better firms choose debt over equity. Furthermore, the success of the project is also positively related to debt.

There are two other papers that study capital structure choice and firm performance simultaneously as we do here. The most closely related to our work is a study by Dessi and Robertson (2003). They examine the effect of leverage on performance of established UK firms while explicitly accounting for endogeneity of capital structure choices. They find that unobserved firm characteristics are important determinants of capital structure and performance. Furthermore, the leverage stops being significant after controlling for capital structure endogeneity suggesting that debt is chosen optimally according with the static trade-off theory. However, the main purpose of our study is to examine entrepreneurial firms at the time of their creation, and to do so we employ a different methodology and different measures of performance. Berger and di Patti (2006) consider the impact of capital structure on firm performance (measured as profit efficiency) and argue that firm performance influences capital structure because either more efficient firms have lower expected costs of bankruptcy (therefore, they acquire more debt) or more efficient firms want to protect the rents that come from higher profit efficiency and shareholders prefer to hold more equity to avoid liquidation leading to less debt. Our paper differs from Berger and di Patti (2002) in that we use survival and growth as the performance measures, and they are measured after the choice of initial capital structure is made, and therefore they cannot influence the choice of initial capital structure. The simultaneity in our paper comes from the idea that some unobservable variables influence both capital structure and performance. As a result, capital structure endogeneity has to be taken into account when estimating the effect of capital structure on performance. Furthermore, our sample is not limited to one particular industry unlike Berger and di Patti (2002) paper that concentrates on banking industry.

3.2.4 Selection and Moral Hazard in Lending and Insurance Markets

The issue of selection and incentive effects is explored in other markets as well. Edelberg (2004) proposes and tests a model of consumer lending that incorporates both adverse selection and moral hazard. Using the data on mortgages and automobile loans from the Survey of Consumer Finances, she shows that consumers self select into contracts that differ in terms of interest rates and levels of collateral. Moreover, higher levels of collateral induce consumers to exert higher effort to ensure the loans are repayed. Using a data set that describes the Kansas voluntary deposit insurance system for banks during 1910-1920, Wheelock and Kumbhakar (1995) show that riskier banks selected to participate in the insurance system, and banks in the system chose to hold less reserves and became more prone to risk.

There is a stream of papers that separate moral hazard and adverse selection in automobile insurance markets. Under adverse selection, higher risk agents are more likely to self-select into contracts with more coverage. They are also more likely to have an accident. At the same time, better coverage has the incentive effect – it induces riskier behavior. In both cases, better coverage is positively correlated with probability of an accident. For example, Dionne et al. (2004) find that low risk individuals self select into contracts that provide less coverage over time, and this induces them to change their unobservable efforts to reduce claims. Chiappori and Salaniè (2000), however, find no evidence of adverse selection or moral hazard in the French market for automobile insurance. But Abbring et al. (2003) argue that dynamic analysis would be more relevant because previous experience could partially reveal hidden effort.

3.3 The Model

3.3.1 General Outline

Entrepreneurs are assumed to have private information about the expected future outcome of their start-up. The observed capital structure (e.g., the level of external debt) will be affected by this private information. For example, it is often hypothesized in the theoretical literature that businesses with better prospects will choose debt over external equity because

debt does not dilute ownership. Conditional on the chosen capital structure, an entrepreneur decides how much effort to supply, which in turn affects performance. For example, high levels of external equity (as opposed to external debt) are generally believed to reduce the effort because the entrepreneur will own just a part of the business. Based on these hypotheses, we expect to have a positive correlation between the level of debt of each individual start-up and its success.

We have two distinct processes that can generate this positive correlation. On the one hand, entrepreneurs who expect to do well may want to have as large an ownership share as possible and not dilute it with external equity by inviting outside investors. If an entrepreneur does not have high expectations for the future payoff (either because the project is expected to generate fairly low returns or the risk is very high), she may prefer to bring in external investors to share the risk. On the other hand, given the chosen capital structure, entrepreneurs with high levels of debt and low levels of external equity will face relatively high incentives to supply more effort than those entrepreneurs with high levels of external equity because their ownership is not diluted (standard moral hazard). Assuming that high levels of effort translate into better performance, we will once again observe positive correlation between debt and success. The direction of the causality is different. In one case, the expectation of success affects the choice of capital structure. In the other case, the capital structure affects incentives, and thus, probability of success.

Each start-up can face different debt (or external equity) contract offers that reflect its observable characteristics. In other words, observably more promising enterprises may be offered more attractive debt contracts thus making them more likely to increase their debt levels. However, the important feature of the decision making process is that firms may have the unobserved by the potential lender propensity to perform well, which would obviously influence both the capital structure decision and the resulting performance. This means that estimating the effect of capital structure on performance will generally produce biased estimates because the error term will correlate with observed debt levels.

These unobserved factors can be accounted for in the following econometric specification.

With the subscript i denoting an individual start-up, we denote by Y_i^* some continuous measure of its performance and by L_i^* some continuous measure of its capital structure:

$$Y_i^* = x_{1i}\beta_1 + \gamma L_i^* + \eta_{1i}, \quad (3.1)$$

$$L_i^* = x_{2i}\beta_2 + \eta_{2i}, \quad (3.2)$$

where x_{1i} and x_{2i} are $1 \times k_1$ and $1 \times k_2$ vectors of exogenous variables; β_1 and β_2 are $k_1 \times 1$ and $k_2 \times 1$ vectors of parameters, γ is a scalar parameter measuring the effect of the endogenous capital structure on performance; and $\eta_i = (\eta_{1i}, \eta_{2i})'$ is the vector of error terms assumed to be normally distributed with zero mean $\mu = (0, 0)'$ and covariance Σ :

$$\Sigma = \begin{pmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{12} & \sigma_{22} \end{pmatrix}. \quad (3.3)$$

Importantly to our analysis, the error terms of the two equations are allowed to be correlated with each other. This correlation term is designed to capture unobserved private information about the quality of the start-up, which directly influences the outcome Y_i^* and also affects the capital structure (borrowing) decision L_i^* of each firm.

We use two distinct measures of performance – survival and a measure of growth. In our data set, which we describe in more detail in the next section, we have data on the survival of business start-ups within three years of the initial interview. We use the change in the number of employees as a measure of growth. Therefore, we have one censored continuous (growth) and one dichotomous (survival) measure of performance of business start-ups. We use the percent of the total investment into business coming from loans as our measure of debt and the percentage of the total investment coming from external individual investors and investment companies as our measure of external equity. Both are continuous censored random variables. When using the discrete measure of performance, we have to make certain adjustments to our error covariance matrix, which will have implications for the estimation procedure. In particular, the variance of the error term of the first equation is normalized to

1. We refer to this specification with the normalized variance of the error in the first equation (dichotomous performance measure) as *Specification 1* and to the model with the unrestricted variance (censored continuous performance measure) as *Specification 2*.

Denoting by Y_i and L_i the observed measures of performance and capital structure, respectively, we have the following system of equations:

$$\begin{aligned} Y_i^* &= x_{1i}\beta_1 + \gamma L_i + \eta_{1i}, \\ L_i^* &= x_{2i}\beta_2 + \eta_{2i}. \end{aligned} \tag{3.4}$$

The latent data Y_i^* and L_i^* are transformed into observed data Y_i and L_i in the following manner for *Specification 1*:

$$\begin{aligned} Y_i &= I[Y_i^* \geq 0], \\ L_i &= I[L_i^* \geq 0] \times L_i^*, \end{aligned} \tag{3.5}$$

and for *Specification 2*:

$$\begin{aligned} Y_i &= I[Y_i^* \geq 0] \times Y_i^*, \\ L_i &= I[L_i^* \geq 0] \times L_i^*. \end{aligned} \tag{3.6}$$

where $I[\cdot]$ is an indicator function taking value 1 if the expression in brackets is true and value 0 otherwise.

The hypotheses formulated in the previous section can be expressed in terms of the parameters of our model as follows:

Hypothesis 1: We expect $\sigma_{12} > 0$ for outside debt and $\sigma_{12} < 0$ for outside equity.

Hypothesis 2: Based on agency theories about effort incentives created by debt, we expect $\gamma > 0$ for outside debt and $\gamma < 0$ for outside equity.

Hypothesis 3: Based on agency theories about risk incentives created by debt, we expect $\gamma < 0$ for outside debt and $\gamma > 0$ for outside equity.

3.3.2 Estimation Details

We use Bayesian estimation procedures to estimate the system of equations (3.4) jointly. The choice of methodology is motivated by the superiority in performance of Bayesian methods over Maximum Simulated Likelihood (MSL) in this type of endogenous treatment models.⁹ The basic idea is to obtain the posterior distribution of the parameters of the model. The posterior distribution is proportional to the product of the likelihood of the observed data and prior distribution of the parameters:

$$p(\text{parameters} | \text{data}) \propto p(\text{data} | \text{parameters}) p(\text{parameters}).$$

We use simulation to obtain samples from the posterior distribution of the parameters because it does not have a form of any recognizable distribution. We develop a straightforward Gibbs sampler with data augmentation to simulate draws from the posterior distribution. The data augmentation step draws the values of latent variables Y_i^* and L_i^* conditional on the observed data and the parameters of the model (see Albert and Chib (1993)). The details of the algorithms are provided below. We derive two separate posterior simulators for each of the two specifications because the sets of the parameters to be estimated are different.

We start by deriving the (augmented) likelihood of the latent variables Y_i^* and L_i^* . We stack the two equations in the following manner:

$$\beta = \begin{pmatrix} \beta_1 \\ \gamma \\ \beta_2 \end{pmatrix}_{(k_1+k_2+1) \times 1} ; \quad X_i = \begin{pmatrix} x_{1i} & L_i & 0_{1 \times k_2} \\ 0_{1 \times k_1} & 0 & x_{2i} \end{pmatrix}_{2 \times (k_1+k_2+1)} ;$$

$$y_i^* = \begin{pmatrix} Y_i^* \\ L_i^* \end{pmatrix}_{2 \times 1} ; \quad \eta_i = \begin{pmatrix} \eta_{1i} \\ \eta_{2i} \end{pmatrix}_{2 \times 1} .$$

⁹See, for example, Munkin and Trivedi (2003).

The system can then be expressed as

$$y_i^* = X_i\beta + \eta_i.$$

We then stack all the observations together as

$$y^* = \begin{pmatrix} y_1^* \\ y_2^* \\ \vdots \\ y_n^* \end{pmatrix}_{2n \times 1}; X = \begin{pmatrix} X_1 \\ X_2 \\ \vdots \\ X_n \end{pmatrix}_{2n \times (k_1 + k_2 + 1)}; \eta = \begin{pmatrix} \eta_1 \\ \eta_2 \\ \vdots \\ \eta_n \end{pmatrix}_{2n \times 1},$$

to produce

$$y^* = X\beta + \eta.$$

We can express the covariance matrix for y^* as

$$\Omega = \begin{pmatrix} H^{-1} & 0 & \dots & 0 \\ 0 & H^{-1} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & H^{-1} \end{pmatrix}_{2n \times 2n} = I_n \otimes H^{-1}, \quad \text{where } H = \Sigma_{2 \times 2}^{-1}.$$

Conditional on the parameters of the model, the augmented likelihood can be expressed as:¹⁰

$$\begin{aligned} p(y^* | \beta, \Sigma) &= (2\pi)^{-\frac{2n}{2}} |I_n \otimes H^{-1}|^{-\frac{1}{2}} \exp\left(-\frac{1}{2}(y^* - X\beta)'(I_n \otimes H^{-1})^{-1}(y^* - X\beta)\right) \\ &\propto |H|^{\frac{n}{2}} \exp\left(-\frac{1}{2} \text{tr}\left[\sum_{i=1}^n (y_i^* - X_i\beta)' H (y_i^* - X_i\beta)\right]\right). \end{aligned} \quad (3.7)$$

For computational simplicity, the latent variables Y_i^* and L_i^* are treated as additional parameters of the model. The appropriate steps are added to our Gibbs sampling algorithm to draw these latent variables conditional on the realized values of the main parameters of the

¹⁰See appendix for details.

model. The latent data are then integrated out to obtain the posterior distribution of the main parameters. The augmented posterior $p(y^*, \beta, \Sigma | data)$, which also contains the latent data, is proportional to

$$p(y^*, \beta, \Sigma | data) \propto p(data|y^*) p(y^*|\beta, \Sigma) p(\beta, \Sigma),$$

where $p(data|y^*)$ is the distribution of the observed data conditional on the latent data (this distribution is going to be different for the two specifications we use), $p(y^*|\beta, \Sigma)$ is the augmented likelihood, and $p(\beta, \Sigma)$ is the prior distribution of the main parameters.

We specify independent priors for β and H^{-1} . The prior for β is normal and given by:

$$\beta \sim N(\beta_0, V_0). \quad (3.8)$$

The conditional posterior of β can be shown to be also normal:¹¹

$$\begin{aligned} \beta | data, H, y_i^* &\sim N(\beta_1, V_1) \\ V_1 &= \left(\sum_{i=1}^n X_i' H X_i + V_0^{-1} \right)^{-1} \\ \beta_1 &= V_1 \left(\sum_{i=1}^n X_i' H y_i^* + V_0^{-1} \beta_0 \right). \end{aligned} \quad (3.9)$$

The prior and the posterior distributions of the precision matrix H depends on what type of dependent variable Y_i we use. We consider two cases.

The first case is when Y_i is a censored continuous random variable. We assume a Wishart prior for the precision matrix H :

$$H \sim W(\nu_0, H_0). \quad (3.10)$$

¹¹See appendix for details.

It can be shown that the conditional posterior distribution of H is also Wishart

$$H|data, \beta, y_i^* \sim W(\nu_1, H_1) \quad (3.11)$$

$$\begin{aligned} \nu_1 &= \nu_0 + n \\ H_1 &= \left(\sum_{i=1}^n (y_i^* - X_i\beta)(y_i^* - X_i\beta)' + H_0^{-1} \right)^{-1}. \end{aligned}$$

The other case that we consider is of the binary survival dependent variable Y_i . The approach developed for the continuous dependent variable is not directly applicable here because the variance parameter in binary discrete choice models is not identified. Only the ratio $\frac{\beta_1}{\sigma_{11}^{1/2}}$ is identified. This follows from the fact that we do not observe the underlying latent propensity Y_i^* . Thus, in the case of dichotomous dependent variable we have to work with the following identified covariance matrix:

$$\begin{aligned} \Sigma = H^{-1} &= \begin{pmatrix} 1 & \rho \\ \rho & \sigma_{22} \end{pmatrix} \\ &= \begin{pmatrix} 1 & \rho \\ \rho & h^{-1} + \rho^2 \end{pmatrix}, \end{aligned} \quad (3.12)$$

The reason to use the parametrization given above is that it allows us to use the properties of the multivariate normal distribution. In particular, $p(\eta_i) = p(\eta_{1i})p(\eta_{2i}|\eta_{1i})$. We restrict $\eta_{1i} \sim N(0, 1)$, while $\eta_{2i}|\eta_{1i} \sim N(\rho\eta_{1i}, \sigma_{22} - \rho^2)$. Following McCulloch et al. (2000), we specify the following prior

$$\begin{aligned} p(\rho, h) &= p(\rho)p(h) \quad (3.13) \\ \rho &\sim N(\rho_0, V_{\rho 0}) \\ h &\sim G(\nu_0, h_0), \end{aligned}$$

where $G(\nu_0, h_0)$ is the gamma distribution with mean h_0 and degrees of freedom ν_0 . The

likelihood with this new parametrization can be rewritten as:

$$\begin{aligned} p(\mathbf{y}^*|\beta, \rho, h) &\propto |H|^{\frac{n}{2}} \exp\left(-\frac{1}{2} \text{tr}\left[\sum_{i=1}^n \eta_i' H \eta_i\right]\right) \\ &\propto h^{\frac{n}{2}} \exp\left(-\frac{1}{2} \sum_{i=1}^n (\eta_{1i}^2 (1 + h\rho^2) - 2\eta_{1i}\eta_{2i}h\rho + \eta_{2i}^2 h)\right). \end{aligned} \quad (3.14)$$

Observe that given ρ, h , the distribution of H is degenerate.¹² Therefore, the conditional posterior for β is exactly the same. The conditional posteriors of (ρ, h) are given by:

$$\begin{aligned} \rho|data, \beta, h, \mathbf{y}_i^* &\sim N(\rho_1, V_{\rho 1}) \\ V_{\rho 1} &= (V_{\rho 0}^{-1} + h \sum_{i=1}^n \eta_{1i}^2)^{-1} \\ \rho_1 &= V_{\rho 1}(\rho_0 V_{\rho 0}^{-1} + h \sum_{i=1}^n \eta_{1i}\eta_{2i}), \end{aligned} \quad (3.15)$$

and

$$\begin{aligned} h|data, \beta, \rho &\sim G(\nu_1, h_1) \\ \nu_1 &= \nu_0 + n \\ h_1 &= \left[\frac{\nu_0}{h_0(\nu_0 + n)} + \frac{\sum_{i=1}^n (\eta_{1i}\rho - \eta_{2i})^2}{(\nu_0 + n)} \right]^{-1}. \end{aligned} \quad (3.16)$$

We use Gibbs sampling algorithms to successively draw from conditional distributions for parameters of the model β , H and latent indices Y_i^* and L_i^* . We denote the s^{th} realization of variable a by a^s . The total number of draws $S = S_0 + S_1$ will be made with the first S_0 discarded as the burn-in. We use the following algorithm for *Specification 1*:

step 0: Set $(Y_i^*)^0 = Y_i$, $(L_i^*)^0 = L_i$, $\rho^0 = 0$, and $h^0 = 1$ (which corresponds to Σ equal to identity matrix);

step 1: draw β^1 from the distribution given in (3.9) conditional on $(Y_i^*)^0$, $(L_i^*)^0$, ρ^0 and h^0 ;

step 2: draw the elements of covariance matrix Σ as a block:

¹²It can be shown that the transformation from (ρ, h) to H is one-to-one.

draw ρ^1 from distribution given in (3.15) conditional on $(Y_i^*)^0$, $(L_i^*)^0$, β^1 and h^0 ;

draw h^1 from distribution given in (3.16) conditional on $(Y_i^*)^0$, $(L_i^*)^0$, β^1 and ρ^0 ;

step 3: draw $(L_i^*)^1$ conditional on $(Y_i^*)^0$, ρ^1 , h^1 and β^1 as:

$$(L_i^*)^1 = \begin{cases} L_i & \text{if } L_i \geq 0 \\ \text{draw from normal distribution} \\ \text{truncated above at 0 with mean equal to} \\ x_{2i}(\beta_2)^1 + \rho^1[(Y_i^*)^0 - x_{1i}(\beta_1)^1 + \gamma^1 L_i] \\ \text{and variance equal to } (h^1)^{-1} & \text{if } L_i < 0; \end{cases}$$

step 4: draw $(Y_i^*)^1$ conditional on $(L_i^*)^1$, ρ^1 , h^1 and β^1 as:

$$(Y_i^*)^1 = \begin{cases} \text{draw from normal distribution} \\ \text{truncated below at 0 with mean equal to} \\ x_{1i}(\beta_1)^1 + \frac{\rho^1}{(\rho^1)^2 + (h^1)^{-1}} ((L_i^*)^1 - x_{2i}(\beta_2)^1) \\ \text{and variance equal to} \\ 1 - \frac{(\rho^1)^2}{(\rho^1)^2 + (h^1)^{-1}} & \text{if } Y_i \geq 0 \\ \text{draw from normal distribution} \\ \text{truncated above at 0 with mean equal to} \\ x_{1i}(\beta_1)^1 + \frac{\rho^1}{(\rho^1)^2 + (h^1)^{-1}} ((L_i^*)^1 - x_{2i}(\beta_2)^1) \\ \text{and variance equal to} \\ 1 - \frac{(\rho^1)^2}{(\rho^1)^2 + (h^1)^{-1}} & \text{if } Y_i < 0 \end{cases}$$

step 5: repeat steps 1-4 S times;

We use the following algorithm for *Specification 2*:

step 0: Set $(Y_i^*)^0 = Y_i$ and $(L_i^*)^0 = L_i$, $H^0 = I$ (where I is an identity matrix);

step 1: draw β^1 from distribution given in (3.9) conditional on $(Y_i^*)^0$, $(L_i^*)^0$, H^0 ;

step 2: draw the precision matrix H^1 from the distribution given in (3.11) conditional on $(Y_i^*)^0$, $(L_i^*)^0$, β^1 , then invert it to obtain $(\Sigma)^1$;

step 3: draw $(L_i^*)^1$ conditional on $(Y_i^*)^0$, H^1 and β^1 as:

$$(L_i^*)^1 = \begin{cases} L_i & \text{if } L_i \geq 0 \\ \text{draw from normal distribution} \\ \text{truncated above at 0 with mean equal to} \\ x_{2i}(\beta_2)^1 + \frac{(\sigma_{12})^1}{(\sigma_{11})^1} [(Y_i^*)^0 - x_{1i}(\beta_1)^1 + \gamma^1 L_i] \\ \text{and variance equal to } (\sigma_{22})^1 - \frac{(\sigma_{12})^1}{(\sigma_{11})^1} & \text{if } L_i < 0; \end{cases}$$

step 4: draw $(Y_i^*)^1$ conditional on $(L_i^*)^1$, H^1 and β^1 as:

$$(Y_i^*)^1 = \begin{cases} Y_i & \text{if } Y_i \geq 0 \\ \text{draw from normal distribution} \\ \text{truncated above at 0 with mean equal to} \\ x_{1i}(\beta_1)^1 + \frac{\rho^1}{(\rho^1)^2 + (h^1)^{-1}} ((L_i^*)^1 - x_{2i}(\beta_2)^1) \\ \text{and variance equal to} \\ 1 - \frac{(\rho^1)^2}{(\rho^1)^2 + (h^1)^{-1}} & \text{if } Y_i < 0 \end{cases}$$

step 4: repeat steps 1-3 S times;

For each run of *Specification 1* we set $S_0 = 50,000$ and $S_1 = 50,000$. For each run of *Specification 2* we set $S_0 = 10,000$ and $S_1 = 50,000$. The reason for this difference is that our preliminary Monte Carlo tests showed that it takes longer for parameters to converge to their true distribution if the dependent variable in equation 1 is dichotomous (implying the necessity to simulate the underlying latent variable). We specify proper but sufficiently diffuse priors in all cases. For both specifications, the prior distribution of β is given by:

$$\beta \sim N(0_{k_1+k_2+1}, 100 * I_{k_1+k_2+1}),$$

where $I_{k_1+k_2+1}$ is the $(k_1 + k_2 + 1) \times (k_1 + k_2 + 1)$ identity matrix. We assume the following

Wishart prior for the precision matrix H in the *Specification 2*:

$$H \sim W\left(\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, 3\right).$$

We choose the following priors for ρ and h in the *Specification 1*:

$$\rho \sim N(0, 100)$$

$$h \sim G(2, 5).$$

We performed some diagnostics of the Markov Chains used in this paper. The post-burn-in first-order autocorrelation coefficient for the covariance parameter was about 0.91 and the autocorrelation for γ was around 0.87. Although both of them are fairly high, the large number of post-burn-in replications (50,000) insures that the sampler covers the posterior distribution sufficiently well (it takes less than 50 iterations for the effects of previous shocks to disappear). Gelman-Rubin statistics were 1.021 for γ and 1.011 for the covariance parameter in the *Specification 1*. The values of Gelman-Rubin statistics in the *Specification 2* were 1.005 for γ and 1.01 for the covariance parameter. These were obtained by running the respective Gibbs sampler 50 times with overdispersed starting values of the parameters of the model (10 values were chosen manually to insure that extreme values were covered, the remaining 40 runs were started at values drawn from a normal distribution centered at zero with a very large variance). These values are within the usual 1.2 cut-off value suggesting that the samplers converged well.

3.3.3 Identification

The system of equations (3.1), (3.2) is nonparametrically unidentified. Technically, the identification can be achieved by using a non-linear transformation from L_i^* to L_i , which we do by employing a censored version of the underlying latent index. Nevertheless, it has long been noted in the literature that it is preferable to achieve identification through exclusion

restrictions. Finding appropriate instruments for the effect of capital structure on performance, however, has proved to be quite difficult. As we already mentioned, there are only a few papers recognizing potential endogeneity of capital structure. Bitler et al. (2005) use the following to instrument for the effect of capital structure on effort (which is assumed to be related to performance): (1) a dummy for whether the business was inherited or given to the present owner, (2) a dummy for whether the business was started-up by the present owner. The omitted category in their analysis consists of those entrepreneurs who bought their current business. In addition, Bitler et al. (2005) use the initial investment by entrepreneur as an instrument. Since we do not model effort in our specification, we have to justify the use of similar instruments in our model. It is arguable that all three are fairly weak instruments. The way in which the business was acquired may tell a lot about the (latent) type of entrepreneur, which in turn may play an important role in determining whether a start-up will be successful. The size of initial investment can be linked to the perception by the owner of the risk involved, which in turn may affect her effort and, thus, performance. We have data on whether the business start-up was inherited or started up by the entrepreneur herself. We, therefore, include a dummy for whether an entrepreneur started the business herself. The omitted category includes those who inherited their businesses, were brought in by other owners, promoted, or purchased it. We found that the initial capital investment is a very strong predictor of both performance and our measures of capital structure.

We propose several new instruments for the effect of capital structure on performance. These variables are related to the access to credit by small businesses. We argue that the differences in the ability of small businesses to obtain bank bank loans affect their capital structure decisions. At the same time, access to credit should not affect performance directly.

Our first instrument is the homestead exemption amount (the value of housing equity that a person can keep in case she declares bankruptcy) for the state where the business is located as an instrument in the capital structure equation. Berkovitz and White (2004) study the effect of state bankruptcy laws on small firms' access to credit. They find that in states with high homestead exemptions, small businesses (both non-corporate and incorporated) are more likely

to be credit rationed and face higher interest rates. Following Berkovitz and White (2004), for those states with unlimited homestead exemption we assign the highest observed homestead exemption level, and we include a dummy variable taking the value 1 for these states.¹³ We also include the Herfindahl index of the concentration of banking industry for the state of each firm's location.¹⁴ This variable measures the competitiveness of the banking industry in each state. In addition to these two supply-side instruments, we include the median debt-to-assets ratio of firms in each industry and state to account for dependence of small firms in each industry on bank loans.¹⁵ This demand-side variable is introduced to control for "standard" leverage levels in each industry/state. Cetorelli and Strahan (2004) use this variable to study the effect of competition in the banking sector on market structure of other sectors. It is, however, possible to argue that different industries in different states may also differ in their inherent ability to survive and succeed. We address this problem by including the change in number of firms in each industry/state from 1985 (year of the first interview) to 1987 (year of the last interview) in the performance equation. This variable should control for the average ability to survive in each industry/state over the sample period.

3.4 Data

The primary data source used in this paper is the national survey of business start-up firms called *New Business in America: The Firms and Their Owners*. This survey was first fielded in 1985 by the National Federation of Independent Business (NFIB) Foundation. There were two more waves fielded in 1986 and 1987, respectively. The sample is drawn from a national sample of new business owners that were members of the NFIB. Given that there does not exist

¹³In 1985, year of the first interview, eight states had unlimited homestead exemption: Arkansas, Florida, Iowa, Kansas, Minnesota, Oklahoma, South Dakota, Texas. The highest homestead exemption was \$90,000 in Nevada.

¹⁴Herfindahl index is calculated as a sum of squared market shares of each institution in the state. Here we have to note that it is possible to define institution as a any type of establishment – branch or main office. Alternatively, we can count each bank holding corporation as institution. The Federal Deposit Insurance Corporation (FDIC) routinely collects data on all lending institutions (branches and main offices). However, the data on these is only available to us starting from 1992, which is too long of a period after the first interview year, 1985. We do have data on the assets of all banks in 1985 grouped by bank. Therefore, our Herfindahl index is based on assets of each bank and not counting each branch as a separate institution.

¹⁵We use the first two digits of the Standard Industrial Classification (SIC) code to define industry.

a benchmark national sample of new businesses, this survey arguably provides close coverage to a nationally representative population of new businesses. Moreover, the characteristics of small businesses based on the NFIB survey closely match other business surveys (that include new and old firms).¹⁶ The survey has some very attractive features, which make it particularly suitable for our research. First, it targets new businesses – the majority of the firms in the sample are younger than 2 years. We restrict our sample to firms that were created since January 1983. The average firm age is just above 14 months. Other major data sets involving small businesses have considerably “older” samples.¹⁷ We hypothesize that in small new businesses the informational problems we study are especially pronounced. In addition, survivorship bias becomes an important consideration in any survey of mature businesses (which survived until the day of the survey). Second, the survey is longitudinal. Although the response rate is greatly reduced by the third (and final) survey year, it still allows us to study the performance of new start-ups within the next three years after the first wave was fielded. Third, the survey collects an array of important information related to the personal characteristics of entrepreneurs themselves. It has been noted in the literature that these characteristics are important in determining both financial structure and performance of business start-ups (see, for example, Cassar (2004)). Last, the survey contains state and industry identifiers, which allows easy matching to other ancillary data sets. The survey, however, has several drawbacks. Probably the biggest one has to do with the fact that the individual item non-response rate is quite high, which makes it costly for us to include large numbers of controls in our equations. Moreover, it does not contain direct (from balance sheets) measures of leverage and outside financing. Instead, we have information on the percentage of total capital invested prior to the first sale coming from different sources, including bank loans and outside investors.

We now describe the ancillary data sets used in this paper. The information on the change of the number of institutions by state/industry over the period 1985-1987 was obtained from

¹⁶See, Cooper et al. (1990).

¹⁷The average age of a firm sampled by the National Survey of Small Business Finances (NSSBF) is equal to 13.4 years (Bitler et al. (2005)).

the US Census Bureau's County Business Patterns data set.¹⁸ We follow Cooper et al. (1990) to control for general economic activity in the states by including the change in unemployment rates by state over the period 1985-1987 as an additional control variable in the performance equation. The data were obtained from the Bureau of Labor Statistics.¹⁹ The median debt-to-assets ratio in each industry/state were computed using the Federal Reserve's National Survey of Small Business Finances (NSSBF).²⁰ This data set was collected in 1987. Ideally, we would want to have this information for the year 1985. However, the 1987 wave of this survey is the closest match available. Dependence on external financing is not likely to change very much over the two-year period. The Herfindahl index reflecting the banking concentration in each state was computed using the Report of Condition and Income data available at the Chicago Federal Reserve Bank's website.²¹ The data contain information on all banks regulated by the Federal Reserve System, FDIC, and the Comptroller of the Currency. The market share of each bank was calculated based on total assets of this bank divided by total assets of all banks in the state. Finally, information on the bankruptcy homestead exemptions that were in effect in 1985 were obtained from Kosel (1985).

It should be noted that despite a significant non-response rate by the third wave of the survey, we still have information on which firms survived and which did not survive. Thus, we have two sample sizes – one for the study of survival (*Specification 1*), which is larger than the sample size used to explain growth (*Specification 2*). In the first case, we just need the information on survival, which is generally available whether or not respondent actually supplied information in third wave (1987). In the second case, we use the change in the total employment in each firm as a measure of growth. After deleting all observations with missing data we arrive at the sample size of 2820 businesses for *Specification 1* and 1522 businesses for *Specification 2* (the sample for *Specification 2* includes respondents in third interview and failed businesses (which lost all their employees)).²² Tables 3.1 and 3.2 list the variables (most

¹⁸The data set itself is also available in easy-to-download electronic form at the University of Virginia Library's webpage: <http://fisher.lib.virginia.edu/collections/stats/cbp/state.html>.

¹⁹Available on the web at: <http://www.bls.gov/lau/home.htm>.

²⁰The dataset is available on the web at: <http://www.federalreserve.gov/pubs/oss/oss3/nssbf87/nssbf87home.html>.

²¹http://www.chicagofed.org/economic_research_and_data/commercial_bank_data.cfm

²²As will be explained in the next section, we addressed the high non-response rate by simulating the em-

of variables are based on responses in the first interview) used in this study along with a brief description and summary statistics based on a larger sample.

3.5 Results

As shown in Table 3.3, only outside equity had a statistically significant effect on survival. One possible explanation of the estimated negative effect is that external equity induces entrepreneurs to exert lower effort, which in turn leads to lower survival chances (Hypothesis 2). Neither debt nor our measure of overall external finance had any statistically significant effect on survival. This is consistent with the findings of Dessi and Robertson (2003). They show that after accounting for capital structure endogeneity, debt becomes an insignificant determinant of performance. They suggest that this is consistent with trade-off theory of capital structure: if the level of debt is chosen optimally it should not influence performance.

Our results also suggest that private information about survival chances plays an important role in the outside equity decisions. Business start-ups with higher survival chances had higher levels of the outside equity financing relative to the overall amount of capital invested. This result is not consistent with Hypothesis 1, which predicts that firms with better prospects would prefer outside debt because it does not dilute ownership. This suggests that outside investors might be able to more successfully overcome informational problems associated with business start-ups. At the same time, it is likely that survival is a crude measure of performance, which does not properly capture the underlying propensity for a long-term success.

When estimating the effects of our measures of capital structure on the employment growth we have a smaller sample size of 1522 firms in the third year (surviving businesses that responded to the interview and failed businesses). We addressed this issue by simulating the employment growth values for non-responders by adding the appropriate step to the Gibbs sampler to draw positive values for non-responders.²³ We report the results based on both samples – the sample of observed firms and the full sample with simulated values for non-responders.

ployment growth values for non-responders.

²³This methodology implicitly assumes that the non-response is random.

Table 3.1 Variable definitions and descriptive statistics

variable	description	mean	st.dev.	min	max
outjob4	Devotes full-time to the business: 1-no outside employment, 0-any outside job	0.843	0.364	0	1
start1	Form of entry: 1- started it, 0-took over existing firm	0.659	0.474	0	1
age1	Age when became principal owner/manager	36.08	9.45	0	68
moved1	Moved residence to go into business: 1-moved, 0-didn't move	0.208	0.406	0	1
partners	Number of full-time business partners	0.418	0.796	0	8
managexp	Any supervisory/managerial experience: 1- yes, 0-no	0.782	0.413	0	1
firequit	Was fired/quit without specific plans: 1-yes, 0-no	0.170	0.376	0	1
diffprod	Product is very different from previous job: 1- yes, 0-no	0.376	0.485	0	1
parenown	Parents owned a business: 1- yes, 0-no	0.443	0.497	0	1
educlev	Highest level of education, scale: 1- less than high school, 9-advanced degree	4.34	1.72	1	9
totjobs	The total number of full-time jobs held prior to business formation	4.48	4.25	0	99
bs11	Percentage of business strategy assigned to "lower prices"	11.82	16.85	0	100
bs21	Percentage of business strategy assigned to "better service"	29.55	21.37	0	100
bs61	Percentage of business strategy assigned to "target missed/poorly served customers"	7.39	11.87	0	80
opercont	Strongly agree(1) to strongly disagree(5) with statement "business operating controls in writing"	2.73	1.12	1	5
sex	Owner's sex: 1-female, 0-male	0.196	0.397	0	1

Table 3.2 Variable definitions and descriptive statistics (cont'd)

variable	description	mean	st.dev.	min	max
race	Owner's race: 1-racial minority, 0-not a minority	0.058	0.233	0	1
capinv	Total capital invested prior to the first sale, categorical: 1- \leq \$5,000, 8- \geq \$500,000	3.38	1.70	1	8
invest	Percentage of capital coming from outside individual investors (not family or friends) and investment companies (/100)	4.70	17.74	0	100
loans	Percentage of capital coming from bank and government loans (/100)	31.78	38.44	0	100
outside	Percentage of capital coming from both loans and outside investors (/100)	36.48	39.46	0	100
oddsyr	The self-perceived chances of success, categorical: 0-no chance, 10-certain success	8.15	2.04	0	10
survive	Survived: 1-yes, 0-no	0.783	0.412	0	1
homestead	Homestead exemption in the state, in \$1,000	37.10	34.48	0	90
h_unlimited	1-Unlimited homestead exemption, 0-otherwise	0.268	0.443	0	1
unemp_change	Change (difference) in unemployment in state over period 1985 to 1987	-0.736	1.20	-2.6	1.7
HH	Herfindahl index of banking concentration in the state	725.6	697.9	65.8	3243.7
est_change	Change in the number of establishments by state/SIC code from 1985 to 1987	0.101	0.092	-0.388	0.671
debt_2_assets	Median debt-to-assets ratio for the SIC code	0.245	0.099	0.055	0.841
agef	Age of the firm, in months	14.4	6.8	1	29
changemp	Change in total employment in the firm from 1985 interview to 1987 interview	0.955	8.30	-1	175.7
lgl_form1.2	1-if partnership, 0-otherwise (proprietorship - omitted category)	0.113	0.316	0	1
lgl_form1.3	1-if corporation, 0-otherwise (proprietorship - omitted category)	0.320	0.467	0	1

Table 3.3 Results: Survival

		Survival	
		γ	Covariance
Capital Structure Measures	Loans	-0.054 (0.182)	0.003 (0.053)
	Outside Equity	-0.957*** (0.356)	0.295** (0.119)
	External Finance	-0.054 (0.272)	0.011 (0.070)

Notes: The sample size is 2,820. Standard errors are displayed in parentheses below coefficients; *** - significant at 1%, ** - significant at 5%.

Table 3.4 Results: Employment Growth

		Employment Growth			
		Respondents only		Full sample	
		γ	Covariance	γ	Covariance
Capital Structure Measures	Loans	-0.071 (1.287)	0.288 (0.353)	0.630 (1.063)	0.049 (0.291)
	Outside Equity	11.271*** (1.768)	-7.479*** (1.086)	13.306*** (1.069)	-6.795*** (0.595)
	External Finance	-0.694 (1.366)	0.315 (0.347)	0.389 (1.124)	0.030 (0.286)

Notes: The sample size is 1,522 for respondents and failed firms, and the full sample size is 2,820. Standard errors are displayed in parentheses below coefficients; *** - significant at 1%, ** - significant at 5%.

Table 3.4 shows that the results of the estimation of *Specification 2* are qualitatively the same for both samples. Only the effect of outside equity was estimated to have a statistically significant effect on the employment growth of those firms that responded to the third interview is positive. The positive relationship between the outside equity and the employment growth is consistent with Hypothesis 3, which predicts that outside equity is preferred to debt because it provides better risk incentives and monitoring. The selection effect of outside equity was also statistically significant. Firms with better growth expectations had significantly lower proportions of outside equity to initial capital investments. This result is consistent with Hypothesis 1. Successful entrepreneurs are less willing to share their returns with outside investors and prefer less outside equity in their businesses.

The consistent result emerging from tables 3.3 and 3.4 is that outside equity has both selection and incentive effects on performance of business start-ups. At the same time the results differ depending on which performance measure we use. It is possible to speculate that risk-sharing and profit-sharing concerns induce entrepreneurs to select different optimal capital structures. It may be that entrepreneurs care more about outside equity when it comes to survival as opposed to growth. This might explain why many new firms first acquire outside equity and later shift into debt financing. The estimated selection effects of debt or our measure of overall external finance are insignificant.

There is very little theoretical research dealing with incentive structures facing business start-ups as opposed to large publicly traded firms. This study suggests directions for future research needed to fully understand the differences among various measures of performance and how they are influenced by capital structure choices in entrepreneurial firms. In particular, it is necessary to explicitly differentiate between risk sharing and profit sharing incentives of entrepreneurs. Given the nature of small business start-ups, which rarely rely on any outside capital, more work is needed to differentiate between questions related to whether to use any outside capital versus questions related to what kind of outside capital to use.

3.6 Concluding Remarks

This paper is the first to examine the relationship between capital structure and performance of business start-ups in the presence of imperfect information. Our results suggest that debt does not have significant incentive or selection effects on performance of business start-ups. In contrast, both selection and incentive effects are present in the case of outside equity indicating that outside investors are able to both overcome informational asymmetries associated with business start-ups and provide better incentives for performance. Our results are qualitatively different depending on which measure of performance we use. In particular, we find that firms with higher survival chances are more likely to have external equity financing, which is inconsistent with the theoretical predictions.

CHAPTER 4. EMPLOYER-PROVIDED HEALTH INSURANCE: PREFERENCES AND CHOICE

Dzmitry Asinski

Abstract

The majority of Americans get health insurance through their employers. One of the potential concerns related to employer-provided health insurance is whether some categories of workers (typically with low educational attainment) actually have access to jobs with health insurance offers. This challenges an assumption imposed in some of the previous literature that workers have equal access to both jobs with and without health insurance offers and sort strategically according to their preferences. A worker can be uninsured either because of lack of choice or because of preferences for a job without a health insurance offer. To be able to separate the effect of lack of choice from the effect of preferences, I develop a structural job search model with two job characteristics - the offered wage and the availability of insurance. I simulate the model to demonstrate the interaction between preferences and availability of choice.

4.1 Introduction

The problem of the uninsured in United States has moved to the forefront of the public policy discussions in the recent years. It is especially important to focus on the employer-provided health insurance (EPHI) because the majority of Americans obtain health insurance from their employers.¹ One of the potential concerns related to the EPHI raised in the litera-

¹The percentage of insured Americans who get their insurance through employer is around 65% (Employee Benefit Research Institute).

ture is whether or not some categories of workers actually have access to jobs with insurance coverage. Some workers (typically workers with low educational attainment) may find it difficult to get a job with an insurance offer. The converse may be true for other types of workers (highly educated workers). They may find it difficult to get a job without insurance coverage. Blumberg and Cancian (2004) classify people into three groups: those who are unlikely to have a job with insurance coverage, those who are unlikely to have a job without insurance coverage, and those in-between the two groups. The classification is based on the distribution of predicted probabilities of having a job with an EPHI offer, which is turn was obtained via fitting a probit specification to the data from the Current Population Survey.² The main problem with this approach is that the job that an individual is observed to have is the result of the job search and not the first job offer that this individual had. The intensity of search for the job that fits individual preferences for health insurance coverage depends on the strength of the preferences for this particular job characteristic. For example, a person who doesn't really care about health insurance will likely accept the first job which comes across and satisfies all other requirements that this individual has. On the opposite end of the spectrum, a person who really values insurance may choose to search until she finds a job that offers it. Therefore, we may have two otherwise identical people with health insurance from their employers, one of whom got it quickly because her probability to find a job with it is very high while the other found the job after a long search.

It is, therefore, important to take into account individual preferences for health insurance. Two papers by Monheit and Vistnes (1999 and 2004) assess the role played by the preferences for health insurance in job search. Either a job with an employer-provided health insurance offer or a job without an EPHI offer has to be chosen. The authors developed a simple search model which yields the logit specification for the probability of having an EPHI offer. The wage differential as well as medical expenses and search costs differentials are approximated by a variety of exogenous variables (e.g., age, education). Both papers utilize the unique feature of the data sets used – they contain the ranked responses to the attitudinal questions on the

²The authors divided the population into three groups using the cutoff points of 0.5, 0.85 of the predicted probabilities.

valuation of the health insurance and on the general risk-taking preferences.³

The typical data that is available to an econometrician consists of accepted wages and indicators of whether chosen jobs have insurance offers. The observed insurance offer state may be the result of either strong “preference” or “choice” (availability of it and not the act of choosing) or both. In other words, some people may have choice and strong preferences and choose accordingly; some others may be stuck in the uninsured state because they do not have much choice even though they may have strong preference. Accordingly, the fact that a person has a job that offers health insurance may have different interpretation depending on preference for having insurance. The important conclusion is that studying either “preference” or “choice” separately will likely lead to biased conclusions as to their contribution to the observed outcome. Suppose, for example, that we intend to estimate the extent of the “choice” that an individual has by estimating a linear-in-parameters probit specification with the observed insurance state as a dependent variable. Two otherwise very similar uninsured persons with drastically different preferences for health insurance coverage will contribute to the probability estimation in the same way thus attributing the effect of preferences to other dependent variables.⁴ At the same time, using the self-reported data about preferences for health coverage as additional explanatory variables may not solve the problem completely. The underlying relationship between preferences and the resulting outcome may be non-linear. A person who strongly prefers to have a health insurance offer and has a relatively low cost to find such an offer (because she has enough “choice”) is likely to behave differently from the person who also has strong preference for health insurance but has relatively high cost associated with finding it (little “choice”). The latter is likely to have had many job offer draws before she found one with EPHI. In addition, simple linear-in-parameters probit is likely to be a very poor approximation to the underlying structural model because it effectively ig-

³First paper uses National Medical Expenditure Survey, which was fielded in 1987. The second paper uses the 2000 wave of Medical Expenditure Panel Survey, which was first fielded in 1996. Both surveys were administered by Agency for Healthcare Research and Quality (AHRQ).

⁴I have to note that it is unlikely that two people with different underlying degrees of “choice” will be observationally identical because this probability correlates with demographics, education and so on. However, the point is that one important dimension of personal characteristics – preference for health insurance coverage – is missing and it is likely to affect the outcome.

nores the possibility of accepting a job because of a very high wage irrespective of its other characteristics.

The purpose of this paper is to demonstrate the interaction of preferences and the availability of choice on the observed outcome (insurance status) in a structural framework. In order to do so, I develop and simulate a structural job search model. Estimation of structural search models has evolved significantly over the last decades.⁵ Although fairly straightforward, the estimation of the model under homogeneity assumption does not produce realistic results. I outline the specific problems that make estimation of the parameters of the structural model difficult to accomplish under the more realistic assumption of unobserved heterogeneity.

The paper proceeds as follows. Section 4.2 presents the formal structural search model. Section 4.3 provides simulations of the model. Section 4.4 gives some Monte-Carlo evidence of the estimation of the model under the homogeneity assumption. Section 4.5 outlines the problems with estimation of the more realistic versions of the model and concludes.

4.2 The Model

I denote *the probability of having an EPHI offer in any single job offer that a person receives* by p and *the probability to be observed with a job that has an EPHI offer* by P . I consider p to be the measure of “choice” that a person has (p close to either 1 or 0 would mean very little choice). The observed outcomes serve as a natural signal of P . However, there are no obvious counterparts of p in the data. These two probabilities are obviously the same in the absence of search. The difference between the two can at least partly be attributed to the preferences for the health insurance offer.

I will now present a static job search model where each job has two characteristics – wage and whether or not it offers health insurance coverage.⁶ The model concentrates on the population of workers – unemployment is not an option. Each individual i chooses a job to

⁵There are a number of good surveys of the empirical search literature. Canals and Stern (2001) is one of the recent comprehensive reviews of the recent developments in the estimation of the structural search models.

⁶Adopted from the product variety choice model of Rosen (1978). The important feature of the model is that it considers search along two dimensions – price of the product and its characteristics. Since the main purpose of this paper is to study the health insurance offers associated with jobs and wage cannot be assumed away from the analysis, we have to have at least two dimensions of search.

maximize utility given by:

$$U_i = \log w_i + \theta_i I_i, \quad (4.1)$$

where w_i is the wage, I_i is the dichotomous variable equal to one if the job has EPHI offer, and θ_i is a parameter reflecting the strength of the preference for EPHI.

The individuals in the model engage in a job search by drawing job offers from the joint distribution $\phi(\log w_i, I_i | \mu_{1i}, \mu_{2i}, p_i)$, which has the following form:⁷

$$\phi(\log w_i, I_i | \mu_{1i}, \mu_{2i}, p_i) = \begin{cases} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(\log w_i - \mu_{1i})^2} p_i, & \text{if } I_i = 1 \\ \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(\log w_i - \mu_{2i})^2} (1 - p_i), & \text{if } I_i = 0 \\ 0, & \text{otherwise,} \end{cases} \quad (4.2)$$

where μ_{1i}, μ_{2i} are the means of the wage distributions for two types of jobs – with and without health coverage. That is, conditional on the realization of I_i the distribution of $\log w_i$ is normal. The means of these two conditional distributions are allowed to be different to account for the fact that at least part of the cost of providing health insurance to employees may be shifted back onto employees in the form of reduced wages. We have the following marginal distributions:

$$\begin{aligned} \phi_1(\log w_i | \cdot) &= \sum_{I_i=0}^1 \phi(\log w_i, I_i | \cdot) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(\log w_i - \mu_{1i})^2} p_i + \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(\log w_i - \mu_{2i})^2} (1 - p_i) \\ \phi_2(I_i | \cdot) &= \int_{-\infty}^{\infty} \phi(\log w_i, I_i | \cdot) d \log w_i = \begin{cases} p_i, & \text{if } I_i = 1 \\ 1 - p_i, & \text{if } I_i = 0 \\ 0, & \text{otherwise.} \end{cases} \end{aligned}$$

The marginal distribution of $\log w_i$ is the normal mixture with component means given by $\{\mu_{1i}, \mu_{2i}\}$ and mixing distribution $\{p_i, 1 - p_i\}$ (the marginal distribution of I_i). The parameter qp_i (\equiv the probability to draw a job with an EPHI offer in any single job application) is of particular interest because it explicitly addresses the question whether different people have different degrees of flexibility as far as choice of whether to have a job with EPHI offer is

⁷The joint distribution is obtained by first specifying the marginal distribution of EPHI offers as a Bernoulli draws with probability p_i . Then, conditional on the outcome of the Bernoulli process for EPHI offer, wage is assumed to be distributed log-normally with means given by μ_{1i} (if $I_i = 1$) and μ_{2i} (if $I_i = 0$).

concerned. This parameter is unobserved by the econometrician. In addition to that, we do not have any direct signals of p (as opposed to P , which is reflected in observed (accepted) jobs with and without EPHI offers). The draws are assumed to be independent, and each draw entails incurring a utility cost of the size c_i .

Now consider the following transformation:

$$U = \log w + \theta I,$$

$$V = I.$$

This transformation is one-to-one, and the inverse is given by:

$$\log w = U - \theta V,$$

$$I = V.$$

The determinant of the Jacobian of this transformation is one:

$$|J| = \begin{vmatrix} \frac{\partial \log w}{\partial U} & \frac{\partial \log w}{\partial V} \\ \frac{\partial I}{\partial U} & \frac{\partial I}{\partial V} \end{vmatrix} = \begin{vmatrix} 1 & -\theta \\ 0 & 1 \end{vmatrix} = 1.$$

Therefore, the joint distribution of (U, V) is given by:

$$\varphi(U_i, V_i | \mu_{1i}, \mu_{2i}, p_i) = \begin{cases} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(U_i - (\theta_i + \mu_{1i}))^2} p_i, & \text{if } V_i = 1 \\ \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(U_i - \mu_{2i})^2} (1 - p_i), & \text{if } V_i = 0 \\ 0, & \text{otherwise.} \end{cases}$$

The marginal distribution of utility U_i is then:

$$\varphi(U_i | \cdot) = \sum_{V_i=0}^1 \varphi(U_i, V_i | \cdot) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(U_i - (\theta_i + \mu_{1i}))^2} p_i + \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(U_i - \mu_{2i})^2} (1 - p_i) \quad (4.3)$$

I now have a one-dimensional search problem, and the familiar reservation value property holds. An individual will accept a job if the utility from it is greater than the reservation

utility \bar{U}_i . The expected utility of the acceptable job offer gross of search costs is:

$$\tilde{U}_i = \int_{\bar{U}_i}^{\infty} U_i \frac{\varphi(U_i|\cdot)}{\pi} dU_i, \quad (4.4)$$

$$\text{where } \pi_i = \int_{\bar{U}_i}^{\infty} \varphi(U_i|\cdot) dU_i. \quad (4.5)$$

The ratio in the equation (4.4) above is the distribution of utility conditional on the offer being acceptable (the utility is above its reservation level \bar{U}_i) and π_i is the probability that the offer is acceptable. Taking the search cost into account, we obtain the following expression for the expected utility of the search process:⁸

$$EU_i = (\tilde{U}_i - c)\pi_i + (\tilde{U}_i - 2c)\pi_i(1 - \pi_i) + (\tilde{U}_i - 3c)\pi_i(1 - \pi_i)^2 + \dots = \tilde{U}_i - c/\pi_i \quad (4.6)$$

This problem has a reservation utility property – each individual maximizes her expected utility given above by accepting the first job that provides utility higher than the reservation utility level. The optimal reservation utility can be found by solving the following first order condition:⁹

$$c = \int_{\bar{U}_i}^{\infty} (U_i - \bar{U}_i)\varphi(U_i|\cdot) dU_i. \quad (4.7)$$

This first-order condition implicitly defines the optimal decision rule as a function of the parameters of the model:

$$\bar{U}_i^* = f(c_i, \theta_i, \mu_{1i}, \mu_{2i}, p_i). \quad (4.8)$$

4.3 Simulation

It is possible to demonstrate some important points using simulations of this simple model. Recall that of the main objectives of this research is to show that joint estimation of θ and p is necessary to be able to draw meaningful conclusions about either of them. Figures 4.1 and 4.2 show the different projections of the plot of the simulated proportions of people with EPHI

⁸Please refer to the appendix for details.

⁹Please, see appendix for details.

offer (P) given different values of θ and p . The plot depicts 5000 points, each of which was obtained by generating data according to the search model described in the previous section. The parameters were set at the following levels:

$$\mu_1 = 6$$

$$\mu_2 = 8$$

$$p \sim UNIF(0.5, 0.8)$$

$$\theta \sim UNIF(1, 3)$$

$$c = 0.3$$

All but two parameters of interest were fixed during simulations. The intention was to

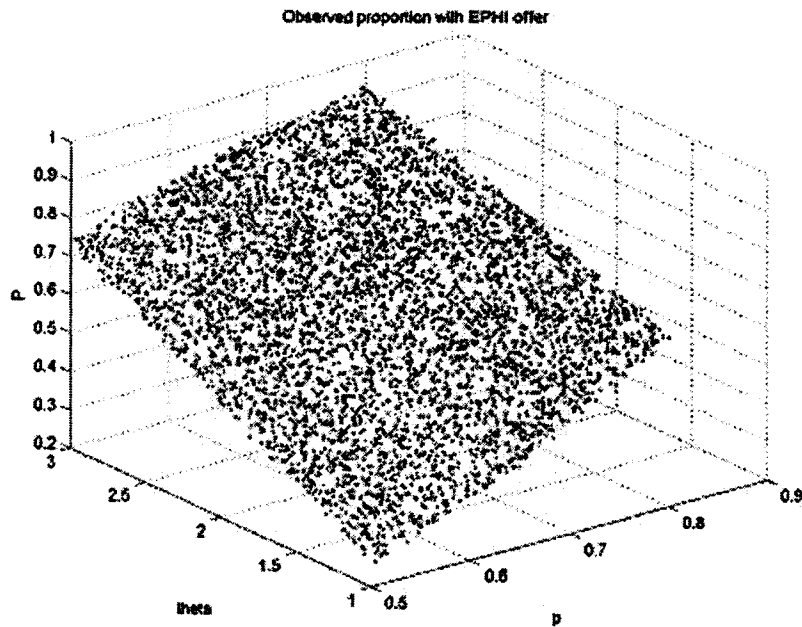


Figure 4.1 The trade-off between preferences and choice

show that observationally identical data can be generated with different combinations of θ and p . This illustrates the idea that some people may be observed with insurance offer because they always get offers (high p) even though they do not care about it (low θ), while others

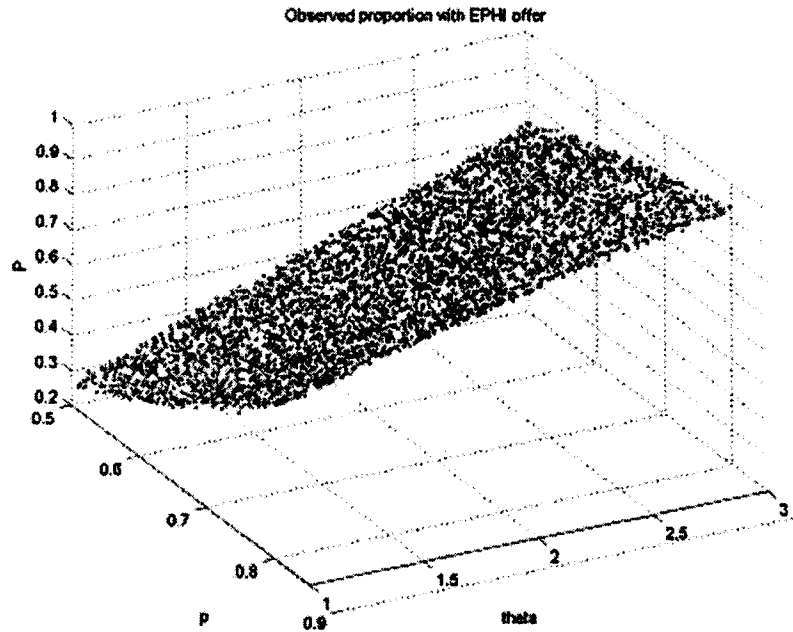


Figure 4.2 The trade-off between preferences and choice

are observed with insurance offer despite their low probability of getting one p but because they wanted it (high θ) and searched for a while. All individuals are assumed to be from the same population – they have exactly the same θ and p , they differ in the amount of luck governed by stochastic process generating job offers. Exactly because of this stochastic nature of the generated data it is impossible to define an analytic relationship between P and (θ, p) . Obviously, as sample size rises, the observed proportion P will converge in probability to EP . It is possible to think of defining the mapping from (θ, p) to EP with level curves that would contain pairs (θ, p) resulting in the same EP . Intuitively, it means that it may inappropriate to draw inferences about p , without explicitly accounting for θ . Conversely, we probably cannot say much about θ , if we do not explicitly account for p .

4.4 Homogeneous Population Case

4.4.1 Maximum-Likelihood Estimation

Given that the individual i acted optimally (the utility level of the accepted job is above \bar{U}_i^* , the probability that utility is above the reservation level is:

$$\begin{aligned}
\Pr[U_i > \bar{U}_i^*] &= \int_{\bar{U}_i^*}^{\infty} \left(\frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(U_i - (\theta_i + \mu_{1i}))^2} p_i + \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(U_i - \mu_{2i})^2} (1 - p_i) \right) dU_i \\
&= p_i \Pr[N(\theta_i + \mu_{1i}, 1) > \bar{U}_i^*] + (1 - p_i) \Pr[N(\mu_{2i}, 1) > \bar{U}_i^*] \\
&= p_i \Pr[N(0, 1) > \bar{U}_i^* - (\theta_i + \mu_{1i})] + (1 - p_i) \Pr[N(0, 1) > \bar{U}_i^* - \mu_{2i}] \\
&= p_i \Pr[N(0, 1) < (\theta_i + \mu_{1i}) - \bar{U}_i^*] + (1 - p_i) \Pr[N(0, 1) < \mu_{2i} - \bar{U}_i^*] \\
&= p_i F((\theta_i + \mu_{1i}) - \bar{U}_i^*) + (1 - p_i) F(\mu_{2i} - \bar{U}_i^*),
\end{aligned}$$

where $N(\mu, \sigma^2)$ is a normal random variable with the mean μ and variance σ^2 , and $F(\cdot)$ is the standard normal cdf.

The joint distribution of the $(\log w_i, I_i)$ is given by (conditional on $\mu_{1i}, \mu_{2i}, p_i, \theta_i, \bar{U}_i^*$):

$$\begin{aligned}
\phi(\log w_i, I_i | U_i > \bar{U}_i^*) &= \phi(\log w_i, I_i | \log w_i > \bar{U}_i^* - \theta_i I_i) \\
&= \frac{\phi(\log w_i, I_i)}{\Pr[\log w_i > \bar{U}_i^* - \theta_i I_i]} = \frac{\phi(\log w_i, I_i)}{\Pr[U_i > \bar{U}_i^*]} \\
&= \begin{cases} \frac{\frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(\log w_i - \mu_{1i})^2} p_i}{p_i F((\theta_i + \mu_{1i}) - \bar{U}_i^*) + (1 - p_i) F(\mu_{2i} - \bar{U}_i^*)}, & \text{if } I_i = 1, \log w_i > \bar{U}_i^* - \theta_i \\ \frac{\frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(\log w_i - \mu_{2i})^2} (1 - p_i)}{p_i F((\theta_i + \mu_{1i}) - \bar{U}_i^*) + (1 - p_i) F(\mu_{2i} - \bar{U}_i^*)}, & \text{if } I_i = 0, \log w_i > \bar{U}_i^* \\ 0, & \text{otherwise.} \end{cases}
\end{aligned}$$

Therefore, conditional on the parameters, the joint distribution of $\log w_i$ and I_i is the mixture of two truncated normals. Denoting the log-wage observations for those people with EPHI offer by $\log w_{1i}$, $i = \overline{1, n_1}$ and the log-wage observations of those without EPHI offer by $\log w_{2j}$,

$j = \overline{1, n_2}$, the joint likelihood is given by:

$$\begin{aligned}
L(\log w_{11}, \dots, \log w_{1n_1}, \log w_{21}, \dots, \log w_{2n_2}, I_1, \dots, I_n | \theta_i, \mu_{1i}, \mu_{2i}, p_i, \overline{U}_i^*) &= \\
&= (2\pi)^{-(n_1+n_2)/2} \prod_{i=1}^{n_1} \frac{p_i}{p_i F((\theta_i + \mu_{1i}) - \overline{U}_i^*) + (1 - p_i) F(\mu_{2i} - \overline{U}_i^*)} \times \\
&\times \prod_{j=1}^{n_2} \frac{1 - p_j}{p_j F((\theta_j + \mu_{1j}) - \overline{U}_j^*) + (1 - p_j) F(\mu_{2j} - \overline{U}_j^*)} \times \\
&\times e^{-1/2[\sum_{i=1}^{n_1} (\log w_{1i} - \mu_{1i})^2 + \sum_{j=1}^{n_2} (\log w_{2j} - \mu_{2j})^2]} \times \\
&\times \prod_{i=1}^{n_1} I[\log w_{1i} \geq \overline{U}_i^* - \theta_i] \times \prod_{j=1}^{n_2} I[\log w_{2j} \geq \overline{U}_j^*]
\end{aligned}$$

where $I[exp]$ is an indicator function equal to 1 if the expression in brackets is true and equal to 0 otherwise.

Let me start by simplifying the model by assuming that the parameters of the model $\{\theta_i, p_i, \mu_{1i}, \mu_{2i}, c_i\}$ are constant across population. The likelihood then becomes:

$$\begin{aligned}
L(\log w_{11}, \dots, \log w_{1n_1}, \log w_{21}, \dots, \log w_{2n_2}, I_1, \dots, I_n | \theta, \mu_1, \mu_2, p, \overline{U}^*) &= \\
&= (2\pi)^{-(n_1+n_2)/2} e^{-1/2[\sum_{i=1}^{n_1} (\log w_{1i} - \mu_1)^2 + \sum_{j=1}^{n_2} (\log w_{2j} - \mu_2)^2]} p^{n_1} (1 - p)^{n_2} \times \\
&\times (p F((\theta + \mu_1) - \overline{U}^*) + (1 - p) F(\mu_2 - \overline{U}^*))^{-(n_1+n_2)} I[\log w_{1(1)} \geq \overline{U}^* - \theta] I[\log w_{2(1)} \geq \overline{U}^*],
\end{aligned}$$

where $\log w_{1(1)}$ and $\log w_{2(1)}$ are the first order statistics for two groups of observations (with and without an EPHI offer respectively).

It is the standard procedure in search models to estimate the reservation value along with the parameters of the wage-EPHI distribution and then use them to get the estimate of search cost (which doesn't appear in likelihood but can be computed using the first-order condition). Consider the continuous part of the likelihood – everything except the last two terms. This part of the likelihood is increasing function of \overline{U}^* , therefore the Maximum Likelihood estimator of reservation utility is:

$$(\widehat{\overline{U}^*})_{MLE} = \min\{\log w_{1(1)} + \theta, \log w_{2(1)}\}.$$

Also, the continuous part of the likelihood is decreasing in θ . Now, if $(\widehat{U^*})_{MLE} = \log w_{2(1)} < \log w_{1(1)} + \theta$, then we can increase likelihood by decreasing θ . If, on the other hand, $(\widehat{U^*})_{MLE} = \log w_{1(1)} + \theta < \log w_{2(1)}$, then by decreasing θ (increase in log-likelihood with the speed proportional to $-\frac{\partial \log L(\cdot)}{\partial \theta} = (n_1 + n_2) \frac{pf((\mu_1 + \theta) - \overline{U^*})}{pF((\theta + \mu_1) - \overline{U^*}) + (1-p)F(\mu_2 - \overline{U^*})}$) we also decrease $(\widehat{U^*})_{MLE}$ (decrease in log-likelihood with the speed proportional to $\frac{\partial \log L(\cdot)}{\partial \overline{U^*}} = (n_1 + n_2) \frac{pf((\mu_1 + \theta) - \overline{U^*}) + (1-p)f((\mu_2 - \overline{U^*}))}{pF((\theta + \mu_1) - \overline{U^*}) + (1-p)F(\mu_2 - \overline{U^*})}$).¹⁰ Therefore, by increasing θ we will increase the likelihood. It follows then that:

$$(\widehat{U^*})_{MLE} = \log w_{1(1)} + \widehat{\theta}_{MLE} = \log w_{2(1)}$$

The first-order conditions with respect to other three parameters are:

$$\begin{aligned} \frac{\partial \log L(\cdot)}{\partial \mu_1} &= -\frac{1}{2}2(-1) \sum_{i=1}^{n_1} (\log w_{1i} - \mu_1) - \frac{(n_1 + n_2)pf((\mu_1 + \theta) - \overline{U^*})}{pF((\theta + \mu_1) - \overline{U^*}) + (1-p)F(\mu_2 - \overline{U^*})} = 0 \\ \frac{\partial \log L(\cdot)}{\partial \mu_2} &= -\frac{1}{2}2(-1) \sum_{j=1}^{n_2} (\log w_{2j} - \mu_2) - \frac{(n_1 + n_2)(1-p)f(\mu_2 - \overline{U^*})}{pF((\theta + \mu_1) - \overline{U^*}) + (1-p)F(\mu_2 - \overline{U^*})} = 0 \\ \frac{\partial \log L(\cdot)}{\partial p} &= \frac{n_1}{p} + \frac{n_2}{1-p}(-1) - \frac{(n_1 + n_2)(F((\theta + \mu_1) - \overline{U^*}) - F(\mu_2 - \overline{U^*}))}{pF((\theta + \mu_1) - \overline{U^*}) + (1-p)F(\mu_2 - \overline{U^*})} = 0. \end{aligned}$$

With some transformations we arrive at the following system of the Maximum Likelihood estimators (some are indirect expressions):¹¹

$$\begin{aligned} \overline{\log w_1} &= (\widehat{\mu_1}) + \frac{(n_1 + n_2)\widehat{p}f((\widehat{\mu_1}) + \widehat{\theta}(\widehat{U^*}))}{n_1(\widehat{p}F((\widehat{\mu_1}) + \widehat{\theta}(\widehat{U^*})) + (1-\widehat{p})F((\widehat{\mu_2}) - (\widehat{U^*})))} \\ \overline{\log w_2} &= (\widehat{\mu_2}) + \frac{(n_1 + n_2)(1-\widehat{p})f((\widehat{\mu_2}) - (\widehat{U^*}))}{n_2(\widehat{p}F((\widehat{\mu_1}) + \widehat{\theta}(\widehat{U^*})) + (1-\widehat{p})F((\widehat{\mu_2}) - (\widehat{U^*})))} \\ \frac{n_1(1-\widehat{p}) - n_2\widehat{p}}{\widehat{p}(1-\widehat{p})} &= \frac{(n_1 + n_2)(F((\widehat{\mu_1}) + \widehat{\theta}(\widehat{U^*})) - F((\widehat{\mu_2}) - (\widehat{U^*})))}{\widehat{p}F((\widehat{\mu_1}) + \widehat{\theta}(\widehat{U^*})) + (1-\widehat{p})F((\widehat{\mu_2}) - (\widehat{U^*}))} \\ (\widehat{U^*}) &= \log w_{2(1)} \\ \widehat{\theta} &= \log w_{2(1)} - \log w_{1(1)}, \end{aligned}$$

¹⁰where $f(\cdot)$ is the standard normal pdf.

¹¹MLE subscript is suppressed in this equation.

where $\overline{\log w_1}$ and $\overline{\log w_2}$ are means of the observed log-wages for both subpopulations (with and without an EPHI offer). The Maximum Likelihood estimator of search cost follows directly from first-order conditions:

$$\hat{c}_{MLE} = \int_{(\overline{U^*})_{MLE}}^{\infty} (U - (\overline{U^*})_{MLE}) \varphi(U_i | (\widehat{\mu_1})_{MLE}, (\widehat{\mu_2})_{MLE}, \widehat{\theta}_{MLE}, \widehat{p}_{MLE}) dU_i$$

4.4.2 Monte-Carlo Experiment

I have performed simulations of this simplified homogeneous search model. All the parameters in the model were randomly drawn from a distributions specified below for each run of the program. In each iteration, the data were generated according to the search model outlined above. The data consist of two vectors – observed wages and dichotomous EPHI offer state variable. The distributions for parameters were:

$$\mu_1 \sim UNIF(5, 8)$$

$$\mu_2 \sim UNIF(7, 10)$$

$$p \sim UNIF(0.5, 0.8)$$

$$\theta \sim UNIF(1, 3)$$

$$c = 0.3.$$

The simulations were run 100 times. Each iteration 10000 data points (individuals) were created. The results are displayed in Figure 1. The horizontal axis on each subplot indicates the true parameter selected at the beginning of each trial by pseudo-random number generator. The vertical axis indicates the estimate of the parameter. It has to be noted that in some cases, quite unusual combinations of parameters were drawn, leading to unexpected distribution of data. In some of these cases, the search algorithm needed to solve the system of first-order conditions failed to converge leading to estimates that were clearly quite far from the true parameter. Yet, the estimation of both means of wage processes as well as preference parameter θ seems to be quite robust. The most sensitive parameter appears to be p , which wasn't

estimated accurately in some cases. This simplified homogeneous population version of the

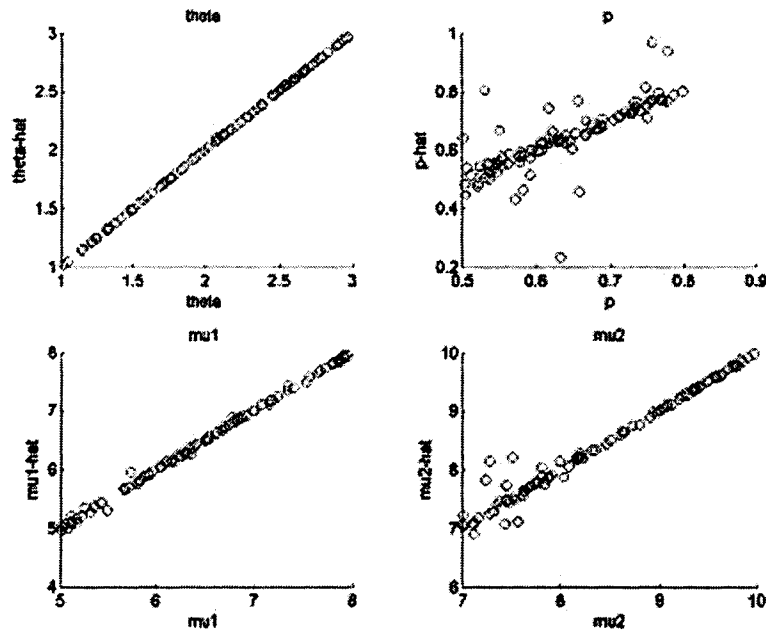


Figure 4.3 MLE, Homogeneous Population

model has one major drawback – it is too simple. It is unrealistic to assume that everyone in the sample draws values from exactly the same distribution.

4.5 Directions for Future Research

In order to model individual heterogeneity in a realistic way, all the main parameters of the model ($\{\theta_i, p_i, \mu_{1i}, \mu_{2i}, c_i\}$) have to be allowed to depend on the observable data. Observed heterogeneity is often modeled in the search models via splitting the sample into finite number of “homogeneous” groups. It is done by identifying the (limited) number of characteristics that are observed by econometrician and creating groups for which these characteristics are the same. All continuous variables have to be discretized into intervals with the assumption that inside the interval differences in the value of a characteristic do not cause heterogeneous behavior. In my opinion, the assumption that all heterogeneity can be accounted for by a few

observed variables as well as discretization may well be too restrictive.¹²

So far I have ignored any potential unobserved heterogeneity. Some people may be more productive and be able to get better jobs in all senses – health insurance and higher wages. To incorporate unobserved heterogeneity, some or all parameters in the model have to be specified as functions of observables and random errors. The problem is that this makes the standard estimation procedure impossible to use because the reservation utility has to be estimated first. If individuals are different based on some unobserved factors, the reservation utility levels are going to be different for everyone in the sample. This means that the order statistics used to identify the preference parameter θ and the reservation utility are no longer useful.

It is a well-known fact that the structural search models are extremely sensitive to the measurement error in observed wages. The reason is that order-statistics are used in estimators of reservation utility levels. Order statistics are known to be sensitive to the measurement error. Inspection of the data leads me to believe that measurement error might be an important issue in this model. It is, therefore, important to consider the issue of measurement error and to try to incorporate it in the estimation procedure.

This paper should serve as a starting point for research aiming to empirically separate the effects of preferences for and availability of choice of the employer-provided health coverage. I show that these two factors can interact in complex ways to produce similar observable outcomes. I also show that there are some problems related to the estimation of the realistic version of my search model, which may require some new solutions.

¹²This is partly based on several attempts to estimate the model for a few population groups with similar demographic characteristics. The estimates in all cases were unrealistic and they are not reported here. In particular, the estimated search cost were very high for almost all groups. This means that effectively there was almost no search. For some subpopulations the estimates of the preference were negative.

CHAPTER 5. GENERAL CONCLUSION

In this dissertation, I explore the role played by private information in different markets. There are two main classes of imperfect information models that provide testable predictions for empirical work – adverse selection and moral hazard. Both adverse selection and moral hazard predict that there should be positive correlation between some choice variable (insurance coverage) and some outcome variable (risk). Although the prediction of both of these models is the same, the mechanisms that generate it are different. In adverse selection models, agents know their risk and choose insurance coverage. In moral hazard models, agents change their own risk depending on the coverage they have.

The main contribution in the first essay is to consider a conceptually different type of private information – risk aversion. Recent theoretical advances demonstrate that the lack of positive correlation between coverage and risk does not signal the absence of informational asymmetries and is consistent with more complex models with unobserved risk-aversion. I extend the empirical literature by testing for selection on more than one type of latent information - risk type and risk-aversion type. I separate the incentive effect from effects of selection on multiple types of unobservables by specifying a hybrid endogenous treatment model with explicit modeling of indicators of latent attitudes toward risk-taking behavior. I use data from the 2000 Medical Expenditure Panel Survey (MEPS), which contains a set of attitudinal variables necessary to estimate the model. Although the overall selection effect appears to be insignificant, the results indicate that individuals do, in fact, possess private information which increases their propensity to be insured and to utilize health care. This suggests that lack of conditional correlation between insurance coverage and health care utilization results from informational asymmetries of multiple types inducing selection in opposite directions. In addition, I find

strong evidence of moral hazard in outpatient and office-based health care utilization but not in inpatient or emergency room utilization.

In the second essay, I consider financing of business start-ups. This is the first paper to examine the relationship between capital structure and performance of business start-ups in the presence of imperfect information. Capital structure and performance of business start-ups are estimated jointly using a unique data set collected by the National Federation of Independent Business (NFIB) Foundation. The results suggest that debt does not have a significant incentive effect on performance in business start-ups after controlling for self-selection. In contrast, both selection and incentive effects are present in the case of outside equity indicating that outside investors are able to both overcome informational opaqueness of business start-ups and provide better incentives for performance.

The third essay examines the role of unobserved preferences for health insurance in the context of job search. The majority of Americans get health insurance through their employers. One of the potential concerns related to employer-provided health insurance is whether some categories of workers (typically with low educational attainment) actually have access to jobs with health insurance offers. This challenges an assumption imposed in some of the previous literature that workers have equal access to both jobs with and without health insurance offers and sort strategically according to their preferences. A worker can be uninsured either because of lack of choice or because of preferences for a job without a health insurance offer. To be able to separate the effect of lack of choice from the effect of preferences, I develop a structural job search model with two job characteristics - the offered wage and the availability of insurance. I simulate the model to demonstrate the interaction between preferences and availability of choice.

APPENDIX A. Posterior distributions for chapter 2

Likelihood of the Latent Data

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

$$\begin{aligned}
p(y^*|\beta, \Sigma) &= (2\pi)^{-\frac{3n}{2}} |I_n \otimes \Sigma|^{-\frac{1}{2}} \exp\left(-\frac{1}{2}(y^* - X\beta)'(I_n \otimes \Sigma)^{-1}(y^* - X\beta)\right) \\
&\propto (|I_n|^3 |\Sigma|^n)^{-\frac{1}{2}} \exp\left(-\frac{1}{2}(y^* - X\beta)'(I_n \otimes \Sigma^{-1})(y^* - X\beta)\right) \\
&\propto |\Sigma|^{-\frac{n}{2}} \exp\left(-\frac{1}{2} \sum_{i=1}^n \varepsilon_i' \Sigma^{-1} \varepsilon_i\right) \\
&\propto |\Sigma|^{-\frac{n}{2}} \exp\left(-\frac{1}{2} \sum_{i=1}^n (y_i^* - X_i \beta)' \Sigma^{-1} (y_i^* - X_i \beta)\right) \tag{A.1}
\end{aligned}$$

Posterior of β

$$\begin{aligned}
p(\beta|y^*, \Sigma) &\propto \exp\left(-\frac{1}{2} \left[\sum_{i=1}^n (y_i^* - X_i \beta)' \Sigma^{-1} (y_i^* - X_i \beta) + (\beta - \mu_{\beta 0})' V_{\beta 0}^{-1} (\beta - \mu_{\beta 0}) \right]\right) \\
&\propto \exp\left(-\frac{1}{2} \left[\sum_{i=1}^n (y_i^{*'} \Sigma^{-1} y_i^* - 2\beta' X_i' \Sigma^{-1} y_i^* + \beta' X_i' \Sigma^{-1} X_i \beta) \right. \right. \\
&\quad \left. \left. + (\beta' V_{\beta 0}^{-1} \beta - 2\beta' V_{\beta 0}^{-1} \mu_{\beta 0} + \mu_{\beta 0}' V_{\beta 0}^{-1} \mu_{\beta 0}) \right]\right) \\
&\propto \exp\left(-\frac{1}{2} \left[\beta' \left(\sum_{i=1}^n X_i' \Sigma^{-1} X_i + V_{\beta 0}^{-1} \right) \beta - 2\beta' \left(\sum_{i=1}^n X_i' \Sigma^{-1} y_i^* + V_{\beta 0}^{-1} \mu_{\beta 0} \right) \right]\right) \\
&\propto \exp\left(-\frac{1}{2} \left[\beta' V_{\beta 1}^{-1} \beta - 2\beta' V_{\beta 1}^{-1} \mu_{\beta 1} \right]\right) \\
&\propto \exp\left(-\frac{1}{2} \left[\beta' V_{\beta 1}^{-1} \beta - 2\beta' V_{\beta 1}^{-1} \mu_{\beta 1} + \mu_{\beta 1}' V_{\beta 1}^{-1} \mu_{\beta 1} - \mu_{\beta 1}' V_{\beta 1}^{-1} \mu_{\beta 1} \right]\right) \\
&\propto \exp\left(-\frac{1}{2} \left[(\beta - \mu_{\beta 1})' V_{\beta 1}^{-1} (\beta - \mu_{\beta 1}) \right]\right) \tag{A.2}
\end{aligned}$$

Therefore,

$$\begin{aligned} p(\beta|y^*, \Sigma) &= N(\mu_{\beta 1}, V_{\beta 1}) \\ V_{\beta 1} &= \left(\sum_{i=1}^n X_i' \Sigma^{-1} X_i + V_{\beta 0}^{-1} \right)^{-1} \\ \mu_{\beta 1} &= V_{\beta 1} \left(\sum_{i=1}^n X_i' \Sigma^{-1} y_i^* + V_{\beta 0}^{-1} \mu_{\beta 0} \right) \end{aligned} \tag{A.3}$$

APPENDIX B. Posterior distributions and posterior output for chapter 3

Likelihood

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

$$\begin{aligned}
p(y|\beta, \Omega) &= (2\pi)^{-\frac{2n}{2}} |I_n \otimes H^{-1}|^{-\frac{1}{2}} \exp\left(-\frac{1}{2}(y - X\beta)'(I_n \otimes H^{-1})^{-1}(y - X\beta)\right) \\
&\propto (|I_n|^2 |H^{-1}|^n)^{-\frac{1}{2}} \exp\left(-\frac{1}{2}(y - X\beta)'(I_n \otimes H)(y - X\beta)\right) \\
&\propto |H|^{\frac{n}{2}} \exp\left(-\frac{1}{2} \sum_{i=1}^n \eta_i' H \eta_i\right) \\
&\propto |H|^{\frac{n}{2}} \exp\left(-\frac{1}{2} \sum_{i=1}^n (y_i - X_i\beta)' H (y_i - X_i\beta)\right) \\
&\propto |H|^{\frac{n}{2}} \exp\left(-\frac{1}{2} \text{tr}\left[\sum_{i=1}^n (y_i - X_i\beta)' H (y_i - X_i\beta)\right]\right) \tag{B.1}
\end{aligned}$$

Posterior distributions

Posterior of β

Since we specify the independent priors for β and H^{-1} , the conditional posterior for β is proportional to

$$\begin{aligned}
p(\beta|data, H^{-1}) &\propto \exp\left(-\frac{1}{2}\left[\sum_{i=1}^n (y_i - X_i\beta)'H(y_i - X_i\beta) + (\beta - \beta_0)'V_0^{-1}(\beta - \beta_0)\right]\right) \\
&\propto \exp\left(-\frac{1}{2}\left[\sum_{i=1}^n (y_i'H y_i - 2\beta'X_i'H y_i + \beta'X_i'H X_i\beta) \right. \right. \\
&\quad \left. \left. + (\beta'V_0^{-1}\beta - 2\beta'V_0^{-1}\beta_0 + \beta_0'V_0^{-1}\beta_0)\right]\right) \\
&\propto \exp\left(-\frac{1}{2}\left[\beta'\left(\sum_{i=1}^n X_i'H X_i + V_0^{-1}\right)\beta - 2\beta'\left(\sum_{i=1}^n X_i'H y_i + V_0^{-1}\beta_0\right)\right]\right) \\
&\propto \exp\left(-\frac{1}{2}\left[\beta'V_1^{-1}\beta - 2\beta'V_1^{-1}\beta_1\right]\right) \\
&\propto \exp\left(-\frac{1}{2}\left[\beta'V_1^{-1}\beta - 2\beta'V_1^{-1}\beta_1 + \beta_1'V_1^{-1}\beta_1 - \beta_1'V_1^{-1}\beta_1\right]\right) \\
&\propto \exp\left(-\frac{1}{2}\left[(\beta - \beta_1)'V_1^{-1}(\beta - \beta_1)\right]\right) \tag{B.2}
\end{aligned}$$

The conditional posterior distribution of β is also normal

$$\begin{aligned}
p(\beta|data, H) &= N(\beta_1, V_1) \\
V_1 &= \left(\sum_{i=1}^n X_i'H X_i + V_0^{-1}\right)^{-1} \\
\beta_1 &= V_1\left(\sum_{i=1}^n X_i'H y_i + V_0^{-1}\beta_0\right) \tag{B.3}
\end{aligned}$$

Posterior of H^{-1}

The first case is continuous Y_i . Independence of prior distributions leads to the following conditional posterior distribution of H

$$\begin{aligned}
p(H) &\propto |H|^{\frac{n}{2}} \exp\left(-\frac{1}{2} \text{tr}\left[\sum_{i=1}^n (y_i - X_i\beta)' H (y_i - X_i\beta)\right]\right) \times \\
&\quad |H|^{\frac{\nu_0 - 2 - 1}{2}} \exp\left(-\frac{1}{2} \text{tr}(H_0^{-1} H)\right) \\
&\propto |H|^{\frac{\nu_0 - n - 2 - 1}{2}} \exp\left(-\frac{1}{2} \text{tr}\left[\sum_{i=1}^n (y_i - X_i\beta)(y_i - X_i\beta)' H + H_0^{-1} H\right]\right) \quad (\text{B.4})
\end{aligned}$$

Therefore, the conditional posterior distribution of H is also Wishart

$$\begin{aligned}
p(H|data, \beta) &= W(\nu_1, H_1) \\
\nu_1 &= \nu_0 + n \\
H_1 &= \left(\sum_{i=1}^n (y_i - X_i\beta)(y_i - X_i\beta)' + H_0^{-1}\right)^{-1} \quad (\text{B.5})
\end{aligned}$$

The other case that we will consider is of a binary dependent variable Y_i . The conditional posterior of ρ is given by

$$\begin{aligned}
p(\rho|data, \beta, h) &\propto \exp\left(-\frac{1}{2} \sum_{i=1}^n (\eta_{1i}^2 h \rho^2 - 2\eta_{1i}\eta_{2i} h \rho)\right) \exp\left(-\frac{1}{2V_{\rho 0}}(\rho^2 - 2\rho\rho_0)\right) \\
&\propto \exp\left(-\frac{1}{2}(\rho^2(V_{\rho 0}^{-1} + h \sum_{i=1}^n \eta_{1i}^2)) - 2\rho(\rho_0 V_{\rho 0}^{-1} + h \sum_{i=1}^n \eta_{1i}\eta_{2i})\right) \\
&\propto \exp\left(-\frac{1}{2V_{\rho 1}}(\rho - \rho_1)^2\right) \quad (\text{B.6})
\end{aligned}$$

Thus, the conditional posterior of ρ is normal

$$\begin{aligned}
 p(\rho|data, \beta, h) &= N(\rho_1, V_{\rho 1}) \\
 V_{\rho 1} &= (V_{\rho 0}^{-1} + h \sum_{i=1}^n \eta_{1i}^2)^{-1} \\
 \rho_1 &= V_{\rho 1}(\rho_0 V_{\rho 0}^{-1} + h \sum_{i=1}^n \eta_{1i} \eta_{2i})
 \end{aligned} \tag{B.7}$$

The conditional posterior of h is given by

$$\begin{aligned}
 p(h|data, \beta, \rho) &\propto h^{\frac{\nu_0-2}{2}} \exp\left(-\frac{h\nu_0}{2h_0}\right) h^{\frac{n}{2}} \exp\left(-\frac{1}{2} \sum_{i=1}^n (\eta_{1i}^2 (1+h\rho^2) - 2\eta_{1i}\eta_{2i}h\rho + \eta_{2i}^2 h)\right) \\
 &\propto h^{\frac{\nu_0+n-2}{2}} \exp\left(-h\left[\frac{\nu_0}{2h_0} + \frac{1}{2} \sum_{i=1}^n (\eta_{1i}^2 \rho^2 - 2\eta_{1i}\eta_{2i}\rho + \eta_{2i}^2)\right]\right) \\
 &\propto h^{\frac{\nu_0+n-2}{2}} \exp\left(-\frac{h(\nu_0+n)}{2} \left[\frac{\nu_0}{h_0(\nu_0+n)} + \frac{\sum_{i=1}^n (\eta_{1i}\rho - \eta_{2i})^2}{(\nu_0+n)}\right]\right) \\
 &\propto h^{\frac{\nu_1-2}{2}} \exp\left(-\frac{h\nu_1}{2h_1}\right)
 \end{aligned} \tag{B.8}$$

Thus, the conditional posterior of h is given by

$$\begin{aligned}
 p(h|data, \beta, \rho) &= G(\nu_1, h_1) \\
 \nu_1 &= \nu_0 + n \\
 h_1 &= \left[\frac{\nu_0}{h_0(\nu_0+n)} + \frac{\sum_{i=1}^n (\eta_{1i}\rho - \eta_{2i})^2}{(\nu_0+n)}\right]^{-1}
 \end{aligned} \tag{B.9}$$

Posterior Output

Table B.1 Results

variable	posterior mean, equation1	posterior st.dev. equation1	posterior mean, equation2	posterior st.dev. equation2
intercept	-0.7091	0.2332	0.0613	0.1299
age1	0.0102	0.0031	-0.0063	0.0016
totjobs	-0.0232	0.0061	-0.0066	0.0038
bs11	-0.0041	0.0016	-0.0000	0.0009
bs21	0.0026	0.0014	-0.0000	0.0007
bs61	-0.0051	0.0023	0.0003	0.0013
race	-0.3578	0.1080	-0.0610	0.0636
outjob4	0.1334	0.0737	-0.0276	0.0401
unemp_change	-0.0762	0.0232		
est_change	0.3950	0.3185		
sex	-0.1935	0.0681	0.1375	0.0371
moved1	-0.0021	0.0697	-0.1336	0.0369
partners	0.0303	0.0377	-0.0283	0.0197
managexp	-0.0521	0.0689	-0.0575	0.0357
firequit	-0.0563	0.0725	-0.0620	0.0398
diffprod	-0.2638	0.0582	-0.0171	0.0311
parenown	0.0782	0.0552	-0.0509	0.0290
educlev	0.0161	0.0169	-0.0028	0.0087
opercont	0.0744	0.0249	0.0095	0.0130
capinv	0.0575	0.0181	0.1317	0.0097
oddsyr	0.0861	0.0129	0.0015	0.0071
agef	0.0088	0.0041	0.0013	0.0021
lgl_form1_2	-0.0684	0.0909	-0.0544	0.0491
lgl_form1_3	-0.0762	0.0662	-0.1980	0.0362
loans	-0.0535	0.1821		
homestead			-0.0025	0.0012
h_unlimited			0.2847	0.0992
HH			-0.0001	0.0000
debt_2_assets			0.0147	0.1505
start1			-0.0479	0.0313
covariance	0.0025	0.0529		

Notes: The sample size is 2880. The dependent variable in equation 1 is *survive* and the dependent variable in the second equation is *loans*. The last line is the posterior mean and standard error of covariance parameter σ_{12} .

Table B.2 Results

variable	posterior mean, equation1	posterior st.dev. equation1	posterior mean, equation2	posterior st.dev. equation2
intercept	-0.6684	0.2242	-1.7095	0.3640
age1	0.0102	0.0031	-0.0027	0.0043
totjobs	-0.0221	0.0061	0.0111	0.0085
bs11	-0.0044	0.0016	-0.0046	0.0025
bs21	0.0025	0.0013	-0.0021	0.0019
bs61	-0.0050	0.0023	0.0022	0.0032
race	-0.3399	0.1077	0.2159	0.1525
outjob4	0.1331	0.0735	-0.0015	0.1121
unemp_change	-0.0748	0.0232		
est_change	0.3929	0.3126		
sex	-0.2044	0.0681	-0.1756	0.1123
moved1	0.0009	0.0693	0.0835	0.0927
partners	0.0638	0.0389	0.2987	0.0418
managexp	-0.0501	0.0692	0.0404	0.1013
firequit	-0.0619	0.0722	-0.1196	0.1147
diffprod	-0.2795	0.0587	-0.2522	0.0868
parenown	0.0673	0.0558	-0.1561	0.0784
educlev	0.0170	0.0167	0.0100	0.0228
opercont	0.0676	0.0249	-0.0820	0.0356
capinv	0.0633	0.0180	0.1032	0.0248
oddsyr	0.0836	0.0128	-0.0228	0.0190
agef	0.0078	0.0041	-0.0077	0.0058
lgl_form1_2	-0.0469	0.0906	0.4040	0.1235
lgl_form1_3	-0.0439	0.0687	0.4347	0.0947
invest	-0.9568	0.3560		
homestead			-0.0012	0.0031
h_unlimited			0.1588	0.2516
HH			-0.0000	0.0001
debt_2_assets			0.2765	0.3842
start1			0.1606	0.0854
covariance	0.2946	0.1188		

Notes: The sample size is 2880. The dependent variable in equation 1 is *survive* and the dependent variable in the second equation is *invest*. The last line is the posterior mean and standard error of covariance parameter σ_{12} .

Table B.3 Results

variable	posterior mean, equation1	posterior st.dev. equation1	posterior mean, equation2	posterior st.dev. equation2
intercept	-0.7057	0.2480	0.1644	0.1187
age1	0.0101	0.0032	-0.0055	0.0015
totjobs	-0.0233	0.0061	-0.0036	0.0034
bs11	-0.0041	0.0016	-0.0008	0.0008
bs21	0.0026	0.0013	-0.0003	0.0006
bs61	-0.0051	0.0023	0.0005	0.0011
race	-0.3570	0.1084	-0.0202	0.0569
outjob4	0.1325	0.0738	-0.0250	0.0367
unemp_change	-0.0750	0.0233		
est_change	0.4023	0.3172		
sex	-0.1914	0.0705	0.0997	0.0339
moved1	-0.0052	0.0717	-0.1128	0.0332
partners	0.0326	0.0382	0.0470	0.0173
managexp	-0.0534	0.0692	-0.0467	0.0325
firequit	-0.0585	0.0732	-0.0720	0.0361
diffprod	-0.2652	0.0587	-0.0442	0.0282
parenown	0.0759	0.0570	-0.0757	0.0263
educlev	0.0159	0.0169	-0.0019	0.0078
opercont	0.0742	0.0246	-0.0058	0.0118
capinv	0.0604	0.0252	0.1277	0.0088
oddsyr	0.0858	0.0129	-0.0020	0.0065
agef	0.0088	0.0041	-0.0003	0.0019
lgl_form1.2	-0.0673	0.0902	0.0057	0.0438
lgl_form1.3	-0.0778	0.0693	-0.1125	0.0322
outside	-0.0537	0.2724		
homestead			-0.0023	0.0011
h_unlimited			0.2773	0.0891
HH			-0.0000	0.0000
debt_2_assets			0.0459	0.1370
start1			-0.0224	0.0286
covariance	0.0111	0.0698		

Notes: The sample size is 2880. The dependent variable in equation 1 is *survive* and the dependent variable in the second equation is *outside*. The last line is the posterior mean and standard error of covariance parameter σ_{12} .

Table B.4 Results

variable	posterior mean, equation1	posterior st.dev. equation1	posterior mean, equation2	posterior st.dev. equation2
intercept	-8.1079	1.7917	-0.1069	0.1842
age1	0.0553	0.0265	-0.0035	0.0023
totjobs	-0.1866	0.0737	-0.0116	0.0059
bs11	-0.0205	0.0153	-0.0007	0.0012
bs21	-0.0013	0.0122	0.0008	0.0010
bs61	-0.0294	0.0206	0.0012	0.0017
race	-2.3178	1.0472	-0.0141	0.0843
outjob4	0.1296	0.6645	0.0151	0.0551
unemp_change	-0.3126	0.2158		
est_change	1.5374	2.0836		
sex	-1.4436	0.6588	0.1488	0.0528
moved1	-0.7656	0.6180	-0.1244	0.0526
partners	0.6354	0.3203	-0.0193	0.0279
managexp	0.2031	0.6132	-0.1060	0.0499
firequit	-0.1470	0.6529	-0.0434	0.0566
diffprod	-1.9227	0.5230	0.0303	0.0435
parenown	0.1483	0.4990	-0.0694	0.0415
educlev	0.0607	0.1402	-0.0022	0.0121
opercont	0.3545	0.2182	0.0055	0.0183
capinv	0.2027	0.1742	0.1270	0.0139
oddsyr	0.4391	0.1227	0.0084	0.0103
agef	0.0415	0.0358	0.0022	0.0031
lgl_form1.2	-0.8413	0.7984	-0.1198	0.0691
lgl_form1.3	-0.1030	0.6007	-0.2309	0.0499
loans	-0.0710	1.2862		
homestead			-0.0021	0.0018
h_unlimited			0.2142	0.1420
HH			-0.0001	0.0000
debt_2_assets			0.2011	0.2115
start1			-0.0615	0.0443
covariance	0.2876	0.3531		

Notes: The sample size is 1522. The dependent variable in equation 1 is *changemp* and the dependent variable in the second equation is *loans*. The last line is the posterior mean and standard error of covariance parameter σ_{12} .

Table B.5 Results

variable	posterior mean, equation1	posterior st.dev. equation1	posterior mean, equation2	posterior st.dev. equation2
intercept	-8.4380	1.7229	-2.3603	0.6071
age1	0.0573	0.0255	-0.0038	0.0071
totjobs	-0.1907	0.0686	0.0224	0.0153
bs11	-0.0184	0.0149	-0.0039	0.0038
bs21	0.0014	0.0122	-0.0017	0.0032
bs61	-0.0319	0.0204	0.0045	0.0048
race	-2.0905	0.9958	0.0940	0.2429
outjob4	0.1027	0.6453	-0.1023	0.1666
unemp_change	-0.3193	0.2018		
est_change	1.4811	2.0158		
sex	-1.2493	0.6244	-0.3384	0.1796
moved1	-0.8581	0.6192	0.2008	0.1486
partners	0.2490	0.3185	0.3205	0.0737
managexp	0.2164	0.6186	0.0777	0.1654
firequit	-0.2236	0.6606	-0.1525	0.1944
diffprod	-1.6375	0.5150	-0.1466	0.1347
parenown	0.2832	0.4795	-0.1931	0.1294
educlev	0.0740	0.1405	0.0059	0.0375
opercont	0.4045	0.2115	-0.1082	0.0556
capinv	0.1420	0.1548	0.1010	0.0404
oddsyr	0.4235	0.1202	-0.0242	0.0296
agef	0.0482	0.0369	0.0071	0.0094
lgl_form1_2	-1.1488	0.8111	0.7495	0.1978
lgl_form1_3	-0.2851	0.5761	0.4652	0.1529
invest	11.2706	1.7684		
homestead			0.0015	0.0051
h_unlimited			-0.0915	0.4155
HH			-0.0001	0.0001
debt_2_assets			0.8023	0.5930
start1			0.2989	0.1372
covariance	-7.4793	1.0855		

Notes: The sample size is 1522. The dependent variable in equation 1 is *changemp* and the dependent variable in the second equation is *invest*. The last line is the posterior mean and standard error of covariance parameter σ_{12} .

Table B.6 Results

variable	posterior mean, equation1	posterior st.dev. equation1	posterior mean, equation2	posterior st.dev. equation2
intercept	-7.9607	1.8010	0.0151	0.1656
age1	0.0529	0.0264	-0.0041	0.0021
totjobs	-0.1831	0.0726	-0.0036	0.0050
bs11	-0.0211	0.0152	-0.0009	0.0011
bs21	-0.0010	0.0122	0.0004	0.0009
bs61	-0.0278	0.0209	0.0018	0.0015
race	-2.3629	1.0375	-0.0252	0.0749
outjob4	0.1441	0.6425	0.0111	0.0503
unemp_change	-0.3053	0.2170		
est_change	1.3930	2.1640		
sex	-1.4079	0.6419	0.0936	0.0469
moved1	-0.7742	0.6102	-0.0775	0.0478
partners	0.6624	0.3176	0.0562	0.0245
managexp	0.1782	0.6318	-0.0903	0.0462
firequit	-0.1611	0.6653	-0.0507	0.0499
diffprod	-1.9113	0.5297	0.0003	0.0392
parenown	0.1021	0.4986	-0.0973	0.0372
educlev	0.0607	0.1417	-0.0044	0.0108
opercont	0.3547	0.2152	-0.0124	0.0166
capinv	0.2385	0.1802	0.1217	0.0124
oddsyr	0.4415	0.1255	0.0030	0.0091
agef	0.0435	0.0363	0.0010	0.0027
lgl_form1.2	-0.8155	0.8067	-0.0064	0.0618
lgl_form1.3	-0.1363	0.5801	-0.1463	0.0452
outside	-0.6937	1.3661		
homestead			-0.0022	0.0016
h_unlimited			0.2249	0.1259
HH			-0.0001	0.0000
debt_2_assets			0.3617	0.1923
start1			-0.0159	0.0397
covariance	0.3154	0.3467		

Notes: The sample size is 1522. The dependent variable is equation 1 is *changemp* and the dependent variable in the second equation is *outside*. The last line is the posterior mean and standard error of covariance parameter σ_{12} .

Table B.7 Results

variable	posterior mean, equation1	posterior st.dev. equation1	posterior mean, equation2	posterior st.dev. equation2
intercept	-0.9181	1.3887	0.0596	0.1318
age1	0.0006	0.0197	-0.0063	0.0016
totjobs	-0.0990	0.0444	-0.0067	0.0038
bs11	-0.0041	0.0112	0.0000	0.0009
bs21	0.0035	0.0089	0.0000	0.0007
bs61	-0.0261	0.0156	0.0003	0.0013
race	-1.9278	0.7678	-0.0605	0.0639
outjob4	0.9418	0.4967	-0.0281	0.0403
unemp_change	-0.2487	0.1565		
est_change	1.1461	1.7195		
sex	-0.8855	0.4791	0.1376	0.0377
moved1	-0.2479	0.4685	-0.1346	0.0374
partners	0.2916	0.2463	-0.0283	0.0200
managexp	-0.1574	0.4609	-0.0583	0.0357
firequit	0.1436	0.4926	-0.0615	0.0397
diffprod	-1.1940	0.3871	-0.0175	0.0314
parenown	0.2724	0.3652	-0.0512	0.0293
educlev	-0.1871	0.1066	-0.0029	0.0087
opercont	0.2733	0.1618	0.0096	0.0131
capinv	0.1917	0.1319	0.1325	0.0099
oddsyr	0.3782	0.0889	0.0014	0.0073
agef	0.0238	0.0270	0.0013	0.0022
lgl_form1.2	-0.3936	0.6005	-0.0541	0.0494
lgl_form1.3	-0.1912	0.4447	-0.1993	0.0355
loans	0.6295	1.0630		
homestead			-0.0025	0.0012
h_unlimited			0.2834	0.0977
HH			-0.0001	0.0000
debt_2_assets			0.0147	0.1526
start1			-0.0484	0.0317
covariance	0.0490	0.2907		

Notes: The sample size is 2880. The dependent variable is equation 1 is *changemp* and the dependent variable in the second equation is *loans*. The last line is the posterior mean and standard error of covariance parameter σ_{12} .

Table B.8 Results

variable	posterior mean, equation1	posterior st.dev. equation1	posterior mean, equation2	posterior st.dev. equation2
intercept	-1.2901	1.3507	-1.6299	0.3198
age1	-0.0011	0.0195	-0.0009	0.0040
totjobs	-0.1038	0.0432	0.0059	0.0089
bs11	0.0040	0.0108	-0.0054	0.0022
bs21	0.0053	0.0087	0.0000	0.0018
bs61	-0.0229	0.0153	0.0017	0.0029
race	-2.0527	0.7346	0.3721	0.1330
outjob4	0.8599	0.4839	-0.0733	0.0974
unemp_change	-0.2735	0.1385		
est_change	0.9565	1.5779		
sex	-0.6211	0.4600	-0.1844	0.1012
moved1	-0.3381	0.4424	0.0969	0.0856
partners	-0.2905	0.2415	0.2351	0.0400
managexp	-0.1658	0.4502	0.0494	0.0935
firequit	0.2015	0.4806	-0.1522	0.1070
diffprod	-0.7545	0.3884	-0.1659	0.0797
parenown	0.4505	0.3576	-0.1082	0.0720
educlev	-0.1810	0.1047	0.0268	0.0212
opercont	0.3470	0.1597	-0.0659	0.0327
capinv	0.0854	0.1136	0.0933	0.0225
oddsyr	0.3768	0.0868	-0.0168	0.0176
agef	0.0396	0.0265	-0.0063	0.0053
lgl_form1.2	-0.7237	0.5878	0.3157	0.1114
lgl_form1.3	-0.7599	0.4308	0.3681	0.0858
invest	13.3058	1.0685		
homestead			0.0014	0.0026
h_unlimited			-0.0649	0.2083
HH			-0.0001	0.0001
debt_2_assets			-0.0729	0.3282
start1			0.1411	0.0703
covariance	-6.7947	0.5946		

Notes: The sample size is 2880. The dependent variable in equation 1 is *changemp* and the dependent variable in the second equation is *invest*. The last line is the posterior mean and standard error of covariance parameter σ_{12} .

Table B.9 Results

variable	posterior mean, equation1	posterior st.dev. equation1	posterior mean, equation2	posterior st.dev. equation2
intercept	-0.8730	1.4125	0.1632	0.1208
age1	0.0003	0.0198	-0.0055	0.0015
totjobs	-0.1001	0.0450	-0.0035	0.0034
bs11	-0.0039	0.0112	-0.0008	0.0008
bs21	0.0035	0.0088	-0.0003	0.0006
bs61	-0.0257	0.0155	0.0004	0.0011
race	-1.9387	0.7591	-0.0217	0.0569
outjob4	0.9316	0.5011	-0.0259	0.0363
unemp_change	-0.2415	0.1560		
est_change	1.0717	1.7403		
sex	-0.8692	0.4731	0.0992	0.0341
moved1	-0.2669	0.4584	-0.1130	0.0336
partners	0.2733	0.2476	0.0470	0.0174
managexp	-0.1663	0.4516	-0.0472	0.0329
firequit	0.1325	0.4936	-0.0722	0.0363
diffprod	-1.1833	0.3899	-0.0445	0.0282
parenown	0.2680	0.3640	-0.0762	0.0264
educlev	-0.1856	0.1073	-0.0019	0.0079
opercont	0.2775	0.1612	-0.0060	0.0120
capinv	0.2028	0.1377	0.1284	0.0089
oddsyr	0.3801	0.0895	-0.0021	0.0066
agef	0.0241	0.0273	-0.0004	0.0020
lgl_form1.2	-0.4030	0.5959	0.0065	0.0441
lgl_form1.3	-0.2381	0.4407	-0.1130	0.0324
outside	0.3892	1.1236		
homestead			-0.0023	0.0011
h_unlimited			0.2782	0.0889
HH			0.0000	0.0000
debt_2_assets			0.0470	0.1369
start1			-0.0226	0.0285
covariance	0.0303	0.2859		

Notes: The sample size is 2820. The dependent variable in equation 1 is *changemp* and the dependent variable in the second equation is *outside*. The last line is the posterior mean and standard error of covariance parameter σ_{12} .

APPENDIX C. First order conditions for chapter 4

Expected utility

The expression of expected utility is derived in the following manner:

$$\begin{aligned}
EU_i &= (\tilde{U}_i - c)\pi_i + (\tilde{U}_i - 2c)\pi_i(1 - \pi_i) + (\tilde{U}_i - 3c)\pi_i(1 - \pi_i)^2 + \dots \\
&= [\tilde{U}_i\pi_i + \tilde{U}_i\pi_i(1 - \pi_i) + \tilde{U}_i\pi_i(1 - \pi_i)^2 + \dots] - \\
&\quad - [c\pi_i + c\pi_i(1 - \pi_i) + c\pi_i(1 - \pi_i)^2 + \dots] - \\
&\quad - [c\pi_i(1 - \pi_i) + c\pi_i(1 - \pi_i)^2 + c\pi_i(1 - \pi_i)^3 + \dots] - \\
&\quad - [c\pi_i(1 - \pi_i)^2 + c\pi_i(1 - \pi_i)^3 + c\pi_i(1 - \pi_i)^4 \dots] - \dots \\
&= \tilde{U}_i\pi_i/(1 - (1 - \pi_i)) - c\pi_i/(1 - (1 - \pi_i)) - c\pi_i(1 - \pi_i)/(1 - (1 - \pi_i)) - \\
&\quad - c\pi_i(1 - \pi_i)^2/(1 - (1 - \pi_i)) - \dots \\
&= \tilde{U}_i - c - c(1 - \pi_i) - c(1 - \pi_i)^2 - \dots \\
&= \tilde{U}_i - c/(1 - (1 - \pi_i)) = \tilde{U}_i - c/\pi_i
\end{aligned}$$

First order condition

$$\begin{aligned}
\frac{\partial EU_i}{\partial \bar{U}_i} &= \frac{(\int_{\bar{U}_i}^{\infty} U_i \varphi(U_i | \cdot) dU_i - c) / \int_{\bar{U}_i}^{\infty} \varphi(U_i | \cdot) dU_i}{\partial \bar{U}_i} = \\
&= \frac{-\bar{U}_i \varphi(\bar{U}_i | \cdot) \int_{\bar{U}_i}^{\infty} \varphi(U_i | \cdot) dU_i - (-\varphi(\bar{U}_i | \cdot)) (\int_{\bar{U}_i}^{\infty} U_i \varphi(U_i | \cdot) dU_i - c)}{(\int_{\bar{U}_i}^{\infty} \varphi(U_i | \cdot) dU_i)^2} = 0 \\
&\Leftrightarrow \bar{U}_i \varphi(\bar{U}_i | \cdot) \int_{\bar{U}_i}^{\infty} \varphi(U_i | \cdot) dU_i = \varphi(\bar{U}_i | \cdot) (\int_{\bar{U}_i}^{\infty} U_i \varphi(U_i | \cdot) dU_i - c) \\
&\Leftrightarrow \bar{U}_i \int_{\bar{U}_i}^{\infty} \varphi(U_i | \cdot) dU_i = (\int_{\bar{U}_i}^{\infty} U_i \varphi(U_i | \cdot) dU_i - c) \\
&\Leftrightarrow \int_{\bar{U}_i}^{\infty} \bar{U}_i \varphi(U_i | \cdot) dU_i - \int_{\bar{U}_i}^{\infty} U_i \varphi(U_i | \cdot) dU_i = -c \\
&\Leftrightarrow \int_{\bar{U}_i}^{\infty} (U_i - \bar{U}_i) \varphi(U_i | \cdot) dU_i = c
\end{aligned}$$

The second line above was obtained by using the derivative of the ratio rule and Leibnitz rule.

BIBLIOGRAPHY

- Abbring, J. H., P. A. Chiappori, J. J. Heckman, and J. Piquet. (2003). Adverse Selection and Moral Hazard in Insurance: Can Dynamic Data Help to Distinguish? *Journal of the European Economic Association, Papers and Proceedings*, 1, 512–521.
- Albert, J. and S. Chib (1993). Bayesian Analysis of Binary and Polychotomous Response Data. *Journal of the American Statistical Association*, 88, 669–679.
- Araujo, A. and H. Moreira (2001). Non-Monotone Insurance Contracts and their Empirical Consequences. *Working paper, Department of Economics, PUC-Rio, Brazil*.
- Bajaj, M., Y.-S. Chan and S. Dasgupta (1998). The relationship between ownership, financing decisions and firm performance: A signalling model. *International Economic Review*, 39(3), 723–744.
- Berger, A. N. and E. B. di Patti (2002). Capital Structure and Firm Performance: A New Approach to Testing Agency Theory and an Application to the Banking Industry. *Board of Governors of the Federal Reserve System, Finance and Economics Discussion Series*, 2002-54.
- Berger, A. N. and G. F. Udell (2003). Small Business and Debt Finance. In Z. J. Acs and D. B. Audretsch (eds.) *Handbook of Entrepreneurship Research: An Interdisciplinary Survey and Introduction*. Kluwer Academic Publishers.
- Berkowitz J. and M. J. White (2004). Bankruptcy and Small Firm's Access to Credit. *RAND Journal of Economics*, 35(1), 69–84.

- Bitler, M., Moskowitz, T., and A. Vissing-Jorgensen (2005). Testing Agency Theory With Entrepreneur Effort and Wealth. *Journal of Finance*, 60(2), 539–576.
- Bloomberg, L. J., and M. Cancian (2004). Do Workers have a Choice and Sort Strategically for Employer Sponsored Insurance Offers? *Working paper, ERIU*.
- Bradley, R. (2002). Is Adverse Selection Simply Moral Hazard? Evidence from the 1987 Medical Expenditure Survey. *Working paper 360, US Bureau of Labor Statistics*.
- Börsch-Supan, A., McFadden, D. L., and R. Schnabel (1996). Living Arrangements: Health and Wealth Effects. In Wise, D. A. (Ed.) *Advances in the Economics of Aging*. The University of Chicago Press.
- Cardon, J., and I. Hendel (2001). Asymmetric Information in Health Insurance: Evidence from the National Medical Expenditure Survey. *RAND Journal of Economics*, 32(3), 408–423.
- Cassar, G. (2004). The Financing of Business Start-ups. *Journal of Business Venturing*, 19, 261–283.
- Cetorelli, N. and P. E. Strahan (2004). Finance as a Barrier to Entry: Bank Competition and Industry Structure in Local U.S. Markets. *NBER working paper 10832*.
- Chiappori, P., Jullien, B., Salaniè, B., and F. Salaniè (2002). Asymmetric Information in Insurance: Some Testable Implications. *Working paper, The University of Chicago*.
- Chiappori, P. A. and B. Salaniè (2000). Testing for Asymmetric Information in Insurance Markets. *Journal of Political Economy*, 108, 56–78.
- Cooper A. C., Dunkelberg W. C., Woo, C. Y., and W. J. Dennis (1990). *New Business in America: The Firms & Their Owners*. Washington, DC: The NFIB Foundation.
- Coulson, N. E. and B. C. Stuart (1995). Insurance Choice and Demand for Prescription Drugs. *Southern Economic Journal*, 61(4), 1146–1157.

- Cressy, R. (1996). Are Business Startups Debt Rationed? *Economic Journal*, 106(438), 1253–1270.
- Darrough M. N. and N. M. Stoughton (1986). Moral Hazard and Adverse Selection: The Question of Financial Structure. *Journal of Finance*, 41(2), 501–513.
- Deb, P. and P. Trivedi (2002). The Structure of Demand for Health Care: Latent Class versus Two-part Models. *Journal of Health Economics*, 21, 601–625.
- Deb, P. and P. Trivedi (2003). Gatekeeping, Self-Selection, and Utilization of Curative and Preventive Health Care Services. *Working paper, Department of Economics, CUNY*.
- Deb, P. and P. Trivedi (2004). Specification and Simulated Likelihood Estimation of a Non-normal Outcome Model with Selection: Application to Health Care Utilization. *Working paper, Department of Economics, CUNY*.
- Dessi, R. and D. Robertson (2003). Debt, Incentives and Performance: Evidence from UK Panel Data. *Economic Journal*, 113, 903–919.
- Dionne, G., Gouriéroux, C., and C. Vanasse (1998). Evidence of Adverse Selection in Automobile Insurance Markets. *Working paper 98-09, HEC Montreal*.
- Dionne, G., Gouriéroux, C., and C. Vanasse (2003). The Informational Content of Household Decisions with Applications to Insurance under Adverse Selection. *Working paper 9701, Institut National de la Statistique et des Etudes Economiques, France*.
- Dionne, G., P.- C. Michaud, and M. Dahchor (2004). Separating Moral Hazard from Adverse Selection in Automobile Insurance: Longitudinal Evidence from France. *Working paper 2004-79, CentER, Tilburg University*.
- Dybvig, P. H. and Y. Wang (2002). Debt and Equity. *Working paper, Boston University*.
- de Meza, D. and D. Webb (2001). Advantageous Selection in Insurance Markets. *RAND Journal of Economics*, 32(2), 249–262.

- Dowd, B., Feldman, R., Cassou, S., and M. Finch (1991). Health Plan Choice and the Utilization of Health Care Services. *The Review of Economics and Statistics*, 73(1), 85–93.
- Driffield, N., V. Mahambrare, and S. Pal (2005). How Ownership Structure Affects Capital Structure and Firm Performance? Recent Evidence from East Asia. *Working paper, Brunel University*.
- Edelberg, W. (2004). Testing for Adverse Selection and Moral Hazard in Consumer Loan Markets. *Finance and Economics Discussion Series, Board of Governors of the Federal Reserve System, 2004-09*.
- Ettner, S. (1997). Adverse Selection and the Purchase of Medigap Insurance by the Elderly. *Journal of Health Economics*, 16, 543–562.
- Finkelstein, A. and K. McGarry (2003). Private Information and its Effect on Market Equilibrium: New Evidence from Long-term Care Insurance. *Working paper 9957, NBER*.
- Gardiol, L., Geoffard, P.-Y., and C. Grandchamp (2003). Separating Selection and Incentive Effects: an Econometric Study of Swiss Health Insurance Claims Data. *Working paper*.
- Geweke, J. (1991). Efficient Simulation from the Multivariate Normal and Student-*t* Distributions Subject to Linear Constraints. In Keramidas E. M. (Ed.) *Computing Science and Statistics: Proceedings of the 23rd Symposium on the Interface*. Fairfax Station: Interface Foundation.
- Grossman, S. J. and O. Hart (1982). Corporate Financial Structure and Managerial Incentives. In J. McCall (ed.) *The Economics of Information and Uncertainty*. The University of Chicago Press.
- Harris, M. and A. Raviv (1991). The Theory of Capital Structure. *Journal of Finance*, 46(1), 297–355.
- Hastings, W.K. (1970). Monte Carlo Sampling Methods Using Markov Chains and Their Applications. *Biometrika*, 57, 97–109.

- Holly, A., Gardiol, L., Domenighetti, G., and B. Bisig (1998). An Econometric Model of Health Care Utilization and Health Insurance in Switzerland. *European Economic Review*, 42, 513–522.
- Hurd, M. and K. McGarry (1997). Medical Insurance and the Use of Health Care Services by the Elderly. *Journal of Health Economics*, 16, 129–154.
- Jensen, M. C. and W. Meckling (2005). Theory of the Firm: Managerial Behavior, Agency Costs, and Capital Structure. *Journal of Financial Economics*, 3, 305–360.
- Jullien, B., Salaniè, B., and F. Salaniè (2005). Screening Risk-Averse Agents Under Moral Hazard. *Working paper, Institut National de la Statistique et des Etudes Economiques, France.*
- Koop, G. (2003). *Bayesian Econometrics*. John Wiley & Sons.
- Kosel, J. (1985). *Bankruptcy: do it yourself*. Berkley: Nolo Press.
- Leland, H. and D. Pyle (1977). Information asymmetries, financial structure, and financial intermediaries. *Journal of Finance*, 32(2), 371–387.
- Li, C. and P. Trivedi (2004). Disparities in Medical Care Utilization, Insurance Coverage and Adverse Selection. *Working paper, Indiana University.*
- Maddala, G. S. (1986). *Limited-dependent and Qualitative Variables in Econometrics*. Cambridge University Press.
- McConnell, J. J. and H. Servaes (2005). Equity ownership and the two faces of debt. *Journal of Financial Economics*, 39(1), 131–157.
- McCulloch, R., Polson, N., and P. Rossi (2000). A Bayesian Analysis of the Multinomial Probit Model with Fully Identified Parameters. *Journal of Econometrics*, 99, 173–193.
- Meer, J. and H. Rosen (2004). Insurance and the Utilization of Medical Services. *Social Science and Medicine*, 58(9), 1623–1632.

- Metropolis, N., Rosenbluth, A. W., Rosenbluth, M. N., Teller, A. H., and E. Teller (1953). Equations of State Calculations by Fast Computing Machines. *Journal of Chemical Physics*, 21, 1087–1092.
- Modigliani, F. and M. Miller (1958). The cost of capital, corporation finance, and the theory of investment. *The American Economic Review*, 48(2), 261–297.
- Monheit, A. and J. Vistnes (1999). Health Insurance Availability at the Workplace: How Important are Worker Preferences? *Journal of Human Resources*, 34(4), 770–785.
- Monheit, A., and J. Vistnes (2004). Health Insurance Enrollment Decisions: Understanding the Role of Preferences for Coverage. *Working paper, ERIU*.
- Munkin, M. and P. Trivedi (2003). A Bayesian Analysis of a Self-Selection Model with Multiple Outcomes Using Simulation-Based Estimation: an Application to the Demand for Health-care. *Journal of Econometrics*, 114, 197–220.
- Myers, S. C. (2001). Capital Structure. *Journal of Economic Perspectives*, 15(2), 81–102.
- Myers, S. C. (2003). Financing of Corporations. In G. M. Constantinides, M. Harris and R. M. Stulz (eds.) *Handbook of the Economics of Finance*. Amsterdam: North Holland.
- Myers, S. C. and N. S. Majluf (1984). Corporate Financing and Investment Decisions When Firms Have Information that Investors Do Not Have. *Journal of Financial Economics*, 13(2), 187–221.
- Rosen, S. (1978). Advertising, Information, and Product Differentiation. In D. Tuerck (ed.) *Issues in Advertising: The Economics of Persuasion*. American Enterprise Institute.
- Rothschild, M. and J. Stiglitz (1976). Equilibrium in Competitive Insurance Markets: An Essay on Economics of Imperfect Information. *The Quarterly Journal of Economics*, 90(4), 629–649.

- Thompson, P. (2005). Selection and Firm Survival. Evidence from the Shipbuilding Industry. *Review of Economics and Statistics*, 87(1), 26–36.
- Tierney, L. (1994). Markov Chains for Exploring Posterior Distributions. *Annals of Statistics*, 22, 1701–1762.
- Train, K. (2003). *Discrete Choice Methods with Simulation*. Cambridge University Press.
- Wahrenburg, M. (1996). Financial Contracting with Adverse Selection and Moral Hazard. In Picot, A. and E. Schlicht (eds.) *Perspectives on Contract Theory*. Physika.
- Walker, J. and M. Ben-Akiva (2002). Generalized Random Utility Model. *Mathematical Social Sciences*, 43, 303–343.
- Wheelock D. C. and S. C. Kumbhakar (1995). Which Banks Choose Deposit Insurance? Evidence of Adverse Selection and Moral Hazard in a Voluntary Insurance System. *Journal of Money, Credit and Banking*, 27(1), 186–201.

ACKNOWLEDGEMENTS

I would like to express my sincere gratitude to all my committee members for their invaluable help and advice. I have to specially thank Brent Kreider for giving me independence to pursue my own ideas and for encouraging me throughout the research process. I am very grateful to Justin Tobias for introducing me to Bayesian econometrics and helping me with numerous problems that I encountered in my research. I have also benefited from the advice from Peter Orazem, Joseph Herriges, Oscar Volij, and Philip Dixon. I would also like to thank Steven Hill for his many comments on my research and for introducing me to the data I used in the second chapter.