ZHENG, MAOZHAO, Ph.D., Employer-sponsored Health Insurance and Worker Productivity. (2017)
Directed by Dr. Ken Snowden and Dr. Martin Andersen. 191 pp.

Health insurance for the working population of the United States is largely provided through employers primarily because of favorable income tax treatments and employment laws that encourage employers, large or small, to provide health insurance to their employees. In fact, a recent survey shows that employer-sponsored insurance covers over 50% of the non-elderly population of the United States, 57% of firms offered health benefits to employees, and 63% of workers accepted the coverage. This dissertation addresses several interesting questions raised by this situation. First, why do some firms offer health insurance to their employees whereas others do not? Second, what determines the number of health insurance plans offered among employers who do offer health insurance? Third, how do employers' decisions concerning how many health insurance plans to offer influence the take-up decisions by employees and, therefore, variations in the extent and quality of health insurance coverage across industries and occupations? To provide at least a partial answer to these questions, this dissertation hypothesizes that employer-sponsored health insurance may affect worker productivity and, as a result, the different types of health insurance policies offered by employers and taken up by their workers. It then empirically investigates whether and how it does so.

The first part of the dissertation addresses this issue from a theoretical perspective by investigating how employers decide the types and costs of the health insurance plans they offer, and their workers decisions whether to take up those plans. Following the existing literature, I examine these issues assuming that each worker's demand for health insurance (and health status) is unobserved by the firm, that health insurance plans are priced competitively, and that workers do not move between employers. The contribution of the theory presented here is to add into

this environment the assumption that a worker's productivity is altered by the provision of employer-sponsored health insurance. The dissertation also explores certain variants of this theoretical model to investigate workers' take-up decisions by modifying the framework so that workers can choose to remain uninsured.

The second part of the dissertation tests whether and how employer-sponsored health insurance affects worker productivity in the real world by conducting an empirical analysis using data from the Medical Expenditure Panel Survey (MEPS). It does so by using a variable measuring health-related absenteeism at the workplace due to sickness as a proxy for productivity and investigating the relationship between this productivity proxy variable and a variable indicating whether a worker is insured through his or her employer. To purge this relationship of the endogeneity effects that may result from selection, the dissertation utilizes employment-related spousal variables as instruments for the potentially endogenous employer-sponsored health insurance variable. The resulting (negative) estimates suggest that employer-health insurance may enhance worker productivity by reducing health-related absenteeism.

The hypothesis that health insurance improves worker productivity helps explain why firms are willing to offer health insurance to their employees and bear part of the premium costs. The dissertation makes several contributions in the field of health economics. First, the dissertation brings about the novel idea that health insurance may affect productivity. Second, it theoretically examines the take-up decisions of workers by allowing them to remain uninsured. Third, the dissertation studies the firms' optimal decisions and equilibrium conditions when workers require reservation wages. Fourth, it finds a statistically significant empirical relationship between a proxy for worker productivity (days missed for health reasons) and employer-sponsored health insurance.

EMPLOYER-SPONSORED HEALTH INSURANCE AND WORKER PRODUCTIVITY

by

Maozhao Zheng

A Dissertation Submitted to
the Faculty of The Graduate School at
The University of North Carolina at Greensboro
in Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy

Greensboro
2017

Approved by

Ken Snowden__________________________
Committee Co-Chair

Martin Andersen________________________
Committee Co-Chair

DEDICATION

This dissertation is dedicated to people without whom my life would have been less meaningful; such people include but are not limited to the Zheng's, the Zhang's, Ray Liao, Chuchu Zheng, Xianghui, Yuxing, Huaichu, Feiyang, and Yiwen Zheng.

The dissertation is also dedicated to those without whom I would have lived like a lonely wolf during these years.

APPROVAL PAGE

This dissertation written by Maozhao Zheng has been approved by the following committee of the Faculty of The Graduate School at The University of North Carolina at Greensboro.

Committee Co-Chair ________________________________
Ken Snowden

Committee Co-Chair ________________________________
Martin Andersen

Committee Members ________________________________
Martijn van Hasselt

____
Stephen Holland

____October 23, 2017__________
Date of Acceptance by Committee

____August 25, 2017________
Date of Final Oral Examination

ACKNOWLEDGEMENTS

I would like to express my sincere gratitude to his academic advisors Dr. Martin Andersen and Dr. Ken Snowden, both Co-chairs of the committee, for their patient guidance, enthusiastic encouragement, and sharp yet helpful critiques of this research work from scratch to completion. Their advice has proven indispensable. Moreover, working under their guidance and supervision has not only been fruitful, but enjoyable as well.

Gratitude is also extended to Dr. Martijn van Hasselt and Dr. Stephen Holland, both members of dissertation committee, for their incisive yet valuable comments. In addition, Dr. van Hasselt and Dr. Holland have provided suggestions that have helped reshape and substantiate the dissertation.

I am grateful to Dr. Dora Gicheva for her insightful comments and helpful suggestions on certain important issues the dissertation has been trying to address.

I would like to thank the Department of Economics as well, for providing me with the necessary resources and convenience throughout these years.

TABLE OF CONTENTS

CHAPTER I

INTRODUCTION

Encouraged by favorable tax treatment and employment law, employers provide health insurance for most of the working population in the United States. According to a survey conducted by the Kaiser Family Foundation (Kaiser) and the Health Research & Educational Trust (HRET), employer-sponsored insurance covers over half of the non-elderly population of the United States, or 147 million people. In 2015, about 57% of firms offered health benefits to at least some of their employees, and 63% of workers were covered at those firms. Of the firms that offered health benefits, about 17% ended up offering a single plan type. However, more than 50% of covered workers were in firms that offered more than one plan type. Statistics from this same survey also show that over 80% of the small firms (with fewer than 200 employees) end up with a single type of plan, whereas about 60% of large firms (with at least 200 employees) offer two or more types. In addition, descriptive statistics using data from Medical Expenditure Panel Survey (MEPS) show that firms' (individual) health insurance offer rates and take-up rates differ across industries for a given occupation. For example, 64% of the workers engaged in service occupations within the manufacturing industry were offered health insurance plans, and 88% of those who were offered insurance plans ended up with a plan. In contrast, the corresponding figures for this same occupation group within the transportation and utilities industry are 72% and 86%, respectively. Offer rates and take-up rates also differ across occupational groups within an industry. For example, in the wholesale and retail trade industry, 59% of service workers were offered health insurance and 69% took up the offers; the corresponding figures for the production, transportation, material moving occupation group, in contrast, are 70% and 80%. While such cross-occupation and cross-industry variations in insurance offer rates (and take-up rates) may be attributable to other unobserved factors, variations in productivity across occupations provide potential explanations. These variations do not seem to disappear even after individual characteristics such as age, sex and health status have been controlled for. While such evidence may not lead to confirmative conclusions, it suggests that offer rates and take-up rates for employer-provided health insurance are not fully explained by tax or employment policies.

These stylized facts raise several questions. First, why do some firms offer health insurance to their employees while others do not? Second, what determines the number of health insurance plans offered among employers who do offer health insurance? Third, how do employers' decisions concerning how many health insurance plans to offer influence the take-up decisions by employees and, therefore, variations in the extent and quality of health insurance coverage across industries and occupations? The tax-deductibility of health insurance may be an explanation for why firms offer insurance (Miller 2005), but does not answer the other questions raised above. The dissertation attempts to provide at least a partial answer by investigating the connections between the provision of health insurance and worker productivity.

How can health insurance affect worker productivity? Prior research provides some explanations. Health can be viewed as a form of human capital (Mushkin 1962; Becker 1964; Fuchs 1966; Grossman 1972) and health care is a special form of investment in human capital. With health insurance plans, health care can be made easily accessible when it is needed. Thus, health and therefore human capital would be affected. When a worker internalizes the benefits from an insurance plan, he either recovers or enhances his health, and thus his productivity (Grossman 1972). David Bloom and David Canning (2003) argue (and present empirical evidence to show) that health as human capital affects worker productivity through both direct and indirect mechanisms. Schulte and Vainio (2010) suggest that improving health enhances workforce well-being, thus increasing productivity. In addition, as O'Brien (2003) argues, good health insurance plans are associated with good jobs, and having a good job enhances a worker's well-being, morale and productivity. Furthermore, improving worker rights (in the general sense) enhances productivity (Buchele and Christiansen (1995)). The implication is that offering a health insurance plan to a worker can mean one of the rights of the worker is respected, and thus his productivity may be improved. Aizawa and Fang (2013) imply that over time, health insurance can generate a positive effect on worker productivity.

If health insurance does alter worker productivity, then how would this influence firms' behavior? The theory part of this dissertation intends to answer this question. To this end, this dissertation assumes that health insurance alters worker productivity but does not presume whether the impact is positive or negative. So it allows for the possibility that health insurance reduces the productive power of a worker. The dissertation presents several models to investigate how the firm's decisions and thus equilibrium outcomes are affected when health insurance has an impact on productivity. The first two models are largely built within the framework of Miller (2005), where the firm offers two health insurance plans (one basic and the other more generous) to its workers who are then induced to choose either of the two plans. One such model assumes that workers would become more (less) productive if they are enrolled in the more generous plan than they would if they are enrolled in the basic one, but their productive power is the same when they are covered under any given plan. The other assumes that productivity changes with the proportion of workers enrolled in any given plan. The analysis shows that it is not the productivity level that changes the equilibrium pattern of employer-sponsored coverage, but the difference in the productivity between workers who choose the more generous plan and those who choose the basic plan. I find, for example, that the firm tends to pool its workers into the more generous plan (high-end pooling equilibrium) if the productivity differential is positive and large. If the productivity differential is small or even negative, on the other hand, the firm is likely to pool its workers into the basic plan (low-end pooling equilibrium). Otherwise, the firm would separate its workers into the two plans, leading to a separating equilibrium.

As mentioned in the preceding paragraph, the framework where workers are all insured may be overly-restrictive and unrealistic. To allow for the possibility that some workers may be uninsured, this dissertation removes the restriction that all workers are insured and assumes that workers may or may not (be induced to) take up the insurance policy offered by their employers or unions. This leads to the so-called take-up models. For simplicity, the model in this dissertation assumes that only one plan is offered. Findings similar to those mentioned in the preceding paragraph are obtained regarding the impact of productivity on firms' decisions and therefore the equilibrium outcomes. However, there is a difference now: a pooling equilibrium means either all workers are insured or all are uninsured. Correspondingly, a separating equilibrium means some workers are insured but others are not.

The descriptive evidence presented in the first two paragraphs of this chapter suggests that the productivity-enhancing effect of health insurance may play a role in the firm's decision to offer insurance. The second part of the dissertation examines whether and how health insurance affects worker productivity in the real world, by estimating a model in which absenteeism at the workplace (a proxy for (negative) productivity) is determined in part by whether an individual was insured through his or her employer or union. If employer-provided health insurance improves worker's health, as O'Brien (2003) hypothesizes, then absenteeism would decrease and, as a consequence, productivity may rise. The ordinary least square regression result for the explanatory variable is positive. However, because of selection issues on the labor market, the through-firm insurance status variable may be endogenous, thus biasing the OLS estimator. After the through-firm insurance status variable is instrumented using relevant employment-related spousal variables, the estimates become negative. This result is consistent with the claim that health insurance affects productivity positively, as health insurance reduces absenteeism. Moreover, the empirical study finds no gender effects: the effects of health insurance on married men (husbands) do not significantly differ from its effects on married women (wives) in terms of the number of days they would miss work due to sickness or for other reasons.

The contributions of the dissertation are four-fold. First, it introduces worker productivity into the models, and as a consequence, the models yield meaningful implications that differ from those of previous work. Prior studies on related subjects, including Miller (2005), do not adequately address the issue of whether and/or how health insurance alters worker productivity. Second, it analyzes the situations where workers are allowed to remain uninsured so that the take-up decisions of workers can be examined theoretically. Third, the dissertation goes one step further to study the firms' optimal decisions and equilibrium conditions when workers face credit and liquidity constraints and thus require reservation wages. These modifications result in more empirically relevant and testable predictions than previous research. Fourth, it empirically analyzes the relationship between worker productivity proxy and employer-sponsored health insurance and finds meaningful results that no previous work has obtained.

The body of the dissertation consists of two parts, a theory part (Part II) and an empirics part (Part III). The theory part has three chapters: Chapter 2, Chapter 3, and Chapter 4. Chapter 2 reviews previous theoretical work on employer-sponsored health insurance, and discusses in some detail the model by Miller (2003) who sets up the general theoretical framework that is used in Chapters 3 and 4. At the end of the chapter, some limitations of Miller's model are discussed and the modifications and extensions to be made in Chapters 3 and 4 briefly described. Chapter 3 presents the variable-productivity models. Specifically, the study presented in the Chapter 3 relaxes the (implicit) restriction that health insurance plans do not alter worker productivity and explores how the productivity effect of health insurance may influence the firm's decisions and thus the equilibrium outcomes. The impact on productivity is modeled in two ways: 1) each worker's productivity changes depending on the particular plan they enroll in, and 2) the productivity of workers depends on the proportions of workers enrolled in a particular plan. Then a third model is examined by modifying Miller's important assumption that all workers are insured in either a "moderate" or a "generous" plan. Chapter 4 presents an alternative one-plan model where the health insurance affects worker productivity and workers require a reservation wage.

The empirical part consists of three chapters. Chapter 5 presents the data used for the empirics. First, I describe the process for selecting the sample and how I construct the main variables of interest. Then I present and discuss high level summary statistics. Finally, I consider several potential instruments for health insurance status and present preliminary evidence of the strength of each of the instruments. In Chapter 6, I validate the instruments by estimating a wage regression equation where the dependent variable is the logarithm of hourly wage. The key explanatory variable of interest is a binary indicator for having health insurance from one's own employer or union. Following Olson (2002), health insurance should reduce one's wage, but because of endogeneity concerns, I instrument for health insurance status using the instruments described in chapter 5. Chapter 7 presents the empirical analysis of the models regressing the variable measuring absenteeism on the binary variable indicating whether or not the worker was insured through the firm. In this chapter, I test whether health insurance has an effect on productivity, which I proxy with absenteeism (days missed work due to illness). I use that same setup as in chapter 6, but now estimate models where the dependent variable is either the number of days a person missed work due to one's own sickness or the number of days a person missed work for other reasons; the latter of the two models is in the spirit of a falsification test.

The last chapter, Chapter 8, restates some of the general framework of Miler's environment, its limitations, and the modifications made in the dissertation. It then summarizes the major conclusions and findings. The appendixes list proofs, alternative models, tables, and figures.

CHAPTER II

REVIEW OF PREVIOUS WORK

2.1. Introduction

This chapter begins the discussion of how employers decide to sponsor health insurance coverage for their employees. The theoretical analysis of this issue is conducted within the general environment examined by Miller (2005). In this chapter, I lay out the basics of the Miller model and review his important results. I then extend the analysis to cases where the health coverage sponsored by the employer improves the worker's productivity (Chapter 3) and to cases where an employer might decide not to sponsor health insurance for any of its workers (Chapter 4).

In Miller's model, an employer offers two health plans to its workers—a basic plan (referred to as the moderate plan) and a higher cost plan that provides employees with preferred or more complete coverage (the generous plan). Workers value the two plans differently because of differences in their own health status and expected health care costs. Employees who choose the generous plan contribute to its higher cost by accepting a lower wage than the wage they would receive with the moderate plan. Because workers value the two insurance policies differently, the employer can determine the proportion and the type of workers who choose each of the plans by varying the wage offered with the generous plan. Moreover, because payments of health insurance premiums by employers are not treated as taxable income for the employees, the employer can change the total compensation (wages plus employer insurance contributions) it pays to its employees by altering the wage offered with the generous plan. Within this framework, Miller's model shows how the pattern of health insurance coverage among workers can be explained by the decisions employers make along these two margins to minimize costs (or maximize profits).

The benchmark model would be the one under full information, the case where the firm knows the expected health insurance cost or the riskiness of its workers and can discriminate among its workers when offering compensation packages. However, the model is analyzed under asymmetric information only, because the full information case is straightforward and does not require elaboration. When no discrimination is allowed, full information is equivalent to asymmetric information because the firm cannot utilize the available information. Thus, there is no difference between asymmetric information and the full information case where discrimination is not allowed.

Miller's environment is highly stylized, but captures key elements of a large segment of the labor market. The second section of the chapter lays out the structure and assumptions that define Miller's framework. I then summarize and explain the equilibrium and Miller's key results within this environment. In Section 2.3, I discuss the conditions for all potential equilibria and the corresponding distribution of rents, and point out some of its key limitations. The last section lays out extensions of the Miller framework.

2.2. Miller's Framework

Miller's model involves three decision-makers: the insurance company[1], the employer, and the workers. The insurance company does not directly sell insurance policies to workers, but rather contracts with the firm which then offers the two policies to its workers. The firm does so

[1] Equivalently, the employer can be thought of as a risk-neutral, self-insured firm that sells insurance to its workers, and so the insurance company is not needed.

by offering packages comprised of a wage and one of the two health insurance plans. Workers then make a choice between the two compensation packages.

Miller's assumptions about the labor market simplify the analysis considerably. To begin with, workers are assumed to be identical except for differences in expected health care costs. This means, in particular, that workers are equally productive and that they have the same preferences over wage income and the two health plans. The firm, meanwhile, is assumed to employ a fixed number of workers and that this labor pool is described by a fixed distribution of health care costs. One interpretation of these assumptions is that workers have no mobility, but two further assumptions ensure that the firm does not act as a monopsonist in the labor market. First, the firm must employ and provide health insurance to all its employees, meaning that it must offer at least one wage-insurance plan package that is acceptable to its workers with the lowest expected health insurance costs. Second, the firm can offer only two wage-policy offers—one for the moderate and one for the generous plan. The firm cannot, therefore, use multiple wage-plan offers to learn about, or take advantage of, the expected health care costs of individual workers. This implies that the firm operates under **asymmetric information** in that it knows the distribution of the expected health insurance costs across its workers but does not know the expected health cost of any given individual worker. This informational assumption is natural in the context of employer-sponsored health insurance because generally employers may not be legally allowed to discriminate among its workers based on their health statuses.

Miller makes three assumptions about the insurance market. First, the insurance company is assumed to be risk-neutral. Second, the insurance company is assumed to know the distribution of the risk types of the workers and the true expected health care cost of each type of workers, though it may not know the risk type of any individual worker. Each worker knows his own type. Third, the health insurance market is assumed to be in perfect competition[2]. Perfect competition, coupled with the risk neutrality assumption about insurance companies, implies that every insurance company makes a zero expected profit on each insurance policy that it sells; this means that every insurance policy an insurance company sells on the market has to be actuarially fair. Hence, when the insurance market is in equilibrium, no insurance company can make a positive profit on any particular insurance policy because, if it did, another insurance company would target workers by offering an otherwise identical policy but at a slightly lower price; this process will continue until no positive expected profit can be made.

An equilibrium in Miller's model is said to have been reached if the insurance market is in equilibrium and the firm maximizes its profit (or minimizes its cost), given all its workers accept one of the compensation packages offered by the firm. While the theory involves three decision-makers, its primary focus will be placed on the interaction between the firm and its workers and the equilibrium outcomes. Given the assumptions, there would be three possible equilibrium outcomes: 1) a separating equilibrium, where some workers are enrolled in the moderate plan and all others in the generous plan; 2) a (low-end) pooling equilibrium, where all workers are enrolled in the moderate plan; and 3) a (high-end) pooling equilibrium, where all workers are enrolled in the generous plan. However, Miller imposes such additional restrictions that only allow a separating equilibrium as defined in 1).

2.2.1. The Worker's Problem

The firm offers two health insurance plans. For each plan, the firm offers a wage, so the total compensation package consists of a health insurance plan and a wage at a level appropriate

[2] Alternatively, similar results can be derived under Bertrand price competition.

for that plan. Every worker receives one and only one health insurance plan, so no one is left uninsured but no one is enrolled in both plans either. Let m denote the moderate plan and w_m denote the wage paid to a worker if the worker is enrolled in the moderate plan; let g denote the generous plan and $w_g$denote the wage paid to a worker if the worker is enrolled in the generous plan. For the convenience of exposition, call the set (w_m, m) the moderate offer and (w_g, g) the generous offer.

The (indirect) utility a worker derives from the compensation package is the sum of his wage and the utility he derives from the health insurance plan he is enrolled in. The utility that a worker derives from his health insurance plan is simply its dollar-value benefit. Each worker is assumed to receive some plan so no one is left uninsured (and no one is enrolled in both either).

Let $c \in [0, \gamma]$ be a continuum and denote the type or the expected health insurance cost of a worker, where $\gamma > 0$. Hence worker's utility function can be written as follows:

$$u(w_i, \mathrm{c}) = w_i + v_i(c)$$

where $\mathrm{i} = \mathrm{m}, g$. Note that c is the expected cost associated with health status so $v_i(c)$ can be interpreted as (and is!) certainty equivalent utility derived by enrolling in a health insurance plan (hence there is no risk premium).

Let $m(c)$ denote the dollar-valued benefit derived by a worker of type c who is enrolled in the moderate plan, where $m(c)$ is strictly increasing ($m'(c) > 0$ for $c \in (0, \gamma)$). Let $g(c)$ denote the additional benefit a worker receives when he is enrolled in the generous plan, where $g(c)$ is strictly increasing ($g'(c) > 0$ for $c \in (0, \gamma)$) and strictly convex ($g''(c) > 0$ for $c \in (0, \gamma)$), with $g(0) = 0$. Thus, $v_m(c) = m(c)$ and $v_g(c) = m(c) + g(c)$.

Therefore, the total utility a worker derives from the compensation package is $w_m + m(c)$ if he is enrolled in the moderate plan and is $w_g + m(c) + g(c)$ if he is enrolled in the generous plan. All workers are assumed to have (after-tax) reservation utility $\overline{u}$, which is independent of their types. The reservation utility ($\overline{u}$) of the workers and how they evaluate each of the plans, i.e., the functions $m(c)$ and $g(c)$, are known to the firm.

Assume from now on as a convention that if one type of worker is indifferent between the generous plan and the moderate one and if there is at least another type of worker who strictly prefer the generous plan, that type of worker will choose the generous plan, otherwise he chooses the moderate one. Hence a worker with $c \in [0, \gamma]$ prefers the generous plan to the moderate one if and only if $w_g + m(c) + g(c) \geq w_m + m(c)$ and $w_g + m(c) + g(c) \geq \overline{u}$, with $g(0) = 0$ and $g'(c) > 0$. The first inequality implies $g(c) \geq w_m - w_g$ if some workers prefer and enroll in the generous plan. The continuity assumption on $c \in [0, \gamma]$ implies that, if some workers prefer and enroll in the generous plan, then there exists a $c \in [0, \gamma]$, denoted as $\mathrm{c_g}$, such that

$$g(\mathrm{c_g}) = w_m - w_g \tag{2.2.1}$$

$\mathrm{c_g}$ represents the lowest-cost workers who choose the generous plan. Thus, workers with expected health insurance cost $c \in [0, \mathrm{c_g})$ elect the moderate plan and receive a wage w_m, whereas workers with expected health insurance cost $c \in [\mathrm{c_g}, \gamma]$ choose the generous plan and receive a wage w_g.

In principle, the employer can induce its employees to choose which plan by adjusting either of the two wages. However, it is technically convenient to assume that the wage associated with the moderate plan is fixed and only the wage associated with the generous plan requires adjusting[3]. Given this assumption, the employer can optimally set the wage associated with the moderate plan in such a way that the total benefits from the compensation package are at least as great as the reservation utility of the lowest-cost workers, i.e., $w_m + m(0) = \overline{u}$. This ensures that no one would left uninsured as required.

2.2.2. The Firm's Problem

A worker's health insurance costs the firm αc for the moderate plan and c for the generous plan, where $\alpha \in (0,1)$ and c is the worker's expected health cost. As pointed out in Miller (2005), the consensus in the literature is that α is somewhere between 0.8 and 0.9 in the case where the moderate plan is an HMO and the generous plan is a PPO. For the generous plan, the firm bears the entire expected health insurance cost.

The government levies a tax at a rate t $(0 < t < 1)$. As a worker cares only about his net wage and the health insurance plan he receives, and a worker's contribution to health insurance made through his firm can be made tax deductible (Gruber, 2000), each dollar a worker must contribute for the moderate plan reduces his net after-tax wage by $1 - t$ dollars. In order to compensate the worker, a wage of one dollar costs the firm $1/(1 - t)$ dollars.

Perfect competition in the insurance market implies that the health insurance policy is actuarially fair, and thus the health insurance premium equals the expected health care cost. Because information is asymmetric, the firm does not know the expected health care cost of any given worker; however, it knows the distribution of the expected health care costs. Let $\mathrm{F(c)}$ represent this distribution and $\mathrm{f(c)}$ the corresponding density. Given that workers with expected health insurance cost $c \in [0, \mathrm{c_g})$ elect the moderate plan and workers with expected health

[3] Making this assumption does not affect the equilibrium outcomes in any way.

insurance cost $c \in [c_g, \gamma]$ receive the generous plan, the firm's cost of the moderate and generous plans are $\int_0^{c_g} \left(\frac{w_m}{1-t} + \alpha c\right) f(c)dc$ and $\int_{c_g}^{\gamma} \left(\frac{w_g}{1-t} + c\right) f(c)dc$, respectively.

Miller (2005) does not explicitly talk about the firm's revenue, but the way he treats the firm's objective function implies that he assumes the firm's revenue is fixed or the production function exhibits constant returns to scale and the labor supply curve is perfectly elastic, for he states that the objective of the firm is to minimize its expected cost. Given the equivalency between cost minimization and profit maximization when the firm's revenue is assumed to be fixed, there is no difference between two approaches. However, for consistency, profit maximization rather than cost minimization will be used throughout the dissertation. The firm's problem is therefore to maximize its profit given that the workers accept its offers.

Given the assumptions and the treatments described above, the firm's problem can be equivalently written as

$$\max_{0 \le w_g, 0 \le w_m} -\int_0^{c_g} \left(\frac{w_m}{1-t} + \alpha c\right) f(c)dc - \int_{c_g}^{\gamma} \left(\frac{w_g}{1-t} + c\right) f(c)dc \tag{2.2.2}$$

s.t.:

$$w_m + m(0) \ge \overline{u}$$

$$w_g + m(c_g) + g(c_g) \ge w_m + m(c_g)$$

where the objective function is the average (per worker) profit. The first constraint is the participation constraint, and the second is the incentive-compatibility constraint, which is necessary only if at least some workers enroll in the generous plan. Profit maximization requires the participation constraint to be binding, i.e., $w_m + m(0) = \overline{u}$ or $w_m = \overline{u} - m(0)$.

Because $\overline{u}$ and $m(0)$ are both constants, w_m must also be a constant. Non-discriminatory policy requires an equal wage for all workers who receive the same health insurance plan. Hence every worker who receives the moderate health insurance plan receives wage $w_m = \overline{u} - m(0)$. The incentive-compatibility constraint can be simplified as $g(c_g) \geq w_m - w_g$. In the preceding section, it was shown that there exists such a $c_g \in [0, \gamma]$ that $g(c_g) = w_m - w_g$. Because w_m is fixed, choosing w_m and w_g is equivalent to choosing $g(c_g)$, and because by assumption, $g(c_g)$ is a monotonic function of c_g, choosing $g(c_g)$ is equivalent to choosing c_g. Assuming that the solutions are interior, the firm's problem becomes

$$\max_{c_g \in [0,\gamma]} \; -\int_0^{c_g} \left(\frac{w_m}{1-t} + \alpha c\right) f(c)dc - \int_{c_g}^{\gamma} \left(\frac{w_m - g(c_g)}{1-t} + c\right) f(c)dc \tag{2.2.3}$$

Expanding terms and substituting $w_m = \overline{u} - m(0)$ yields the following expression

$$\max_{c_g \in [0,\gamma]} \; -\frac{\overline{u} - m(0)}{1-t} - \alpha \int_0^{c_g} cf(c)dc - \int_{c_g}^{\gamma} cf(c)dc + \frac{g(c_g)}{1-t} \int_{c_g}^{\gamma} f(c)dc \tag{2.2.4}$$

The first term represents the pre-tax wage if all workers elect the moderate plan, the second term is the expected health care cost of enrolling in the moderate plan, the third term is the expected health care cost of workers who choose the generous coverage, and the last term represents the wage savings that would be obtained if workers take the generous plan.

2.2.3. Equilibrium

Differentiating the firm's objective function with respect to c_g yields the first order conditions, which can be written as

$$\left((1-\alpha)\mathrm{c}_{\mathrm{g}}^{*}-\frac{g\left(\mathrm{c}_{\mathrm{g}}^{*}\right)}{1-t}\right)f\left(\mathrm{c}_{\mathrm{g}}^{*}\right)+\frac{g'\left(\mathrm{c}_{\mathrm{g}}^{*}\right)}{1-t}\left(1-F\left(\mathrm{c}_{\mathrm{g}}^{*}\right)\right)\begin{cases}\leq 0, if\ \mathrm{c}_{\mathrm{g}}^{*}=0,\\ =0, if\ 0<\mathrm{c}_{\mathrm{g}}^{*}<\gamma,\\ \geq 0, if\ \mathrm{c}_{\mathrm{g}}^{*}=\gamma\end{cases} \quad (2.3.1)$$

where $g\left(\mathrm{c}_{\mathrm{g}}^{*}\right)=\mathrm{w}_{\mathrm{m}}-\mathrm{w}_{\mathrm{g}}$. Equation (2.3.1) is the equivalent of Equation (7) in Miller (2005). Note again that Miller (2005) defines the firm's problem as cost-minimization whereas I re-define it, equivalently, as profit-maximization so the signs on the left-hand side of (7) in Miller (2005) are the opposite of the signs in (2.3.1) and the inequalities are reversed. The term $-\frac{g\left(\mathrm{c}_{\mathrm{g}}^{*}\right)}{1-t}$ can be interpreted as pre-tax wage savings resulting from the marginal worker enrolling in the generous plan rather than in the moderate plan. $(1-\alpha)\mathrm{c}_{\mathrm{g}}^{*}$ represents the (pre-tax) expected health insurance cost the firm has to pay for the marginal worker enrolled in the generous plan rather than in the moderate one. This is positive because the generous plan requires the firm to pay the full cost rather than a part of it. $f\left(\mathrm{c}_{\mathrm{g}}^{*}\right)$ is the density (proportion) of marginal workers. The last term $\frac{g'\left(\mathrm{c}_{\mathrm{g}}^{*}\right)}{1-t}\left(1-F\left(\mathrm{c}_{\mathrm{g}}^{*}\right)\right)$ in the equation is the reduction in pre-tax wage savings that results from an infinitesimal rise in $\mathrm{c}_{\mathrm{g}}^{*}$ because all workers who were already enrolled in the generous plan would have to be paid a little higher wage. Note that the left-hand of the first order condition represents the pre-tax marginal net profit from enrolling workers with a cost of $\mathrm{c}_{\mathrm{g}}^{*}$ and so a negative cost is a profit.

As Miller (2005) shows, if $\mathrm{c}_{\mathrm{g}}^{*}=0$, the first term equals zero because $g(0)=0$ by assumption; then the first order condition becomes $\frac{g'(0)}{1-t}\left(1-F(0)\right)\leq 0$; but because $F(0)=0$, the first order would not hold unless $\frac{g'(0)}{1-t}\leq 0$, which contradicts the assumption that $\frac{g'(c)}{1-t}$ is positive for all c, including $c=0$. Hence, it must be true that $\mathrm{c}_{\mathrm{g}}^{*}>0$. This says that the profit-maximizing firm would enroll at least some workers in the moderate plan. Hence no high-end pooling equilibrium can exist where all workers are enrolled in the generous plan.

To better understand the first order condition, let's suppose an interior solution so that the above first order condition holds with equality. Then it can be written as

$$\frac{g(\mathrm{c}_{\mathrm{g}}^*)}{1-t} f(\mathrm{c}_{\mathrm{g}}^*) = \frac{g'(\mathrm{c}_{\mathrm{g}}^*)}{1-t}\left(1 - F(\mathrm{c}_{\mathrm{g}}^*)\right) + (1-\alpha)\mathrm{c}_{\mathrm{g}}^* f(\mathrm{c}_{\mathrm{g}}^*)$$

The term on the left-hand side represents the surcharge for the generous plan (relative to the moderate plan). This surcharge is numerically equal to the wage savings. The terms on the right-hand side represent the marginal cost of enrolling marginal workers with expected health cost at the cut-off level.

The first order condition is necessary but may not sufficient for the problem to have a unique solution. The second order condition for a global maximum can be satisfied with the specifications of $\mathrm{g(c)}$ and a variety of distributions about the expected health cost c.

If $\mathrm{c}_{\mathrm{g}}^* = \gamma$, then the second term is zero because $F(\gamma) = 1$. This implies $g(\gamma) \leq (1-t)(1-\alpha)\gamma$. This says that no worker values the generous plan more than its incremental cost. This implies it would not be socially desirable to offer the generous plan to any worker. On the other hand, if $\frac{g(\gamma)}{1-t} > (1-\alpha)\gamma$, then the left-hand side is positive, contradicting the first order condition that it is non-negative. Thus $\mathrm{c}_{\mathrm{g}}^* = \gamma$ cannot be an optimal solution. Hence $\mathrm{c}_{\mathrm{g}}^* < \gamma$. This result states that the firm would enroll at least some workers in the generous plan as long as these workers are willing to pay for it. If this assumption is relaxed, it is possible that $\mathrm{c}_{\mathrm{g}}^* = \gamma$. That is, the profit-maximizing firm may enroll all its workers in the moderate plan. Therefore, a pooling equilibrium where all workers are enrolled in the moderate plan is possible.

When the equilibrium is separating, the zero-cost (lowest-cost) workers receive the moderate plan and earn the reservation utility, but all other workers receive more than their reservation utility (workers enrolled in the generous plan receive greater benefits than those enrolled in the moderate plan). In the pooling equilibrium, the firm extracts no rent as all rents accrue to workers other than the zero-cost (lowest-cost) ones, but the zero-cost workers have no risk and cannot yield rents.

2.2.4. Miller's Discussion of Equilibrium and Efficiency

Before examining the firm's problem under asymmetric information (when workers' types are unobservable to the firm), Miller (2005) first discusses the socially optimal allocation of plans to workers (935-936):

> Before considering the employer's decision, we first characterize the socially optimal allocation of workers to plans. The incremental cost of enrolling a type-c worker in the generous plan is $(1-\alpha)c$, and the incremental benefit is $g(c)$. Hence, the surplus-maximizing allocation of workers to plans is for a type-c worker to elect generous coverage if and only if $g(c) \geq (1-\alpha)$.
> To focus in the interesting case where each plan is efficiently provided to some workers, we assume that:
>
> $$g'(0) < (1 - \alpha) \text{ and } g(\gamma) > (1 - \alpha)\gamma \quad (1)$$
>
> which implies that there exists a unique worker type $c_E \in (0, \gamma)$ such that:
>
> $$g(c_E) = (1 - \alpha)c_E \quad (2)$$
>
> That is, for type c_E the marginal benefit from the generous plan just equal its marginal cost.
> Under the efficient allocation, workers of type $0 \leq c < c_E$ receive the moderate plan and workers of type $c_E \leq c \leq \gamma$ receive the generous plan, where without loss of generality we adopt the convention that workers indifferent between the two plans choose the generous one.

When discussing efficiency, Miller (2005) assumes an interior solution exists. Given that, there still two sources of inefficiency in the equilibrium in the Miller model. The first is the tax distortion, which increases the number of agents getting the generous plan (936-937)

> Under full information, the cost-minimizing employer chooses a type-specific wage for each worker so that the worker earns exactly his reservation utility from employment.... the employer prefers that the type-c worker receive generous coverage whenever
>
> $$g(c) \geq (1-t)(1-\alpha)c \tag{3}$$
>
> Let c_F be the lowest-cost employee for which the employer prefers generous coverage to moderate, i.e., c_F satisfies (3) with equality. Comparing (3) with (2) shows that $c_F < c_E$. That is, with full information the employer gives generous coverage to some employees for whom the incremental benefit is less than its incremental cost. The reason for this is that, due to the tax advantage afforded employer-provided health benefits, the employer's cost of providing generous coverage is less than the true cost. Hence, c_F can also be thought of as the cut-off point for the tax-preferred socially optimal allocation of workers to plans, i.e., treating the employer's tax-preferred cost as the true cost of care.

The introduction of tax leads to the so-called tax-preferred socially optimal cut-off level of worker type c, which is lower than what is socially optimal. The second source of inefficiency in the equilibrium in the Miller model is informational asymmetry. After presenting the model, Miller talks about inefficiency resulting from informational asymmetry.

> In the full information case, the entire benefit of treating the employer's expenditure on health insurance as non-taxable accrues to the employer….If the employer either does not know workers' types or is unable to act upon this knowledge, then its compensation plan will consist of a choice between moderate coverage and a higher wage or generous coverage and a lower wage. Thus, the difference in the wages can be thought of as the surcharge imposed on those who choose the generous plan. The employer's task is then to choose wages for workers electing each health plan (i.e., the surcharge) in order to minimize the expected compensation cost of its workers, subject to the constraints that each worker receives at least his reservation utility and chooses the health plan that maximizes his net benefit from employment.
>
> To gain further insight into the employer's problem, it is useful to rewrite (7) (which is equivalent to Equation (2.3.1) presented earlier) as (for an interior solution (942):

$$g(c_g^*)\left(1+\frac{1}{\varepsilon}\right) = (1 - t)(1 - \alpha)c_g^* \qquad (10)$$

where ε is the elasticity of demand (willingness to pay) for the generous plan, $\varepsilon = -\left[f(c_g^*)/g'(c_g^*)\right] \times \left[g(c_g^*)/\left(1 - F(c_g^*)\right)\right]$. The left-hand side of (10) is the monopolist's marginal revenue.

Comparing (10) with (3) establishes that fewer workers receive generous coverage under private information than under full information, i.e., $c_g^* > c_F$. Extending the monopoly analogy, in the full information case, the employer is a perfectly price discriminating monopolist, reducing the wage of each worker who receives generous coverage by his willingness to pay for it. Because of this, the employer has an incentive to offer generous coverage to all workers who value generous coverage more than its (tax-subsidized) incremental cost. When the monopolist cannot price discriminate, it charges a price above the competitive price. The result is that fewer workers receive generous coverage.

Expression (10) is useful in thinking about how the preferential tax treatment afforded employer-provided health benefits impacts employer policy and through it employee welfare. Even when there is no tax advantage to providing health benefits, the employer still has an incentive to act as a monopolist, which results in the employer enrolling fewer workers in the generous plan than is socially optimal. That is, if $t = 0$, $c_g^* > c_E$. Relative to this benchmark, making employer-provided health benefits tax advantaged decreases the firm's marginal cost of providing generous coverage to more workers, and therefore induces the employer to charge less for the generous plan and provide generous coverage to more workers. Thus, while the employer's monopoly power leads it to charge a high price for generous coverage and enroll too few workers (from a social perspective) in the generous plan, the tax deductibility of employer-provided health insurance reduces this distortion. Indeed, it is straightforward to show that there exists a tax rate that induces socially optimal sorting.

Miller's analysis shows that, relative to the tax-preferred socially optimal level, informational asymmetry reduces the number of workers receiving the generous plan (937); when there is no tax, it reduces the number of workers receiving the generous plan (937) relative to the social optimum.

The discussion of efficiency by Miller is primarily about the allocation of workers to plans. He does not talk too much about efficiency in the distribution of rents (if any) between the firm and its employees. The discussion of efficiency in the next section will be on rents distribution rather on the allocation of workers to plans. However, allocation of workers types will be briefed in later chapters where productivity is introduced. The following discussion is not presented in Miller (2005).

2.3. Discussions

In his analysis, Miller examines only the separating equilibrium, where some workers are enrolled in the generous plan and others in the moderate plan. Miller's model, however, can support pooling as well as separating equilibria. For example, if the pre-tax wage savings for the worker with the highest expected health costs from the generous plan is less than its additional premium, then the firm will induce all its workers to choose the moderate plan. I refer to the resulting equilibrium as low-end pooling. In the asymmetric information case as presented here, all workers receive the same wage which is equal to the reservation utility minus the health benefit of the moderate plan for the lowest-cost workers. In this low-end equilibrium, all workers except for the lowest-cost ones receive rents and the firm extracts no rents. It is the pooling of different types of workers in a single plan that causes rents to accrue to workers.

If the pre-tax wage savings generated by enrolling the highest-cost workers in the generous plan exceeds the corresponding additional premium cost, then at least some type of workers will be enrolled in the generous plan. If enrolling the lowest-cost workers in the generous plan involves pre-tax transfers from the lowest-cost to higher-cost workers, then the firm will leave the lowest-cost workers in the moderate plan and enroll others in the generous one. This outcome constitutes a separating equilibrium. The specific separating equilibrium depends on the parameters of the model. In every separating equilibrium, the firm pays an additional benefit beyond the wage associated with the generous plan to workers with costs higher than the cost of the marginal workers and none receives that additional part. As an illustration, start with the low-end pooling equilibrium and think of moving the highest-cost workers into the generous plan. Then the wage associated with the generous plan has to be raised so that the marginal workers become indifferent between the generous and the moderate plans. But by the non-discriminatory policy (or asymmetric information), not just the marginal workers but also all higher-cost workers must be paid a higher wage. The excess wage paid to the higher-cost workers is something additional relative to the full information case where the firm pays the higher wage only to the marginal workers. The additional part that the firm has to pay under asymmetric information is the cost of information distortion (or the cost of non-discriminatory policy distortion). As we move down the line until (but not including) the point that represents the lowest-cost workers, different separating equilibria with more and more workers enrolled in the generous plan will be obtained, and every such equilibrium involves an extra cost as described above. Thus, in any separating equilibrium, the after-tax benefit of the marginal workers will be greater than the cost of the generous plan. In this case, all except the lowest-cost ones receive rents, but the firm does not.

2.3.1. Graphical Presentation of the Miller Model: A Closer Look at Equilibrium and Efficiency

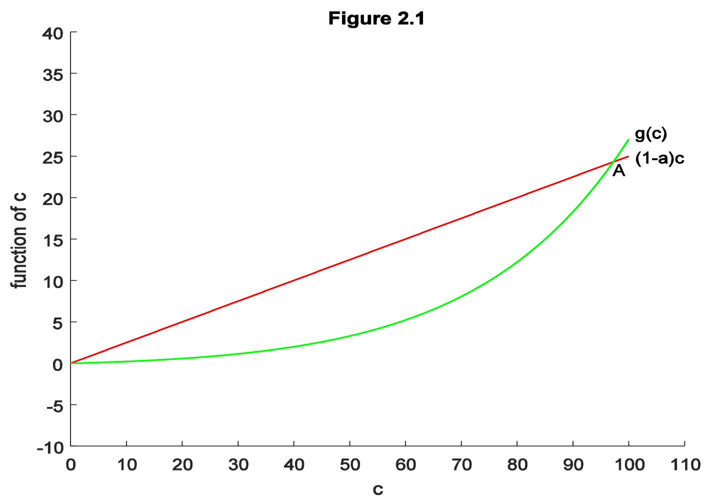
The graphical analysis, including all figures, is not included in Miller (2005), but it is useful because it provides a convenient unifying framework for illustrating comparative statics and different models. Specifically, it is intended to serve two purposes: 1) to illustrate the conditions under which separating versus pooling equilibrium occurs, and 2) as a useful framework to examine the impact of productivity in Chapter 3.

Figure 2.1 shows the full information case without taxation. As shown in the figure, the curve for $(1 - \alpha)c$ is an upward sloping straight line, which captures the premium cost of enrolling workers in the generous plan in excess of the premium cost of enrolling workers in the moderate plan (henceforth the excess premium cost curve). The curve for $g(c)$ captures both the benefit of the generous plan in excess of the benefit of the moderate one (henceforth the excess benefit curve), which is increasing and convex in expected health care cost. Note that the excess benefit curve passes through the origin. A sensible excess benefit function has to pass through the origin, or in other words, the zero-cost workers should not value health insurance positively because they do not face any health risk; nor should such a function have negative values because the generous plan is not worse than the moderate plan for any workers.

The two curves intersect each other at both the origin and point A. But in this case, only point A represents the social optimum because the other point does not satisfy condition (2) in Miller (shown earlier). Point A as illustrated occurs between the origin (where $c = 0$) and the terminal (where $c = 100$). In this case, it is socially optimal to separate workers into two the health plans (separating social optimum). It should be pointed out that **Figure 2.1** only shows one possible socially optimal condition; there are two other possibilities: 1) The excess benefit curve $g(c)$ is above the excess premium cost curve $(1 - \alpha)c$ for all c except at the origin, where the two curves intersect; 2) the excess benefit curve $g(c)$ is below the excess premium cost curve $(1 - \alpha)c$ for all c except at the origin, through which both curves pass. In the case of 1), the social optimum occurs at the origin and so it is socially desirable for the firm to pool all workers in the generous plan. In the case of 2), the social optimum occurs at the terminal and hence it is socially desirable to pool all workers in the moderate plan.

The excess benefit of the generous plan versus its excess premium may affect social optimality. For example, if the excess benefit is lower than the excess premium cost for all workers (except for the lowest-cost ones), then the resulting social optimum would be low-end pooling. If the excess benefit of the generous plan is higher than the corresponding excess premium cost for some workers but lower than the excess premium for all others, the social optimum is separating, where workers with expected health costs higher than the cut-off level would prefer the generous plan and those with expected health care costs below this level prefer the moderate plan. Otherwise, the resulting social optimum would be high-end pooling.

Figure 2.1. Full Information Case without Taxation



The analysis presented above requires the excess benefit function to be convex. Strict convexity is a desirable property, because it is consistent with the usual assumption made in economics textbook that individuals are risk-averse, but it is not sufficient for a desired social optimum to exist, wherein workers with lower costs desire the moderate plan and those with higher costs desire the generous plan. Another property of the excess benefit function is its monotonicity, which is implied.

Figure 2.2 illustrates the impact of taxation on equilibrium and efficiency. The $g(c)/(1-t)$ (where $0 < t < 1$) curve represents the pre-tax excess benefit; it intersects the excess premium cost curve at B, which is to the left of A. This means that taxation causes the firm to cover more workers under the generous plan than socially desired and, as a result, some workers become over-insured relative to the socially optimal allocation. This implies that the firm will become less motivated to enroll workers in the generous plan when the tax rate drops. Therefore, if the tax rate decreases toward zero, then B will moves toward A; when the rate is equal to zero, B will coincide with A. Hence, the social optimal tax rate under full information is zero. If the social optimum occurs at the terminal, the analysis and results will be essentially the same.

If, however, the social optimum happens at the origin, the introduction of taxation will not change the allocation relative to the socially optimum because the introduction of tax will move the excess benefit curve further leftward and away from the excess premium cost curve so the intersection still occurs at the origin.

Figure 2.2. Impact of Taxation on Equilibrium and Efficiency

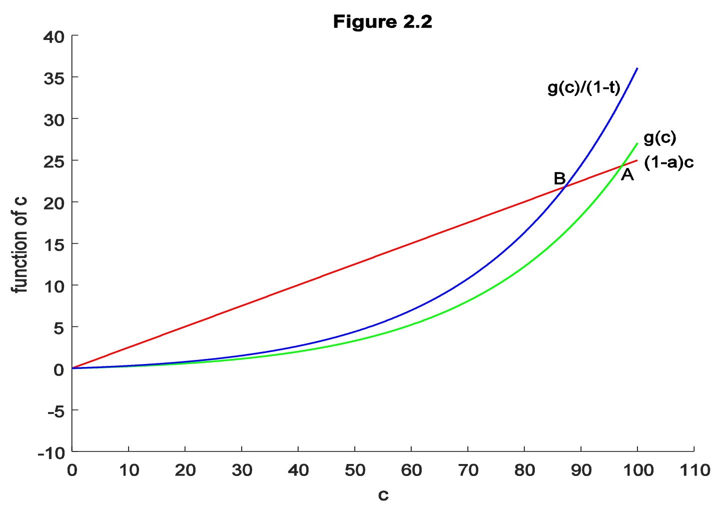
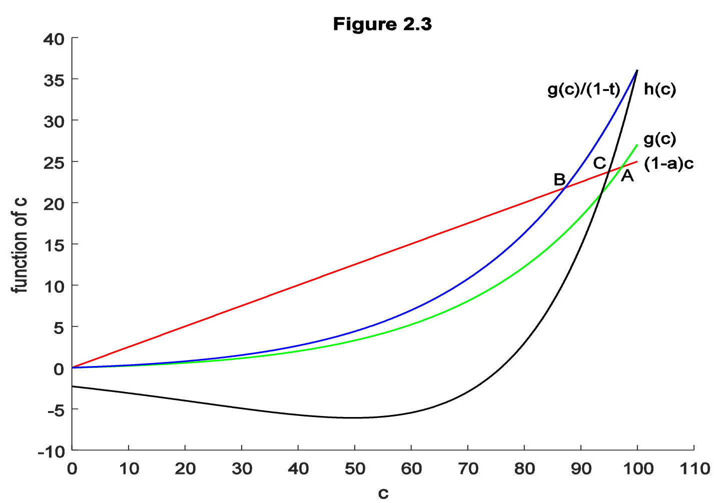


Figure 2.3 illustrates the impact of informational asymmetry on equilibrium. The curve that is shown in Figure 2.3 but not in Figure 2.2 is $h(c)$, which represents $\{\frac{g(c)}{1-t} - \frac{g'(c)}{1-t}\frac{(1-F(c))}{f(c)}\}$, and captures both the benefit of the generous plan in excess of the benefit of the moderate one and the effect of informational asymmetry. The effect of informational asymmetry is negative, so it

moves the curve down and away from the origin[4]. It intersects the excess premium cost curve at point C, which is located between A and B. This means that the firm allocates fewer workers to the generous plan than the tax-preferred socially optimal level but more than what is socially optimal. Because informational asymmetry tilts the pre-tax excess benefit curve leftward from B to C, it reduces the distortion induced by taxation. Thus, under asymmetric information, there exists a positive tax rate that can induce socially optimal sorting. This can be illustrated by gradually reducing the tax rate; as the tax rate decreases, the pre-tax excess benefit curve will tilt right-downward so B and C will move closer and closer to A; C will reach A (before B does) when the tax rate drops to a certain level. The tax rate that makes C arrive at A is the rate that induces socially optimal sorting.

Figure 2.3. Impact of Informational Asymmetry on Equilibrium



[4] If the derivative of the excess benefit function for the generous plan is zero, i.e., $g'(c) = 0$, then $h(c) = \frac{g(c)}{1-t}$; then this curve passes through the origin.

Note that the intersection point C corresponds to the first order condition for an interior solution under asymmetric information, so it represents a separating equilibrium under asymmetric information, where the types to the left of C are enrolled in the moderate plan and those to the right of C are covered under the generous plan.

As illustrated in Figure 2.3, the curve for $h(c)$ does not intersect the curve for $(1 - \alpha)c$ (unless the derivative of the excess benefit function of the generous plan is zero or negative for the lowest-cost workers, but assuming that this derivative is zero or negative seems strange and quite unreasonable), so the pooling equilibrium where all workers are covered under the generous plan (henceforth high-end pooling equilibrium) cannot exist.

The above analysis implies that taxation can change the likelihood of a specific type of equilibrium in addition to causing over-insurance in a particular equilibrium. Specifically, it may increase the likelihood of a high-end pooling but decrease the likelihood of a low-end pooling equilibrium. In particular, if the equilibrium is separating in the case where there is no taxation, the introduction of taxes may turn it into a high-end pooling equilibrium if the impact of the taxation is large enough. However, the introduction of taxes cannot turn a separating equilibrium into a low-end pooling one because taxation encourages the firm to move workers away from the moderate plan and toward the generous one, not the other way round. If the no-taxation equilibrium is pooling at the low end, then introducing taxes may change it into a separating equilibrium; but if the no-taxation equilibrium is pooling at the high end, then the introduction of taxes cannot change it because taxation drives the equilibrium further away from separating.

The analysis of taxes generates several empirical predictions. First, firms become more willing and thus more likely to offer health insurance. Hence, the proportion of firms that offers health insurance to their workers is greater when health insurance premium paid by the employers is tax-deductible than when it is not. In other words, a higher proportion of firms offering health insurance should be observed in cases where the premium cost is tax-deductible. Second, firms would become more willing and so more likely to share the insurance premium costs if health insurance cost is tax-deductible than if it is not. As a consequence, the (individual) health insurance take-up rate or the portion of workers enrolled in employer-provided health insurance plans is likely to be higher when the premium costs are tax-deductible than when they are not, holding everything else the same.

Informational asymmetry, on the other hand, has an opposite effect. It discourages the firm to enroll its employees in the generous plan because the firm has to pay an extra cost (the information rent). In addition, the asymmetry of information may change the likelihood of a pooling equilibrium. For example, it can increase the likelihood of the low-end pooling equilibrium (note that the low-end pooling equilibrium can exist in Miller's framework if one restrictive assumption is relaxed) and decrease the likelihood of a high-end pooling one (it actually makes a high-end pooling equilibrium impossible under any reasonable circumstances).

The empirical prediction generated by the analysis of informational asymmetry is that firms become less willing and less likely to offer health insurance and to share health insurance costs because informational asymmetry causes them to bear an additional burden. As result, the health insurance offer rates and take-up rates are likely to be lower in situations where the issue of informational asymmetry is more serious.

2.4. A Look Ahead to the Extensions of the Miller Model

The Miller framework seems to be a reasonable approximation to the real world situations. For example, there are several well-known influences that restrict worker mobility even if they do not indicate complete immobility. Labor market frictions, more often than not, hinder worker mobility. Looking for a job requires a lot of time and energy. The interviewing process is costly and may be painful. Moving from one place to another to take a job is costly too. The new job may require starting from scratch, leaving a (sunk) cost unrecoverable; in addition, there is a risk that the new job or the new firm may ultimately turn out to be worse. Hence this assumption captures the major characteristics of the real-world job market. The assumption that all workers are insured also reflects to certain extent the reality, since in the United States, over half of the workers are insured. The one-dimensional worker heterogeneity assumption captures and summarizes the primary characteristics of workers and leads to models that yield suggestive equilibrium outcomes.

Despite its advantages, the Miller's framework has a number of limitations. These limitations primarily arise from the restrictive assumptions the models presented in this chapter rest on. First, the worker immobility assumption does not adequately capture the reality, because in the real world, workers are more or less mobile across firms and industries. It is not quite clear what the equilibrium outcomes if this assumption fails to hold without making further assumptions. It can be imagined that the firm's workforce may be replaced in part or in its entirety, depending on how the wage-health benefit package is structured, how good each of the two parts comprising the compensation package is, and what other firms are offering. Also, the assumption that all workers are offered health insurance and accept the offers is not realistic, because empirically many firms end up offering no insurance and many workers may remain uninsured. Furthermore, Miller's model rests upon additional several restrictive assumptions, including an increasing and convex excess benefit function, identical preferences, perfect competition on the health insurance market, fixed distribution of workers' expected health costs, fixed worker productivity, coverage of all workers, and no liquidity constraints for workers. While these assumptions hold in some cases, there are situations they do not.

To overcome some of the limitations of the Miller model, some of these assumptions will be relaxed or otherwise modified in the next two chapters. Additionally, workers may also face liquidity constraints so that they require certain levels of wages. However, all other assumptions for the Miller model will remain unchanged.

In Chapter 3, the Miller model will be extended to include situations where worker productivity alters with the health insurance plans under which workers are covered. The two-plan models presented in this chapter are truly extensions or generalizations of the Miller model in that they incorporate the fixed worker productivity assumption underlying the Miler model as a special case. However, because altering productivity only affects the firm's profit function and has nothing to do with the workers, the worker's problem is not impacted.

Assuming variable worker productivity, as presented in Chapter 3, has two noticeable consequences. First, a pooling equilibrium would be possible even in the case where it is not under the fixed productivity assumption as in Miller. Second, in the case of a separating equilibrium, more (or fewer) workers would be enrolled in the generous health insurance plan if the generous plan boosts (reduces) productivity relative to the moderate one. However, the distribution of rents between the firm and the workers remains the same qualitatively in any equilibrium as in the fixed productivity case, except that the firm may share the increased (decreased) productivity with its workers. Chapter 3 also covers another situation where the firm is assumed to offer only one health insurance plan and workers may choose to either accept or reject the plan and thus remain uninsured.

Chapter 4 explores an alternative model where workers may face credit or liquidity constraints. When workers face such constraints, those who elect to be uninsured (if any) may require a wage that is at least as high as their reservation wage, which is above their reservation utility. If this is the case, then in all possible equilibria, workers receive rents. The firm, however, may or may not extract rents, depending on the type of the equilibrium, as well as on whether the average benefit of the insurance plan the uninsured workers would obtain if they chose to be enrolled in that plan exceeds their reservation wage.

CHAPTER III

MODELS IN WHICH HEALTH INSURANCE AFFECTS WORKER PRODUCTIVITY

3.1. Introduction

The Miller model briefed in Chapter 2 implicitly assumes that the firm's productivity does not vary with the insurance coverage workers elect. The models to be presented in this chapter make a critical change to this assumption by allowing worker productivity to alter with the workers' choices of health insurance plans. Although the study primarily deals with health insurance, the implications of the model results may be applicable to other employer-sponsored wellness programs that potentially affect productivity. Except for this modification, the general framework and all other assumptions used in the Miller model will remain unchanged, including the assumptions of an increasing and convex health insurance excess benefit function, identical workers' preferences, a perfectly competitive health insurance market, and a firm that operates under asymmetric information regarding the expected health care cost of individual workers.

Three models are presented in this chapter. The first two retain Miller's assumptions that the firm offers workers a basic plan (the moderate) and a more generous (the generous) and that these must be structured so that all workers choose to be enrolled in one of them. Within this framework, I assume that each worker's productivity differs depending on which plan is chosen and examine how these productivity differentials across plans alter the likelihood of a pooling versus separating equilibrium and affect the equilibrium cutoff levels of the expected health care cost for a separating equilibrium. The third model examines how the productivity effects of health insurance influence a worker's decision on whether to enroll in insurance or to become uninsured.

To focus on this issue, I assume that the employer offers workers only one insurance policy and explore how the productivity effects of health insurance change the number of workers who accept or reject the conditions under which the employer offers the policy. Although many firms in the United States offer workers more than one plan as Miller assumes, many others offer only one. By examining this environment, it is possible to focus on the important issue of how the productivity impact of health insurance influences which workers remain uninsured.

In all three models, health insurance is assumed to alter worker productivity in one way or the other. In the first, I assume that productivity shifts across plans but remains fixed for any given plan, regardless of how many workers or how large the proportion of workers is enrolled in that plan. This implies that per worker productivity may vary across plans but is equal across all workers who are enrolled in the same plan. In the second framework, I assume that productivity of all workers covered with insurance varies with the proportion of workers covered under the plan; as more workers move from one plan to the other, the two productivity levels diverge. I examine both frameworks because they capture different mechanisms through which health insurance could affect productivity depending on the nature of the tasks or functions workers need to perform at a job.

An important way that jobs differ from one another is the extent to which worker productivity depends on interdependence. For example, tasks that are complementary may have higher degrees of interdependence than others. Accomplishing interdependent tasks requires a high degree of cooperation and/or mutual interaction. On the other hand, cooperation and/or mutual interaction is not as important for tasks that are largely separate and independent. In professions where tasks or functions are interdependent and so cooperation is required, the productivity of one worker also affects the productivity of another. For example, if a worker has to use as input for his task the output generated by an upper-stream worker, then the worker in the downstream cannot finish his task until the upper-stream worker has completed his. However, if interdependence among tasks or functions is weak or nonexistent, then the productivity of one individual worker has little or no impact on the productivity of another.

Thus, if workers are in professions where tasks or functions are interdependent, then when health insurance affects the productivity of one worker, it may also affect the productivity of another. As a consequence, health insurance has two effects on the overall productivity. First, it changes the productivity of each individual worker; second, it indirectly (through some mechanism) changes the productivity of other workers as well. This second effect may arise out of the interaction between workers. The more workers are covered, the more likely or the stronger the health insurance is to affect the productivity of other workers. Hence in such professions (or industries), productivity varies with the proportion of workers enrolled in a health insurance plan. Industries where interdependent tasks or functions may prevail include (but are not limited to) construction, transportation, and manufacturing and occupations of such nature include production, transportation and construction.

If workers are doing independent jobs that do not require much cooperation, the productivity of one worker has little impact on the productivity of another worker and thus the health insurance of one worker does not affect the productivity of another. Therefore, the productivity may shift with health insurance plans as it affects the individual productivity only and does not depends on how many workers are covered under a particular health insurance policy. Industries where there is little or no interdependence among tasks or functions include (but are not limited to) financial service, leisure and hospitality, and professional and business services.

Such occupations may include (but are not limited to) management, sales, and service operation. In these industries or occupations, there are no spillover effects among workers.

The remainder of the chapter is organized as follows. Section 2 presents the model in which productivity alters with the health insurance plans only, covering the worker's problem, the firm's problem, and the equilibrium. Section 3 briefs an alternative model where the productivity varies with the proportion of workers covered under the generous plan. Section 4 introduces and analyzes a one-plan model. Section 5 presents a discussion of the models. Section 6 concludes the chapter.

3.2. A Model where Productivity Changes across Health Insurance Plans

The model presented in this section assumes that worker productivity shifts across the two health insurance plans but is fixed within a given plan. Hence, there are (potentially) only two levels of worker productivity, each a constant for either of the two plans. It is possible that the same health insurance plan may affect different workers, and thus their productivity differently. The model presented in this section, however, focuses only on a single, plan-specific productivity impact. There are two reasons for making such a simple assumption. First, this assumption captures the primary aspects of the impact of health insurance on productivity. In terms of productivity impact, the difference between the types of plans should be more important than the difference across types of workers, given that they are otherwise identical. Second, this assumption simplifies the analysis considerably. I refer to this case as a plan-shifting productivity model and expect it to be most applicable in professions or jobs where only the productivity of individual workers is affected by health insurance coverage and there is no spillover effect among workers. Such professions or jobs may include those that require physical or mental strengths (or both) or that encourage independence and creativity.

3.2.1. The Worker's Problem

Throughout this chapter, it is assumed that productivity differentials have no direct impact on a worker's utility, but on the other hand, the worker is affected by how productivity differences change the wage-insurance plan combinations offered by the employer. The worker then evaluates these packages in the same way as in Chapter 2: the utility a worker obtains from the moderate plan is $w_m + m(c)$ and the utility a worker obtains from the generous plan is $w_g + m(c) + g(c)$. Workers would accept employment at the firm if and only if either $w_m + m(c) > \bar{\mathrm{u}}$ or $w_g + \mathrm{m(c)} + g(c) > \bar{\mathrm{u}}$ for $\mathrm{c} \in [0, \gamma]$, and they would elect the generous one if and only if $w_g + m(c) + g(c) \geq w_m + \mathrm{m(c)}$.

3.2.2. The Firm's Problem

As stated at the beginning of this section, worker productivity varies across the two plans but is equal for all workers covered under the same plan. So, let e_k denote the productivity of a worker who has chosen plan k, where $k \in \{m, g\}$. The expected profit the firm obtains from the worker enrolled in the moderate plan is $\int_0^{c_g} \left(e_m - \frac{w_m}{1-t} - c\right) f(c)dc$ and the expected profit it receives from the worker enrolled in the generous plan is $\int_{c_g}^{\gamma} \left(e_g - \frac{w_g}{1-t} - c\right) f(c)dc$, given the distributional assumption about the expected health care cost of the workers.

The introduction of worker productivity does not change the way individual workers make choices of the health insurance plan relative to the Miller case. Hence, the participation constraint, as well as the incentive-compatibility constraint, remains the same. The firm's problem can therefore be expressed as

$$\max_{0\leq w_g, 0\leq w_m} \int_0^{c_g} \left(e_m - \frac{w_m}{1-t} - \alpha c\right) f(c)dc + \int_{c_g}^{\gamma} \left(e_g - \frac{w_g}{1-t} - c\right) f(c)dc \tag{3.2.1}$$

s.t.:

$$w_m + m(0) \geq \overline{u}$$
$$w_g + m(c_g) + g(c_g) \geq w_m + m(c_g)$$

where, again, the first is the individual rationality or participation constraint and the second the incentive-compatibility constraint. As in the Miller case, profit maximization requires both constraints to be binding, so $w_m = \overline{u} - m(0)$, which is a constant, and $g(c_g) = w_m - w_g$ for $c_g \in [0,\gamma]$. Moreover, choosing $g(c_g)$ is equivalent to choosing c_g. Therefore, the firm's problem can be re-expressed as

$$\max_{c_g\in[0,\gamma]} \int_0^{c_g} \left(e_m - \frac{w_m}{1-t} - \alpha c\right) f(c)dc + \int_{c_g}^{\gamma} \left(e_g - \frac{w_m - g(c_g)}{1-t} - c\right) f(c)dc \tag{3.2.2}$$

The above expression can be further re-written as

$$\max_{c_g\in[0,\gamma]} \int_0^{\gamma} (e_g - c)f(c)dc - \frac{w_m}{1-t} + \int_0^{c_g} \left((1-\alpha)c - (e_g - e_m\right) f(c)dc$$
$$+ \frac{g(c_g)}{1-t} \int_{c_g}^{\gamma} f(c)dc \tag{3.2.3}$$

where the first two terms are constants, which do not affect the firm's optimal decisions. Define $\delta = e_g - e_m$. This is the productivity differential between the two types of plans. It is this differential that separates this model from Miller's.

3.2.3. Equilibrium

Differentiating the profit function with respect to c_g yields the first order conditions as follows.

$$\left(-\delta+(1-\alpha)c_g^*-\frac{g(c_g^*)}{1-t}\right)f(c_g^*)+\frac{g'(c_g^*)}{1-t}\left(1-F(c_g^*)\right)\begin{cases}<0, if\ c_g^*=0,\\=0, if\ 0<c_g^*<\gamma,\\>0, if\ c_g^*=\gamma\end{cases} \quad (3.2.4)$$

where $\delta=e_g-e_m$. The optimal choice of $c_g$is determined implicitly by the first order conditions. Productivity indeed affects the firm's optimal choice as it appears in the first order conditions. This is in a contrast with the Miller (2005) model in which productivity is implicitly assumed to be the same for both plans; in that case, productivity differential δ is zero, which is a special case of this model.

The left-hand side is the net profit. Suppose the first order condition holds with equality, i.e., there is an interior solution. Then the term $(-\delta)$ captures the productivity loss as a result of shifting those individuals who used to be enrolled in the generous plan to the moderate plan. All other terms have the same interpretations as in the Miller case. The first part on the left hand side of the first order conditions, $\left(-\delta+(1-\alpha)c_g-\frac{g(c_g)}{1-t}\right)f(c_g)$, captures the total benefit of shifting marginal workers from the generous plan to the moderate one.

3.2.3.1. Analysis of the First Order Conditions

If $0<c_g^*<\gamma$, then the first order condition holds with equality, and because by assumption, $f(c_g^*)\neq 0$, it can be written as

$$\delta = (1-\alpha)\mathrm{c}_{\mathrm{g}}^{*} - \frac{g(\mathrm{c}_{\mathrm{g}}^{*})}{1-t} + \frac{g'(\mathrm{c}_{\mathrm{g}}^{*})}{1-t}\frac{\left(1-F(\mathrm{c}_{\mathrm{g}}^{*})\right)}{\mathrm{f}(\mathrm{c}_{\mathrm{g}}^{*})} \tag{3.2.5}$$

If $\delta > 0$, and if the firm's problem has a unique interior solution so that the second order conditions for a maximum are satisfied, then $\mathrm{c}_{\mathrm{g}}^{*} < \mathrm{c}_{\mathrm{g}}^{\mathrm{M}*}$, where $\mathrm{c}_{\mathrm{g}}^{\mathrm{M}*}$ is the (unique) interior solution in Miller's model presented in Chapter 2. This states that the firm induces more workers to elect the generous plan than if worker productivity is fixed.

To help interpret (3.2.5), define $h(c) = \frac{\mathrm{f}(\mathrm{c})}{1-F(c)}$. The term $1-F(c)$ is (analogous to) the survival probability. The survival probability measures the proportion of workers who would be enrolled in the generous plan as the cut-off level varies. $\mathrm{f}(\mathrm{c})$ is the event density. Hence $h(c)$ is hazard function or hazard rate and $\frac{\left(1-F(\mathrm{c}_{\mathrm{g}}^{*})\right)}{\mathrm{f}(\mathrm{c}_{\mathrm{g}}^{*})}$ the inverse hazard rate. Monotonicity requires $\frac{d}{d\mathrm{c}_{\mathrm{g}}^{*}}\left(\frac{\left(1-F(\mathrm{c}_{\mathrm{g}}^{*})\right)}{\mathrm{f}(\mathrm{c}_{\mathrm{g}}^{*})}\right) \leq 0$.

If $\mathrm{c}_{\mathrm{g}}^{*} = 0$, then the first order condition can be written as $(-\delta)f(0) + \frac{g'(0)}{1-t}\left(1-F(0)\right) < 0$. Note that $F(0) = 0$. By assumption, $f(0) > 0$. Hence $\frac{g'(0)}{1-t}\frac{1}{f(0)} < \delta$. Because the left-hand side is by assumption positive, this condition holds only if $\delta > 0$. If δ is positive, i.e., the generous plan has a larger effect on worker productivity than does the moderate one, then it is possible (but not necessary) that this condition holds. Thus, this condition is more likely to hold if δ is large, $g'(0)$ is small, $f(0)$ is large, and/or the tax rate t is small.

Thus, it is possible for this model to have a pooling equilibrium where the firm enrolls all workers in the generous plan (high-end pooling equilibrium) if the generous plan has a larger positive effect on worker productivity than does the moderate one. In contrast, there is no high-end pooling in the Miller model under any (reasonable) circumstances! If the left-hand side of the first order condition is less than zero, i.e.,

$$\left(-\delta + (1-\alpha)\mathrm{c}_{\mathrm{g}}^{*} - \frac{g\left(\mathrm{c}_{\mathrm{g}}^{*}\right)}{1-t}\right) f\left(\mathrm{c}_{\mathrm{g}}^{*}\right) + \frac{g'\left(\mathrm{c}_{\mathrm{g}}^{*}\right)}{1-t}\left(1 - F\left(\mathrm{c}_{\mathrm{g}}^{*}\right)\right) < 0$$

then by complementary slackness, $\mathrm{c}_{\mathrm{g}}^{*} = 0$. Hence $\frac{g'(0)}{1-t}\frac{1}{f(0)} < \delta$.

Therefore, this condition is necessary but not sufficient for $\mathrm{c}_{\mathrm{g}}^{*} = 0$. This states that the profit-maximizing firm would enroll all its workers in the generous plan only if $\frac{g'(0)}{1-t}\frac{1}{f(0)} < \delta$. If, on the other hand, $\delta \leq \frac{g'(0)}{1-t}\frac{1}{f(0)}$, i.e., if the generous plan does not have enough impact on worker productivity relative to the moderate one, then the profit-maximizing firm would enroll at least some workers in the moderate plan. Furthermore, if the generous plan has a smaller effect on worker productivity than the moderate one so that the productivity differential is negative, the firm would be more likely in this model than in Miller (2005) to enroll at least some workers in the moderate plan.

If $\mathrm{c}_{\mathrm{g}}^{*} = \gamma$, then the second term would be zero since $F(\gamma) = 1$, and the first order condition becomes $(1-\alpha)\gamma - \frac{g(\gamma)}{1-t} > \delta$. If δ is positive, then this condition is less likely to hold than is the corresponding condition in Miller (2005). Thus, the model under study is less likely to have a pooling equilibrium where the firm would enroll all its workers in the moderate plan (low-end pooling). However, if the productivity differential is negative so the right-hand side of the above condition is negative, then this condition is more likely to hold and thus the firm is more likely to pool its workers in the moderate plan than in the Miller case.

The following proposition summarizes the results.

Proposition 3.1. *A separating equilibrium is characterized by* $\delta = (1-\alpha)c_g^* - \frac{g(c_g^*)}{1-t} + \frac{g'(c_g^*)}{1-t}\frac{\left(1-F(c_g^*)\right)}{f(c_g^*)}$*, whereas a pooling equilibrium is characterized by either* $\delta < (1-\alpha)\gamma - \frac{g(\gamma)}{1-t}$ *or* $\delta > \frac{g'(0)}{1-t}\frac{1}{f(0)}$*. In particular, a low-end pooling equilibrium requires* $\delta < (1-\alpha)\gamma - \frac{g(\gamma)}{1-t}$*, whereas a high-end pooling equilibrium requires* $\delta > \frac{g'(0)}{1-t}\frac{1}{f(0)}$*.*

Proof. These are the first order conditions derived above. The analysis above proves the proposition. Note that the first order conditions are necessary but not sufficient for a maximum (and thus an equilibrium) to exist.

How does a change in the productivity differential affect the optimal cut-off level? The following proposition summarizes a result that answers this question.

Proposition 3.2. *If* $c_g^* \in (0,\gamma)$ *is an interior solution to* (3.2.4)*, then* $\frac{\partial c_g^*}{\partial \delta} < 0$.

Intuitively, this lemma states that if the generous (moderate) plan boosts productivity more than the moderate one, then the firm would enroll more workers in the generous plan.

Proof[5]. If $c_g^* \in (0,\gamma)$ is a solution to (3.2.4), then $\delta = (1-\alpha)c_g^* - \frac{g(c_g^*)}{1-t} + \frac{g'(c_g^*)}{1-t}\frac{\left(1-F(c_g^*)\right)}{f(c_g^*)}$; taking c_g^* as a function of δ and differentiating both sides with respect to δ yields

$$1 = (1-\alpha)\frac{\partial c_g^*}{\partial \delta} - \frac{g'(c_g^*)}{1-t}\frac{\partial c_g^*}{\partial \delta} + \frac{g'(c_g^*)}{1-t}\frac{\left(1-F(c_g^*)\right)}{f(c_g^*)}\frac{\partial c_g^*}{\partial \delta} + \frac{g'(c_g^*)}{1-t}\left(2 - \frac{\left(1-F(c_g^*)\right)f'(c_g^*)}{\left(f(c_g^*)\right)^2}\right)\frac{\partial c_g^*}{\partial \delta}$$

[5] This proof can be simplified by using the implicit function theorem and the fact that the second order condition is negative.

Collecting and shifting terms yields

$$\frac{\partial c_g^*}{\partial \delta} = \left[(1-\alpha) + \frac{g'(c_g^*)}{1-t}\frac{\left(1-F(c_g^*)\right)}{f(c_g^*)} - \frac{g'(c_g^*)}{1-t}\left(2 - \frac{\left(1-F(c_g^*)\right)f'(c_g^*)}{\left(f(c_g^*)\right)^2}\right)\right]^{-1} < 0$$

This expression is negative because the term in the bracket is the second order condition (SOC) for a maximum, and c_g^* is a solution implies this SOC for a maximum is satisfied. Intuitively, this result says that, if the firm finds it optimal to separate its workers into the two health insurance plans, then it would be in the firm's best interest to change the optimal cut-off level in the opposite direction when the productivity differential experiences a small change.

Proposition 3.2 implies $\frac{\partial g(c_g^*)}{\partial \delta} = g'(c_g^*)\frac{\partial c_g^*}{\partial \delta} < 0$, because $g'(c_g^*) > 0$. This says that if the productivity differential somehow experiences a small change, the wage differential (recall that $g(c_g) = w_m - w_g$) would have to change in the opposite direction: a rising (declining) productivity differential would therefore widen (shrink) the wage differential whereas a deceasing productivity differential would shrink the wage differential. Furthermore, Proposition 3.2 also implies $\frac{\partial w_g^*}{\partial \delta} = -g'(c_g^*)\frac{\partial c_g^*}{\partial \delta} > 0$, because $g(c_g^*) = w_m - w_g^*$ so $g'(c_g^*)\frac{\partial c_g^*}{\partial \delta} = -\frac{\partial w_g^*}{\partial \delta}$ as w_m is fixed. This states that if the productivity differential widens (narrows), the firm would increase the wage paid to workers who would be enrolled in the generous plan.

The first order conditions can be re-expressed as

$$\left((1-\alpha)c_g^* - \frac{g(c_g^*)}{1-t}\right)f(c_g^*) + \frac{g'(c_g^*)}{1-t}\left(1-F(c_g^*)\right)\begin{cases} < \delta, if\ c_g^* = 0, \\ = \delta, if\ 0 < c_g^* < \gamma \\ > \delta, if\ c_g^* = \gamma \end{cases}$$

If $\delta > 0$, then $c_g^* < c_g^{M*}$, where c_g^{M*} is the cut-off level in the Miller case; this implies more workers will enroll in the generous plan when workers covered under the generous plan are more productive than they would if they were covered under the moderate plan.

The impact of the (positive) productivity differential on the cutoff level and thus equilibrium is illustrated in the following figures.

Figure 3.1. Equilibrium in the Case where Insurance Does Not Affect Productivity

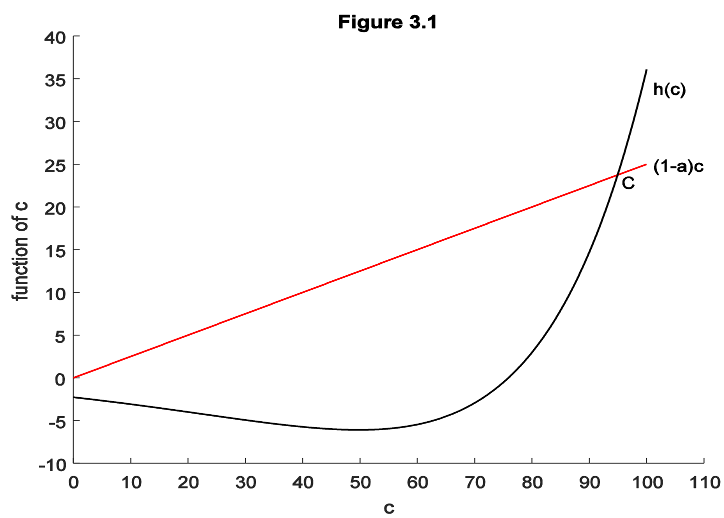
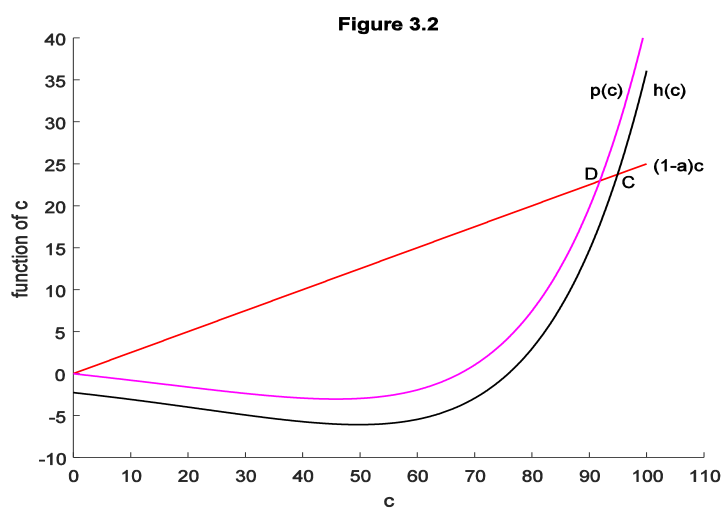


Figure 3.1 shows the case where there is no productivity effect. The $(1-\alpha)c$ line and the $h(c)$ curve are the same as in the Miller case; the former, again, captures the premium cost of enrolling workers in the generous plan in excess of the premium cost of enrolling workers in the moderate plan, and the latter represents $\{\frac{g(c)}{1-t} - \frac{g'(c)}{1-t}\frac{(1-F(c))}{f(c)}\}$, capturing both the benefit of the generous plan in excess of the benefit of the moderate one and the impact of informational asymmetry. As in the Miller case, the asymmetry of information shifts down the excess benefit curve for the generous plan from the origin to a point on the negative part of the vertical axis and so its intercept with the vertical axis is negative. The $(1-\alpha)c$ line and the $h(c)$ curve intersect

each other at C, which is located between the origin and the terminal point. Hence, it represents a separating equilibrium, where workers of types located to the left of C are enrolled in the moderate plan and workers of types located to the right of C are covered under the generous plan.

Figure 3.2 illustrates a case where the productivity differential is positive. The line $(1-\alpha)\mathrm{c}$ and the curve for $h(\mathrm{c})$ are the same as shown in the **Figure 3.1**. The curve $p(\mathrm{c})$ represents $\{\, h(c)+\delta \,\}$ or $\{\frac{g(c)}{1-t} - \frac{g'(\mathrm{c})}{1-t}\frac{(1-F(\mathrm{c}))}{f(\mathrm{c})} + \delta\,\}$, where δ is a positive productivity differential. This positive δ shifts the curve $h(c)$ upward to become curve $p(\mathrm{c})$. As a consequence, the intersection point is shifted leftward to D. Thus, the optimal cutoff level $\mathrm{c}_{\mathrm{g}}^{*}$ for a separating equilibrium is lowered and more workers are enrolled in the generous plan than without such a productivity effect.

Figure 3.2. Impact of A Constant Productivity Differential on Equilibrium



The intersection point C as illustrated in the figures is close to the terminal point, but this does not need to be the case, because it is quite possible that the $h(\mathrm{c})$ curve intersects the $(1-\alpha)\mathrm{c}$ line at a point close to the origin. Then, if the productivity effect is large enough, the

$p(\mathrm{c})$ curve may intersect the $(1-\alpha)\mathrm{c}$ line at the origin. In this case, the resulting equilibrium is high-end pooling. If, on the other hand, the $p(\mathrm{c})$ curve intersects the $(1-\alpha)\mathrm{c}$ line at the terminal point (where $\mathrm{c}=\gamma=100$), then the $p(\mathrm{c})$ curve would intersect the $(1-\alpha)\mathrm{c}$ line at an interior point, and the resulting equilibrium is separating. Hence, the productivity effect makes a high-end pooling equilibrium more likely and a low-end pooling one less likely than in the no-productivity effect case.

In the analysis above, the productivity differential δ is assumed to positive. While it may not be realistic to think that the generous plan has a smaller productivity effect than the moderate plan so the productivity differential is negative, I assume this is possible for the sake of analysis. If the productivity differential δ is negative, the direction of the shift stated above would be reversed, and so the optimal cutoff level $\mathrm{c}_{\mathrm{g}}^{*}$ would rise, thereby reducing the number of workers enrolled in the generous plan. Moreover, the conclusions regarding the likelihood of a pooling equilibrium should be reversed as well: a negative productivity effect makes a low-end pooling equilibrium more likely and a high-end pooling equilibrium less likely.

3.3. An Alternative Variable-Productivity Model

The model presented in the preceding section assumes that worker productivity varies across the health insurance plans but remains the same for any given plan. This section deals with a case where worker productivity changes with the proportion of workers enrolled in a particular plan. This may happen in a profession or job that requires interaction and emphasizes cooperation among workers. For example, a particularly desirable health insurance plan may somehow help boost morale or, through a causal chain, ultimately reduce absenteeism and thus interruption to a job or a production line, thereby smoothing cooperation, alleviating the burden of coordination, and increasing per worker output. In such cases, it makes sense to assume that the productivity of workers varies with the proportion of them covered under the desirable health insurance plan. This model can be interpreted as one where health insurance alters team productivity and the overall productivity rises (drops) faster than does the proportion of workers enrolled in a plan. To simplify the analysis here, I assume that the productivity for one of the plans (the moderate one) is fixed. Thus, only the productivity associated with the other plan (the generous one) changes as the proportion of workers enrolled in the generous plan changes.

3.3.1. The Worker's Problem

When the way in which productivity varies with the health insurance plans is changed, only the firm's decisions are affected and so the worker's problem remains the same as presented in the preceding section.

3.3.2. The Firm's Problem

As mentioned earlier, the productivity of workers enrolled in the moderate plan is assumed to be fixed and normalized to zero for simplicity, whereas the productivity of workers covered under the generous plan is assumed to be a function of the proportion of workers enrolled in the plan. Let $\delta(\rho_g)$ denote that productivity and assume that $\delta(0) = 0$. That is, if nobody is enrolled in the generous plan, then the plan does not generate a productivity differential. Hence, if $\delta(\rho_g^*) > 0$, then $\delta'(\rho_g^*) > 0$; if $\delta(\rho_g^*) < 0$, then $\delta'(\rho_g^*) < 0$; if $\delta(\rho_g^*) = 0$, $\delta'(\rho_g^*) = 0$. The converse also holds. For example, if $\delta'(\rho_g^*) > 0$, then $\delta(\rho_g^*) > 0$. All other assumptions remain unchanged. Thus, by definition,

$$\rho_g = \int_{c_g}^{\gamma} f(c)dc = 1 - F(c_g) \tag{3.3.1}$$

The firm's problem can be written as

$$\max_{c_g\in[0,\gamma]} \int_0^{c_g}\left(0-\frac{w_m}{1-t}-\alpha c\right)f(c)dc + \int_{c_g}^{\gamma}\left(\delta(\rho_g) - \frac{w_g}{1-t}-c\right)f(c)dc \qquad (3.3.2)$$

Expanding and shifting terms yields the following

$$\max_{c_g\in[0,\gamma]} -\frac{w_m}{1-t}-\alpha\int_0^{c_g} cf(c)dc-\int_{c_g}^{\gamma} cf(c)dc+\left(\delta\left(1-F(c_g)\right)+\frac{g(c_g)}{1-t}\right)\int_{c_g}^{\gamma} f(c)dc \quad (3.3.3)$$

As in the previous case, the domain is closed and bounded. Thus it has a maximum.

3.3.3. Equilibrium

Differentiating the function with respect to c_g yields the first order condition:

$$\left((1-\alpha)c_g^*-\frac{g(c_g^*)}{1-t}-\delta(\rho_g^*)\right)f(c_g^*)+\left(\frac{g'(c_g^*)}{1-t}-\delta'(\rho_g^*)f(c_g^*)\right)\left(1-F(c_g^*)\right)\begin{cases}\le 0, if\ c_g^*=0,\\ =0, if\ 0<c_g^*<\gamma\\ \ge 0, if\ c_g^*=\gamma\end{cases}$$

where $\rho_g^*=1-F(c_g^*)$. Shifting terms and dividing both sides by $f(c_g^*)$ ($f(c_g^*)\neq 0$ by assumption) yields the following first order conditions (FOC)

$$(1-\alpha)c_g^*-\frac{g(c_g^*)}{1-t}+\frac{g'(c_g^*)}{1-t}\frac{\left(1-F(c_g^*)\right)}{f(c_g^*)}\begin{cases}\le \delta(\rho_g^*)+\delta'(\rho_g^*)\left(1-F(c_g^*)\right), if\ c_g^*=0,\\ =\delta(\rho_g^*)+\delta'(\rho_g^*)\left(1-F(c_g^*)\right), if\ 0<c_g^*<\gamma\\ \ge \delta(\rho_g^*)+\delta'(\rho_g^*)\left(1-F(c_g^*)\right), if\ c_g^*=\gamma\end{cases}$$

This FOC differs from the FOC for the model analyzed in Section 3.2 in that the terms on the right-hand side of this one contain an additional term, $\delta'(\rho_g^*)$, which can be less than, equal to, or greater than zero. Suppose $\delta'(\rho_g^*)>0$. Then $\delta(\rho_g^*)>0$, and thus $\delta(\rho_g^*)+\delta'(\rho_g^*)\,(1-$

$F\left(c_g^*\right)\Big) > 0$. Then again, $c_g^* < c_g^{M*}$, where c_g^{M*} is the cut-off level in the Miller case. This means that the firm would enroll more workers in the generous plan when workers enrolled in the generous plan are more productive than they would if they were covered under the moderate plan.

Figure 3.3. Impact of A Variable Productivity Differential on Equilibrium

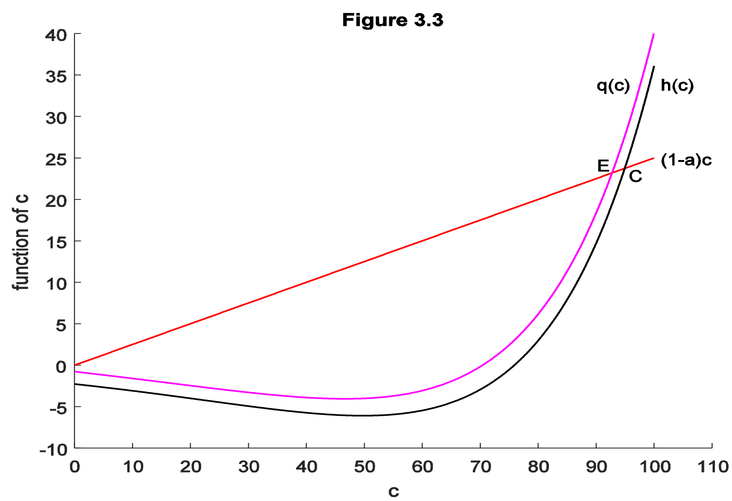


Figure 3.3 illustrates the effect of the productivity differential in this case. The curves for $(1-\alpha)c$ and $h(c)$ remain the same as in Figure 3.2. Curve $q(c)$ here is similar to $p(c)$ in Figure 3.2, and represents $\{\frac{g(c_g^*)}{1-t} - \frac{g'(c_g^*)}{1-t}\frac{\left(1-F(c_g^*)\right)}{f(c_g^*)} + \delta(\rho_g^*) + \delta'(\rho_g^*)\left(1-F(c_g^*)\right)\}$, where $\delta'(\rho_g^*)$ is assumed to be positive and therefore $\delta(\rho_g^*)$ is also positive. Thus, the overall impact of the shifting part consisting of these two terms is positive. Hence the productivity differential causes an upward shift, so curve $q(c)$ is above curve $h(c)$, thereby shifting leftward the point where the two curves intersect each other. As a consequence, the productivity effect reduces the optimal cut-off level c_g^*, as illustrated in the figure.

The difference between **Figure 3.2** and **Figure 3.3** is that the distance between the $h(\mathrm{c})$ and $p(\mathrm{c})$ curves in **Figure 3.2** is the same everywhere because the parameter δ for the $p(\mathrm{c})$ curve in **Figure 3.2** is a constant, whereas the distance between the $h(\mathrm{c})$ and $q(\mathrm{c})$ curves in **Figure 3.3** change with $\mathrm{c_g}$because $\delta(\rho_g^*)$and $\delta'(\rho_g^*)\rho_g$ incorporated in the $q(\mathrm{c})$ curve in Figure 3.3 each vary with $\mathrm{c_g}$[6]. However, the similarities the between the two models by far outweigh their discrepancies. They have similar implications for equilibrium outcomes; both have more workers enrolled in the generous plan than the Miller model if the generous insurance coverage enables workers to be more productive than does the moderate coverage.

The empirical predictions generated by the analysis of the productivity effects of health insurance using each of the two models depend upon whether such productivity effects (differentials) are positive or negative. Suppose the productivity effect is positive. Then firms would become more willing and more likely to offer health insurance and to share the costs of health insurance with their employees. Therefore, the empirical health insurance offer rates and take-up rates tend to be high for industries where the productivity effects are positive and low for industries where the productivity effects are negative or there are no such effects, everything else held the same. In cases where multiple insurance plans are offered and taken up, the offer rates and take-up rates for each type of health insurance plans in industries where each type of plans has a positive productivity effect relative to an inferior type should be higher than those for a comparable type in industries where none of the health insurance plans has a positive productivity impact. Although not necessarily true, high offer rates and/or high take-up rates

[6] For the curve shown in Figure 3.3, the productivity differential $\delta(\rho_g^*)$ is assumed to take this form: $\delta(\rho_g) = \delta_0 + \sqrt{\rho_g}$, where $\rho_g = 1 - F(\mathrm{c_g})$ and $\delta_0 \geq 0$. It can be proven that $\delta(\rho_g^*) + \delta'(\rho_g^*)\rho_g$ decreases as $\mathrm{c_g}$ rises (See Appendix). Thus, the distance between $h(\mathrm{c})$ and $p(\mathrm{c})$ curves shrinks as $\mathrm{c_g}$ rises.

should in general imply positive productivity effects. The empirical predictions of the two models are very similar.

3.4. Productivity Impacts and the Take-up Rate

This section studies a model where the firm offers only one health insurance policy to its workers, but still offers two compensation packages, one comprising a wage and the health insurance plan, and the other a wage only. Workers may take up the plan or turn it down. If a worker takes it up, the firm would share the expected health insurance cost with him by bearing a fixed portion of it, just as it would in the models presented earlier. If the worker chooses to reject the insurance offer, he may still be hired but will receive a wage-only compensation package; this contrasts the assumption made in the two-plan model that accepting an offer means accepting both an insurance plan and the wage. However, this model is the same as the one presented in Section 3.2 in all other aspects.

3.4.1. The Worker's Problem

Let $\mathrm{w_m}$ denote the wage associated with the insurance policy and $\mathrm{w_u}$ the wage that would be paid to workers who remain uninsured. As before, let $c \in [0, \gamma]$ denote the type or the expected health insurance cost of a worker, where $\gamma > 0$.

The health insurance provides a benefit to the enrollee. Let $m(\mathrm{c})$ denote the dollar-valued benefit a worker of type c would receive from the health insurance plan if he opts to enroll in the plan, where $m(c)$ is strictly increasing ($m'(c) > 0$ for $c \in (0, \gamma)$). The total (indirect) utility a worker receives from the compensation package offered by the firm equals the sum of a wage and the benefit from the insurance if the worker is insured and equals the wage if he is uninsured. Thus, if a worker is insured, the total utility he would receive from the compensation package is

$w_m + m(c_m)$. Otherwise, his total utility is w_u, the wage without insurance. Again, all workers are assumed to have (after-tax) reservation utility $\bar{u}$, which is independent of their types.

A worker would accept employment with the firm if and only if $w_u \geq \bar{u}$ or $w_m + m(c) \geq \bar{u}$ for $c \in [0, \gamma]$. A worker would choose to be insured if and only if $w_m + m(c) \geq w_u$ and $w_m + m(c) \geq \bar{u}$. The wage paid to workers who choose to be insured must be such that for the lowest-cost workers who are insured, $w_m + m(0) \geq w_u$ and $w_m + m(0) \geq \bar{u}$. Moreover, for the lowest-cost ones who opt to be insured, it must be true that $m(c_m) = w_u - w_m$. Thus, workers with expected health care cost $c \in [0, c_m)$ elect to be uninsured and receive a wage w_u, whereas workers with expected health cost $c \in [c_m, \gamma]$ opt to be insured and receive a wage w_m.

3.4.2. The Firm's Problem

The pretax cost that a worker would incur to the firm equals the pretax wage the firm pays to its workers, i.e., $\frac{w_u}{1-t}$, if the worker chooses to remain uninsured and equals the pre-tax wage (which may be different from the wage paid to uninsured workers) plus the share of the expected health insurance expense that the firm bears if the workers elect to be insured, i.e., $\frac{w_m}{1-t} + \alpha c$.

Productivity differs between insured and uninsured workers. Assume for simplicity that the difference in the their productivity is a constant. Now suppose that a worker's productivity is e if he is uninsured, and his productivity is $e + \delta$ if the worker is insured. The profit the firm obtains from each insured worker is $\int_{c_m}^{\gamma} \left(e + \delta - \frac{w_m}{1-t} - \alpha c\right) f(c)dc$ and the profit from each uninsured worker is $\int_0^{c_m} \left(e - \frac{w_u}{1-t}\right) f(c)dc$, given the distributional assumption about the expected health care cost or the type of workers as described in Section 3.2. The firm's objective is to maximize its profit given that workers accept its offers. Hence, the firm's problem can be expressed as

$$\max_{0\le w_m, 0\le w_u} \int_0^{c_m}\left(e-\frac{w_u}{1-t}\right)f(c)dc + \int_{c_m}^{\gamma}\left(e+\delta-\frac{w_m}{1-t}-\alpha c\right)f(c)dc \qquad (3.4.1)$$

s.t.:

$$w_u \ge \bar{u}$$

$$w_m + m(c_m) \ge w_u$$

The first is the participation constraint and the second the incentive-compatibility constraint. The two constraints do not need to hold simultaneously. In fact, the first constraint does not need to hold if the firm wants all its workers to be insured. If some workers opt to be insured and others choose to remain uninsured, then both constraints may hold. Optimality requires that $w_u = \bar{u}$ for the uninsured workers and that the lowest-cost workers who are insured are indifferent between insured and uninsured. So $w_u = w_m + m(c_m)$, or $m(c_m) = w_u - w_m = \bar{u} - w_m$. Because w_u is constant, choosing w_u and $w_m$is therefore equivalent to choosing $m(c_m)$. Furthermore, by assumption, $m(c_m)$ is a monotonic function of c_m. Thus, choosing $m(c_m)$ is equivalent to choosing c_m. Therefore, the firm's problem can be expressed as

$$\max_{c_m\in[0,\gamma]} \int_0^{c_m}\left(e-\frac{w_u}{1-t}\right)f(c)dc + \int_{c_m}^{\gamma}\left(e+\delta-\frac{w_m}{1-t}-\alpha c\right)f(c)dc \qquad (3.4.2)$$

Expanding the expression and collecting terms leads to the following expression

$$\max_{c_m\in[0,\gamma]} e-\frac{\bar{u}}{1-t}+\delta\int_{c_m}^{\gamma}f(c)dc-\alpha\int_{c_m}^{\gamma}cf(c)dc+\frac{m(c_m)}{1-t}\int_{c_m}^{\gamma}f(c)dc \qquad (3.4.3)$$

The term $e - \frac{\bar{u}}{1-t}$ is the total productivity net of the total wage cost if all workers remain uninsured, the second term is the additional profit generated by insured workers, the third term is the expected health insurance cost the firm has to pay for those workers who are insured, and the last term is the wage savings from insured workers.

3.4.3. Equilibrium

Differentiating the profit function with respect to c_m yields the first order conditions, which can be written as

$$\left(-\delta + \alpha c_m^* - \frac{m(c_m^*)}{1-t}\right) f(c_m^*) + \frac{m'(c_m^*)}{1-t}\left(1 - F(c_m^*)\right) \begin{cases} \leq 0, if\ c_m^* = 0, \\ = 0, \text{if } 0 < c_m^* < \gamma \\ \geq 0, if\ c_m^* = \gamma \end{cases}$$

Note that the constant vanishes after differentiation. This expression is very similar to the first order conditions for the two-plan model presented in Section 3.2 of Chapter 3. Interpretations of the terms are also similar and so no elaboration is necessary.

If the insurance plan does not have any effect or even has an adverse effect on worker productivity, then $\delta \leq 0$. Then it is not optimal to insure all workers, or $c_m^* = 0$, because in this case, the left-hand side would be positive, contradicting the first order condition that it is negative. If $\delta = 0$, then $c_m^* < \gamma$ as long as $\alpha\gamma < \frac{m(\gamma)}{1-t}$, for otherwise the first order condition would be violated.

If $\delta > \alpha\gamma - \frac{m(\gamma)}{1-t}$, then it is not optimal to leave all workers to remain uninsured, i.e., $c_m^* = \gamma$, because then the last term would vanish. However, since the term $-\delta + \alpha\gamma - \frac{m(\gamma)}{1-t} < 0$, the left-hand would be less than zero, contradicting the first order condition that it is non-negative. Hence, $\delta > \alpha\gamma - \frac{m(\gamma)}{1-t}$ implies $c_m^* < \gamma$. This states that as long as the productivity differential is greater than the difference between cost savings and the foregone wage savings, then the firm should insure at least workers toward the high-cost end.

Conversely, $c_m^* < \gamma$ does not imply $\delta > \alpha\gamma - \frac{m(\gamma)}{1-t}$, because $c_m^* < \gamma$ implies either $\frac{m'(c_m^*)}{1-t}\left(1 - F(c_m^*)\right) \leq \delta$ or $\left(\alpha c_m^* - \frac{m(c_m^*)}{1-t}\right) f(c_m^*) + \frac{m'(c_m^*)}{1-t}\left(1 - F(c_m^*)\right) = \delta$. In the former case, it is possible that $\delta \leq \alpha\gamma - \frac{m(\gamma)}{1-t}$, whereas in the latter case, c_m^* can be such that $\left(\alpha c_m^* - \frac{m(c_m^*)}{1-t}\right) f(c_m^*) + \frac{m'(c_m^*)}{1-t}\left(1 - F(c_m^*)\right) \leq \alpha\gamma - \frac{m(\gamma)}{1-t}$, so $\delta \leq \alpha\gamma - \frac{m(\gamma)}{1-t}$.

If the productivity differential is positive, or alternatively, insurance boosts worker productivity, then it may be optimal to insure all workers, because if $c_m^* = 0$, then $\frac{m'(c_m^*)}{1-t}\left(1 - F(c_m^*)\right) \leq \delta$, but since $\delta > 0$, the condition can hold for some c_m^*. Thus, the productivity enhancing effect of health insurance makes it more likely for the firm to insure its workers than if health insurance does not affect productivity.

This can be called a take-up model because workers may choose to either take up the insurance plan or not to do so. It is similar to and can be treated as a special case of the two-plan model where the value of the moderate plan is set to zero. However, the there is an important difference, which involves the reservation utility of workers: the reservation utility is the same for all workers in this model, whereas in the two-plan model, the moderate plan plays the role of relaxing the equal reservation utility restriction: the higher valuation of the moderate plan to the high-cost workers means that the high-cost workers have a lower reservation utility than the low-cost workers. Another difference between this one-plan and the two-plan models lies in the distribution of information rents: workers stay on the low-end (uninsured) receive no rents, whereas in the two-plan case, workers who remain on the low-end receive information rents.

3.5. Discussions

The first variable productivity models make the assumption that productivity varies with different plans but is equal across individual workers enrolled in the same plan. A potential argument against this assumption is that if workers are different in terms of their expected health care costs, the benefits each worker would obtain from the health insurance plan should also be different and therefore the effects of the plan on productivity should be different as well. The justification for equal productivity impact for a given plan is that the productivity impact is purely a result of enrolling in a health insurance plan and that the characteristics of individual workers or the benefits they obtain from the plan are not quantitatively relevant. The assumption underlying the alternative model where productivity varies with the proportion of workers enrolled in the generous plan seems to be more natural, but relative to the Miller model, the results derived from this model are directionally the same as those derived from the other model.

The two-plan models assume all workers are offered health insurance and accept the offers. This assumption may not be realistic, because in the real world, many firms end up with offering no insurance and many workers may end up with no insurance coverage either, though it is not quite clear whether this is because the firm did not offer any insurance plan or because none has been taken up. This limitation makes the two-plan models unsuited for empirical analysis. The one-plan model overcomes the limitation.

3.6. Conclusions

The variable-productivity models can potentially have both pooling (high-end) and separating equilibria even if other assumptions remain the same as in the Miller model; in contrast, Miller's model only allows for a separating equilibrium, given its assumptions. When the productivity differential is positive, the high-end pooling equilibrium is possible, but the low-end equilibrium is less likely to exist than in the Miller case. When the productivity differential is negative, the opposite is true. For the separating equilibrium, a positive productivity differential incentivizes the firm to induce more workers into the generous plan, whereas a negative productivity differential encourages the firm to do the opposite. The case where the productivity differential is zero is equivalent to the Miller case.

When the (generous) health insurance plan boosts productivity more than does the moderate plan (or no plan), the existence of a high-end pooling equilibrium (where all workers are covered under the (generous) health insurance plan) does not require that the marginal cost (transfer cost) of enrolling the zero-cost workers in the (generous plan) be zero, because the productivity enhancement effect of enrolling workers in the (generous) health insurance plan may add to the incentive for the firm. As long as the (generous) health insurance plan boosts worker productivity relative to the moderate plan (or no-plan) to such an extent that the productivity differential between the two plans exceeds the marginal cost of enrolling the zero-cost workers in the (generous) plan, regardless of whether this marginal cost is zero or positive, the firm would find it profitable to enroll all workers in the (generous) plan. In other words, a zero marginal cost of enrolling the zero-cost workers in the generous plan is not a necessary condition for the model to have the high-end pooling equilibrium. This feature differentiates the models studied in this chapter from Miller's, where the marginal cost of enrolling the zero-cost workers in the generous plan has to be zero in order for a high-end pooling equilibrium to exist. Whether a low-end pooling (where workers are all enrolled in the moderate plan in the case of the two-plan model or are all insured in the case of the one-plan model) may exist depends upon whether health insurance savings from the moderate plan outweighs its additional wage cost plus the additional productivity foregone as a result of shifting workers from the (generous) plan to the moderate one (no-plan).

If information is symmetric or if discrimination is not allowed, all workers except for the lowest-cost ones receive rents, but the firm does not, as in the Miller case. This is because, in any equilibrium, the participation constraint is binding only for the lowest-cost workers. Whether the productivity differential quantitatively affects the distribution of benefits (rents) among workers depends on the type of the equilibrium that would be obtained without such a productivity differential. In the case where the equilibrium is separating when health insurance does not affect worker productivity, the distribution of benefits to workers is changed by the productivity effect relative to the no-productivity effect case, as the firm shares the productivity differential (if any) with its workers, be it positive or negative. In the case where the equilibrium is high-end pooling when health insurance has no productivity impact, the distribution of rents would not be changed by a positive productivity differential relative to the no-productivity case, because other incentives such as the tax-deductibility of health insurance are already enough to motivate the firm to cover all workers under the generous plan. As a result, the firm would receive all the increased productivity benefits. However, if the productivity differential is negative, the distribution of (information) rents may be altered by a negative productivity differential, because the negative productivity differential may change the high-end pooling equilibrium in the no-productivity case to a separating one, in which case the negative productivity differential would be shared between the firm and the workers. In the case where the equilibrium would be low-end pooling, the distribution of rents may or may not be changed either, depending on whether the productivity effect is large enough to alter the conditions to such an extent that the low-end pooling equilibrium in the no-productivity case would be changed to a separating one.

The variable-productivity models studied in its various forms has addressed the potential issue that Miller's fails to address. However, these models do not accommodate situations where people face credit or liquidity constraints or where the firm possibly affects the choice of a worker by using a mechanism other than through manipulating the wage associated with a particular health insurance plan. Another important issue with the take-up model presented in Section 3.4 is the assumption that all workers have the same reservation utility. The consequence is that uninsured workers have to live at their reservation utility only and have no buffer against potential adverse shocks. This is undesirable because workers with different health risks are assumed to live at the same minimum level. The next chapter will present a model to address these issues.

CHAPTER IV

AN ALTERNATIVE ONE-PLAN MODEL

4.1. Introduction

The models presented in the preceding chapters, including the one-plan model, are based on an implicit assumption that people are indifferent between wage benefits and benefits gained through health insurance, as long as the two types of benefits are quantitatively the same. As a consequence of assuming workers are indifferent between the two types of benefits, the firm can "trade" wage for health insurance cost. While this assumption is appropriate in many cases, it fails to consider situations where people may view these two types of benefits differently and prefer one type to the other. For example, people may have other needs that require cash payments and so may prefer wage to health insurance, at least to some extent.

There are possibly several reasons why people prefer wage to health insurance to come extent. First, the benefits gained through health insurance are individual-specific, so even the same health insurance plan has different values to different people. Second, a health insurance plan cannot be divided into several smaller ones. Third, health insurance may not be (legally) transferred from one person to another once it has been offered to and accepted by an individual. Even without the third property, the first two properties, individual-specific valuation and indivisibility, already makes health insurance plan a hard-to-trade commodity. The non-transferability property of health insurance leads to its non-convertibility: once purchased, health insurance cannot be converted back into cash or cash equivalent. Wage, on the other hand, is itself cash or cash equivalent or can easily be converted into the same, and thus can be used to meet individual's any other needs. The non-transferability and therefore non-convertibility of health insurance plan makes it less preferred for some people or to some extent.

When a person faces credit or liquidity constraints, the individual may not be able to pay for their other, alternative needs if, in the context of Miller's model, the worker is paid at the bare minimum level, which is the reservation utility at the most. While there might be multiple ways to address the issue associated with or arising out of credit or liquidity constraints, this study does this by postulating that, in order to improve his or her ability to pay for alternatives in life and overcome the difficulty when facing such constraints, the person requires a level of wage that is above the bare minimum utility, rather than simply requires a satisfactory total compensation package, because the health insurance benefit cannot be used to address the issue arising out of credit or liquidity constraints.

This chapter will present a simple, one-plan framework to study the issue raised above. The one-plan models provide a convenient way to attain the same goal as would any other form of models such as two-plan models. By comparison, a two-plan model would complicate the analysis but would not offer much additional benefits. An additional advantage of the one-plan model to be presented in this chapter is that, by assuming workers require reservation wages, this model can bypass the sort of pitfall faced by the model presented in Section 3.4 that uninsured workers live at the bare minimum level.

The remainder of the chapter is structured as follows. Section 2 presents and analyzes the model. Section 3 concludes the chapter.

4.2. The Model

This section presents a model that assumes workers require reservation wages. Workers require reservation wages because they may face credit or liquidity constraints so health benefits cannot be used as substitutes for wages. When workers face credit or liquidity constraints, the workers who choose to be uninsured (if any) may require a wage that is at least as high as their reservation wage which is above their reservation utility.

For analytical purposes, the roles the reservation wages play are similar to those of the reservation utility. Hence, the reservation wages are not important in and of themselves. What makes them critical to this model is the difference between the two reservation wages. For example, if the difference is zero, or equivalently, the two reservation wages are equal, then the equilibrium outcomes are entirely driven by wage differentials, and thus there would be no difference between this model and the one-plan model analyzed in Section 3.4 of Chapter 3.

4.2.1. The Worker's Problem

As in the preceding sections, assume workers are of continuous types. Moreover, if a worker is insured, he would receive a total benefit equal to $\mathrm{w_m} + m(\mathrm{c}\,)$; otherwise, he only receives wage $\mathrm{w_u}$. All other assumptions and notations remain the same as in the preceding chapter, whichever are applicable, unless otherwise noted.

Workers have to be compensated sufficiently in order for them to join the firm. This means the compensation level not only has to be at least as large as their reservation utility, but also should be at least at high as their required (reservation) wages, depending on which package they would receive from the firm. Suppose a worker has to be paid a reservation wage at $\overline{\mathrm{w_l}}$ if he chooses to be insured, otherwise, he would have to be paid a wage at least equal to $\overline{\mathrm{w_h}}$. As mentioned at the beginning of this section, these reservation wages are assumed to be above their reservation utility. Thus, workers would accept employment with the firm if and only if $w_u \geq \overline{w_h} \geq \bar{u}$ or $w_m \geq \overline{w_l} \geq \bar{u}$. They would choose to be insured if and only if $w_m + m(c_m) \geq \bar{u}$ and $w_m + m(c_m) \geq w_u$.

4.2.2. The Firm's Problem

The firm offers two packages, one comprising a wage which is not lower than the reservation at the lower bound and a health insurance plan, and the other consisting of only a wage, which is at least as high as the reservation wage at the upper bound. In principle, the firm may adjust the levels of the two wages, but for practical purpose, it may optimally fix one wage at the upper bound of the reservation wage and change the other wage to achieve an optimal level of insurance rate.

Let α denote the fixed proportion of the expected health insurance cost to be borne by the firm. So the cost that the firm would need to assume is αc if a worker of type c chooses to be insured, and the portion of the expected health insurance cost worker would need to bear is $1 - \alpha$.

Productivity is assumed to alter with the insurance status of workers. If a worker is uninsured, his productivity would be e, but if the worker is insured, his productivity would be $e + \delta$. Given the same distributional assumption about the worker types (or workers' expected health care cost) presented in earlier, the profit from each uninsured worker is $\int_0^{c_m} \left(e - \frac{w_u}{1-t}\right) f(c)dc$ and the profit the firm obtains from insured worker is $\int_{c_m}^{\gamma} \left(e + \delta - \frac{w_m}{1-t} - \alpha c\right) f(c)dc$. Given all the assumptions, the firm's problem can be expressed as follows.

$$\max_{w_m, w_u} \int_0^{c_m} \left(e - \frac{w_u}{1-t}\right) f(c)dc + \int_{c_m}^{\gamma} \left(e + \delta - \frac{w_m}{1-t} - \alpha c\right) f(c)dc$$

s.t.:

$$w_u \geq \overline{w_h} > \overline{w_l} \geq \bar{u}$$

$$w_m \geq \overline{w_l}$$

$$w_m + m(c_m) \geq \bar{u}$$

$$w_m + m(c_m) \geq w_u$$

Again, the layout of the constraints is just for convenience or for aesthetics reasons. Some of the constraints may be redundant and even if they are not redundant they do not need to hold altogether. The first inequality in the first chain of constraints represents the participation constraint for workers who would opt to be insured. This constraint says that if a worker is uninsured, he would have to be paid a wage at least as high as the required reservation wage $\overline{w_h}$, which is greater than the reservation utility. The relationship between the two levels of reservation wages against the reservation utility holds by assumption. The second constraint restricts the level of wage that would have to be paid to workers who take up the insurance policy. The third is the participation constraint for any worker to accept the wage-insurance package, but a quick examination reveals that this is not necessary because it is implied by the first and the fourth combined; the fourth is the incentive compatibility constraint: the total benefit from the wage-insurance package has to be at as good as the wage-only package that would be offered to workers who choose to be uninsured. Optimality requires that the first constraint be binding, that is, $w_u = \overline{w_h}$. Thus, the firm's problem can be re-written as

$$\max_{w_m} e - \frac{\overline{w_h}}{1-t}\int_0^{c_m} f(c)dc + \delta\int_{c_m}^{\gamma} f(c)dc - \int_{c_m}^{\gamma}\left(\frac{w_m}{1-t} + \alpha c\right) f(c)dc$$

s.t.:

$$w_m \geq \overline{w_l} \geq \bar{u}$$

$$w_m + m(c_m) \geq \bar{u}$$

$$w_m + m(c) \geq \overline{w_h} > \overline{w_l}$$

By continuity of the wage w_m, at least one of the two constraints has to be binding for some $c \in [0, \gamma]$ at the optimum, for otherwise the firm can lower the wage by a small amount; as long as neither of the two is binding, this process can continue. Which of the two constraints would be binding depends on $\overline{w_l}$ relative to $\overline{w_h}$, and the health insurance benefit function $m(c)$. By continuity of w_m and c, there exists w_m (and c) that can make the second constraint binding, so $w_m + m(c_m) = \overline{w_h}$. The constraint on c_m is determined by the condition that $m(c_m) \leq \overline{w_h} - \overline{w_l}$ because $w_m \geq \overline{w_l}$. Therefore, the restriction on the wage paid to the insured workers can be equivalently imposed on $m(c_m)$, and finally equivalently on c_m, because $m(c)$ is monotonic in c. Since $c = m^{-1}(c)$, where $m^{-1}(c)$ is the inverse function of c; in addition, $c \in [0, \gamma]$. Hence, $0 \leq c_m \leq \min\left(\gamma, m^{-1}(\overline{w_h} - \overline{w_l})\right)$.

On the other hand, $m(c_m)$ is bounded from below by the second constraint, that is, $w_m + m(c_m) \geq \bar{u}$ for any w_m. The minimum level of w_m is determined by the constraint $w_m \geq \overline{w_l}$. So if $\overline{w_l} + m(c_m) \geq \bar{u}$ holds, the second constraint can hold as well. Thus, $m(c_m) \geq \bar{u} - \overline{w_l} \geq 0$. Because $\overline{w_l} \geq \bar{u}$, the condition that $m(c_m) \geq 0$ dominates. The last inequality derived in the preceding paragraph holds, which is $0 \leq c_m \leq \min\left(\gamma, m^{-1}(\overline{w_h} - \overline{w_l})\right)$.

The firm's profit maximization problem can then be modified as follows.

$$\max_{c_m} e - \frac{\overline{w_h}}{1-t}\int_0^{c_m} f(c)dc + \delta\int_{c_m}^{\gamma} f(c)dc - \int_{c_m}^{\gamma}\left(\frac{\overline{w_h} - m(c_m)}{1-t} + \alpha c\right) f(c)dc$$

s.t.:

$$0 \leq c_m \leq \min\left(\gamma, m^{-1}(\overline{w_h} - \overline{w_l})\right)$$

The upper bound of the integration interval remains unaffected by the imposition of restrictions on the upper bound of c_m. The objective function can be expanded and re-written as

$$\max_{c_m} e - \frac{\overline{w_h}}{1-t} + \delta\int_{c_m}^{\gamma} f(c)dc + \int_{c_m}^{\gamma}\left(\frac{m(c_m)}{1-t} - \alpha c\right) f(c)dc$$

where the first two terms are constants. This problem is similar to the one presented in Section 4.2. and is pretty straightforward to solve.

4.2.3. Equilibrium

Differentiating the aforesaid objective function with respect to c_m, given the constraints on c_m, yields the the following first order conditions:

$$\left(-\delta + \alpha c_m^* - \frac{m(c_m^*)}{1-t}\right) f(c_m^*) + \frac{m'(c_m^*)}{1-t}\left(1 - F(c_m^*)\right) \begin{cases} \leq 0, if\ c_m^* = 0 \\ = 0, \text{if } 0 < c_m^* < \min\left(\gamma, m^{-1}(\overline{w_h} - \overline{w_l})\right) \\ \geq 0, if\ c_m^* = \min\left(\gamma, m^{-1}(\overline{w_h} - \overline{w_l})\right) \end{cases}$$

The first order condition implicitly determines c_m^*. The productivity differential δ is present in the first order condition so it affects the optimal cut-off level c_m^* (in the same way at it does in the models presented in Chapter 3). Note that the reservation constraint does not affect the condition for the existence of a high-end pooling because the two reservation wages do not enter the first order condition for $\mathrm{c}_{\mathrm{m}}^{*} = 0$. Hence, there is no difference between this model and any of the models presented earlier when $\mathrm{c}_{\mathrm{m}}^{*} = 0$. The difference between this and any of those presented in Section 3.4 of Chapter 3 is that $\mathrm{c}_{\mathrm{m}}^{*}$ may have a narrower range if $\overline{w_h} - \overline{w_l}$ is small. But if $\min\left(\gamma, m^{-1}(\overline{w_h} - \overline{w_l})\right) = \gamma$, then there is no difference between this model and the one presented in Section 3.4. This implies that $\overline{w_h} - \overline{w_l}$ is too large and the constraint on the lower bound of the wage for insured workers may not be binding. Suppose $\overline{w_h} - \overline{w_l}$ is not too large so $\min\left(\gamma, m^{-1}(\overline{w_h} - \overline{w_l})\right) = m^{-1}(\overline{w_h} - \overline{w_l}) < \gamma$. Then an immediate conclusion can be drawn that at least some workers would choose to be insured. The reason is that when insured workers are guaranteed a reservation wage that is not too far below the level for uninsured ones, some workers would always choose to be insured. However, on the lower bound there is not much difference between this model and the ones analyzed in Section 3.4 of Chapter.

An evident difference between this model and the one-plan model presented in Section 3.4 of Chapter 3 lies in the distribution of rents when the equilibrium is low-end pooling, where no workers are insured. In the low-end pooling equilibrium for the model analyzed in Section 3.4, the firm extracts all rents and workers receive no rents. In this model, however, all workers, including the lowest-cost ones, may receive rents even if the equilibrium is low-end pooling because their reservation wage is above their reservation utility. When the equilibrium is separating or when it is high-end pooling where all workers are insured, all workers, except for the lowest-cost ones, receive rents. Whether the firm obtains rents depends not only on the type of an equilibrium, but also on whether the average benefit of the insurance plan the uninsured workers would obtain if they chose to be enrolled in the plan exceeds their reservation wage. If the equilibrium is low-end pooling where all workers are uninsured or if it is separating, the firms would extract rents from uninsured workers, provided that the average benefit of the insurance plan that would be obtained if such workers chose to be insured exceeds their reservation wage. The total amount of rents equals the difference between the average benefit uninsured workers receive from the health insurance plan and the reservation wage multiplied by the number of uninsured workers who would be enrolled in the plan. If the equilibrium is high-end pooling in which all workers are insured, all rents would accrue to the workers.

4.3. Conclusions

The model presented in this chapter shares some similarities with models presented earlier in that all are within the broad, general framework. Notwithstanding the similarities, there is a critical difference in terms of rents distribution when the equilibrium is pooling in which workers are uninsured or when the equilibrium is separating; in either case, the firm extract rents from uninsured workers, whereas in the Miller case, the firm extracts no rents in the corresponding equilibrium. The model studied in this chapter has an equilibrium outcome in which every worker receives some rents because the worker's required reservation wage is above his or her reservation utility. Another important benefit of assuming workers require a reservation wage is that it overcomes the problem that the one-plan model presented in Chapter 3 faces: uninsured workers live at the bare minimum reservation utility level.

Workers demand for reservation wages for a variety of reasons, including liquidity or credit constraints or other practical considerations. In this model, workers receive rents, whatever the type of the equilibrium is. Whether the firm extracts rents depends on the type of the equilibrium and whether the average benefit of the insurance plan that the uninsured workers would receive if they chose to be enrolled in the plan exceeds their reservation wage.

CHAPTER V

ANALYTICAL DATA

5.1. Introduction

This chapter presents the analytical dataset for the empirical work that follows. The data are drawn from the Household Component (HC) of the Medical Expenditure Panel Survey (MEPS), a large overlapping panel of the U.S. civilian noninstitutionalized population. The MEPS-HC is designed so that each person is interviewed for five rounds over two consecutive calendar years. The MEPS-HC was initiated in 1996 with a new panel recruited each year from the previous year's respondents to the National Health Interview Survey. The MEPS-HC contains detailed information on demographic characteristics, health status, health insurance coverage, and labor market participation. Demographic characteristics include age, sex, race, marital status, and education measured by years of schooling. Health status includes disability days and self-reported (physical) health status and mental health status. Health insurance coverage data includes whether an individual was insured, whether the employer of the person offered insurance, and whether the individual was insured through his or her employer or union. Labor market participation includes employment status at the time the survey was conducted, hourly wages the individual earned, the size of the firm or location that employed the individual at the time of the survey, the industry the firm belonged to, and the individual's occupation.

While the MEPS-HC began in 1996, I only use data starting from 2002 because the categorization of industries and occupations changed in 2002; the data ends in 2014 as this is the latest data available. I observe most people for two years, except for people who were recruited into the MEPS in 2001 and those who were recruited in 2014. In the former case I only observe their second year of data and in the latter the first year. Moreover, the MEPS-HC only includes information on people living the household. The survey design ensures that each individual person appears only once in a year for at most two consecutive years. The spouse, if present in the household, appears in both years, though the spouse may not be the same person because a person might be married during the initial round of survey, get divorced in the second round, and then remarry in the third round; if a person changed his or her spouse, this change would be reflected in the information collected each round, but only the spousal information as of December 31 of the survey year will be used for the analysis in this part. The MEPS-HC also contains spousal information for married people. The information about the spouse of an individual was collected at the same time when the information on that individual was collected, provided that the spouse was living in the household (spouses may live apart for a variety of reasons, including military deployment, different locations of employment, expatriation, etc.). If a person changed marital status across rounds of survey, this information would be indicated in rounds; however, only the year-end status would be used for the purpose of analysis in this dissertation.

Statistics show that the extent that the data used in the analysis represent the national population is very similar to the extent represented by information from other sources. For example, the insurance offer rate (simply the proportion of the number of people offered health insurance through their employer or unions to the total number of people in the sample) is close to 69%, compared with the national average of around 72% over the years as reported in the Employer Health Benefits Survey.

The remainder of the chapter is structured as follows. Section 2 describes the data cleaning and selection process. Section 3 presents descriptive statistics and cross-tabs for certain variables believed to be endogenous. Section 4 briefs the implications of the available information for identifying strategies. Section 5 discusses the advantages and limitations of the data and concludes the chapter.

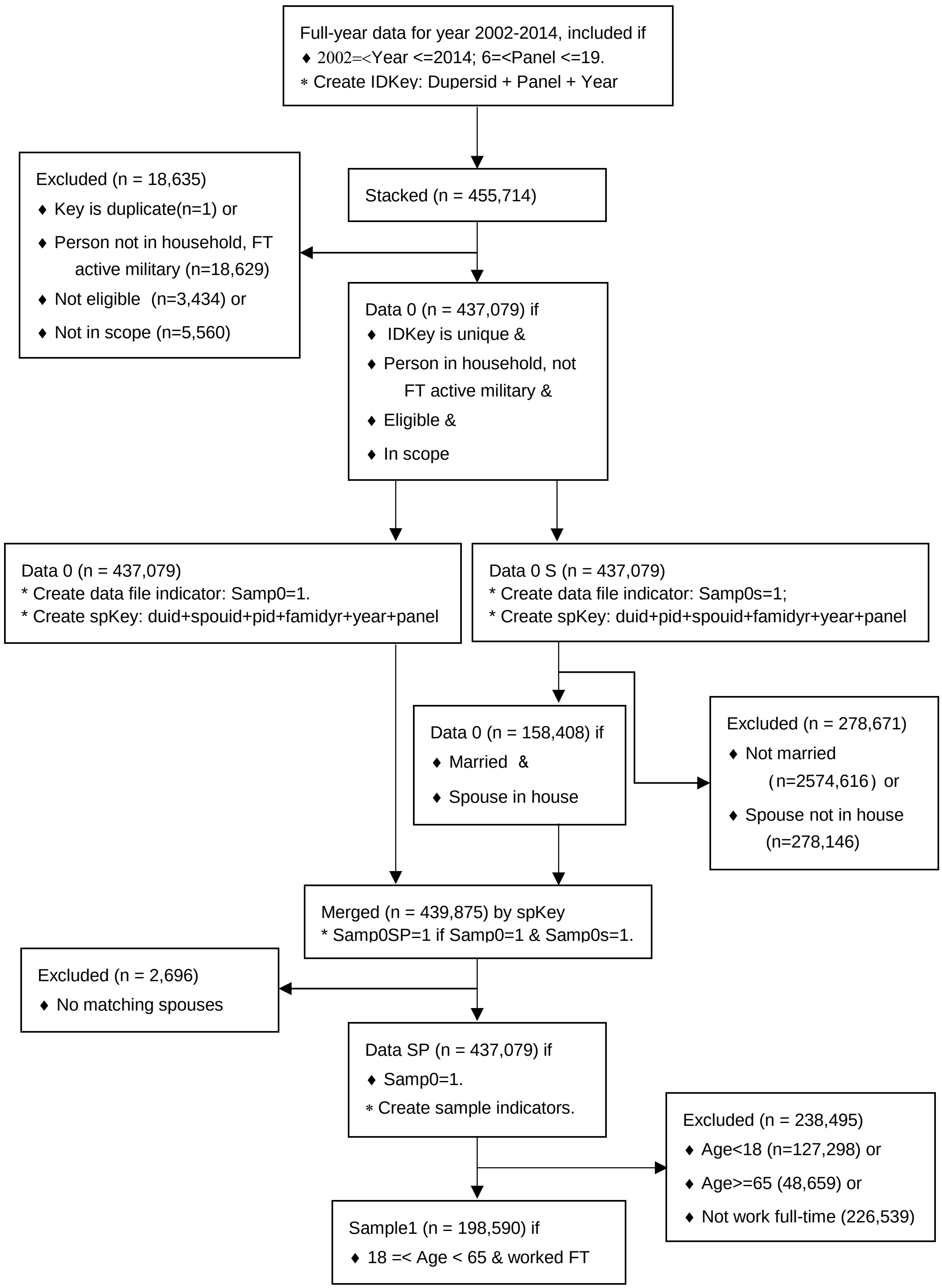
5.2. Sample Selection and Variable Construction

5.2.1. Sample Selection Process

Figure 5.1 illustrates the data cleaning and sample selection process. A final analytical dataset (Sample 1) is obtained in the end. The defined sample indicators can be used to extract the required datasets for various analytical purposes. The Sample Construction CONSORT Diagram is **Figure 5.1**.

In **Figure 5.1**, each box represents a step; for example, the first box represents the first step (Step 1), etc.. Each arrow represents a transition from one step to the next. Parallel arrows represent independent transitions that may happen simultaneously. A box located at the start of an arrow precedes the step represented by a box to which the arrow points. A star * represents an action and a diamond represents a condition. The procedures for creating the analytic dataset are as follows.

Figure 5.1. Sample Construction CONSORT Diagram



5.2.1.1. Step 1: Obtaining Source Data Files and Create an Identifier

Data files for years 2002 through 2014 and Panels 6 through 19 were extracted from the MEPS raw data files available at AHRQ's (Agency for Healthcare Research and Quality) website. The variable Year should have a value between 2002 and 2014, inclusive, whereas the variable Panel should have a value between 6 and 19, inclusive. Any observation that does not have a value falling within the range for the two variables are excluded. In the meantime, a unique identifier that consists of the individual's identifier (a variable already available in the source data files that consists of the dwelling identifier and the person identifier), Year, and Panel was created by concatenating these variables. This unique identifier variable (called IDKey in the diagram) was used in subsequent steps to obtain unique observations for each combination of year and panel. "Unique observation" as used here means that an individual appears only once in each combination of year and panel; it does not mean that an individual appears only once in the entire combined data file.

5.2.1.2. Step 2: Stacking Data Files

The data files obtained in Step 1 were then stacked by vertically concatenating these data files. The resulting combined data file contains 455,714 observations, which equals the sum of the number of observations in all data files obtained in Step 1.

5.2.1.3. Step 3: Creating Data File without Duplicates and with Non-institutionalized Persons Who Were Eligible for and In Scope of MEPS Data Collection

The combined data file created in step 2 was screened using the inclusion criteria that the person was in the household (not full-time active in military), eligible for inclusion in MEPS data collection, and in the scope of data collection. Moreover, the data file was also screened for duplicate observations using the unique identifier (IDKey) created in Step 1 to obtain a data file that does not contain duplicates. As a consequence of this step, a total of 18,635 observations were excluded and the resulting data file contains 437,079 unique observations.

5.2.1.4. Step 4: Creating Own Data File with Spousal Match Key

A copy of the data file created in Step 3 was saved as a separate data file. This data file was then used to create a match Key (spKey in Diagram 5.1) that consists of the individual's dwelling identifier, spousal identifier, own person identifier, family identifier, year, and panel. The match key will be used to merge the own data file and the spousal data file that is to be created later. A data file indicator that equals 1 for this file was created for the data file created in Step 3.

5.2.1.4'. Step 4': Creating Spousal Data File with Spousal Match Key

The data file created in Step 3 was used to create the spousal data file by renaming certain variables as the spousal variables. Moreover, a spousal match key (spKey in Diagram 5.1) and a spousal data file indicator (called Samp0s in Diagram 5.1) were also created. The spousal match key consists of the individual's dwelling identifier, own person identifier, spousal identifier, year, and panel. The spKey created in the spousal data file is not exactly the same the IDKEY created earlier. Only the renamed spousal variables, the spousal data file indicator, and the spousal match key created in this step were kept in this data file.

5.2.1.5. Step 5: Creating Data File for Married People whose Spouses Were in Household

Persons who were not married or persons who were married but whose spouses were not in the household should be excluded from the spousal data file created in Step 4' and only married people whose spouses were in the household should be included in the data file. Based on these criteria, a total of 278,671 observations were excluded, and the resulting data file contains 158,408 observations.

5.2.1.6. Step 6: Merging Own Data file and Spousal Data File

The own data file created obtained in Step 4 and the spousal data file created in Step 4' were merged by the match Key. The resulting combined data file has 439,875 observations. A spousal sample indicator was created, which equals 1 if the own data file indicator created in Step 4 and the spousal data file indicator both equal 1 and equals 0 otherwise.

5.2.1.7. Step 7: Excluding Persons without Matching Spouses

Observations from the spousal data file that do not have matching spouses are excluded from the merged data file created in Step 6 by setting Samp0 = 1. Based on this criterion, a total of 2,696 observations were excluded and the resulting data file has 437,079 observations. This data file contains not only all the observations contained in the own data file created in Step 4, but also the observations contained in the spousal data file with matched spouses. The (only) difference between this merged data file and the one created in Step 3 is that the newly merged data file contains spousal variables, whereas the data file created in Step 3 does not. Certain sample indicators were created using the own and spousal data files created within this Step.

5.2.1.8. Step 8: Creating the Analytical Data File

The combined dataset has two components: the original data, which contains the original variables, and the spousal data, which contain spousal variables. Observations that are appropriate for the analytical purposes include individuals who worked full-time and were at least 18 but younger than 65 years of age. Imposing these restrictions yields a data file, labeled Sample1, which contains 198,690 observations. This is the analytical data file from which various modeling samples can be obtained by applying relevant sample indicators.

The data selection process ensures that the data file does not have duplicates because the merge key (a combination of the dwelling-unit-person identifier, year, and Panel) uniquely identifies each individual. In addition, the missing values generated (in the spousal data, persons who do not have spousal identifiers and whose spousal identifiers are inconsistent do not have matches) can be easily identified by using the appropriately defined indicators.

5.2.2. Construction of Variables

Variables were constructed during the data extraction and selection process. Because most variables in the source data files are either round- or year-specific, variables with names common to all data files were created. Since the survey was conducted over five rounds, certain information changed from round to round, either because of changes in economic or other circumstances, or because the variable is cumulative in nature. In addition, the MEPS-HC specifically asked about year-end values (e.g. employment status and health insurance coverage) for some variables.

For most (non-cumulative) variables, the year-end values were used whenever possible, and if that was not possible, then the next most recent value was used. For variables whose values are cumulative, the sum of the values for that year over the rounds is used, except in 2002 (the starting period of the data) when the round 3 survey was conducted across years, some data for this round pertain to current year and the remaining part to the previous or the following year: Round 5/3 was conducted across years 2001 and 2002 (Panel 6) and Round 1/3 was conducted across years 2002 and 2003 (Panel 7). The proportion for each year was not assigned and left to the analyst to determine. I split the data for round 3 equally between the two years to calculate the number of disability days. For all subsequent years, the data collected each round pertain to the current year because there was a change in the questionnaire in 2003. Another variable that receives special treatment is the variable measuring levels of education. This variable was numeric in some years and categorical in other years. To make the variable consistent and suitable for the analysis, I re-define it as numerical if the original one is categorical by assigning a number measuring years of education to each category. This number is determined according to the US education system.

Certain spousal variables are also defined, including Spouse Offered, Spousal firm size, and spousal union status. Spouse Offered is a binary variable indicating whether the spouse was offered health insurance through his/her own employer or union. Spousal firm size is redefined as a binary variable that equals 1 if the number of workers the firm hired is at least 100 and 0 otherwise. There are two reasons for using 100 as the cutoff. First, firms with 100 or more employees consistently differ from those with fewer than 100 workers in terms of offer rates (the ratio of the number of firms offering health insurance to the total number of firms) (Kaiser Employer Health Benefits Survey 2015). Second, using this number as a cutoff can separate the sample into two reasonably sized parts. Spousal Union Status is a binary variable indicating whether the spouse was unionized or not.

5.2.3. Exclusion/Inclusion Decisions

Exclusion/inclusion decisions depend on the variables involved, and were made as the data selection process went on. The goal is to include as many observations as possible so that as little information is lost as possible, while at the same time a sample that works well enough can be obtained. To attain this goal, a number of exclusion/inclusion criteria were imposed.

First, only married people with their spouses in the household are included in the data file because certain spousal variables are needed as instruments for the identification of the econometric models to be analyzed in subsequent chapters.

Second, only full-time employed persons (self-employed individuals are automatically excluded when certain exclusion criteria such as positive hourly wage are applied because for self-employed persons, the question on hourly wage was not asked so the information on wage is missing), including their spouses, are included in the sample. While full-time employment is not required for spouses because their variables will be used as instruments only, it is desirable to impose the full-time employment restriction as an exclusion/inclusion criterion, since doing so will result in a cleaner dataset that has fewer missing values for the instrumental variables (simply imposing the full-time employment restriction may not eliminate all missing values for the instrumental variables).

Third, observations with missing values for many of these variables are excluded. Such variables include demographics such as sex, age, race, employment-related variables such as firm size, industry and occupation codes, union status, and insurance-related variables such as insurance status (insured or not), insurance category (public insurance, private insurance, or no insurance), employer-sponsored insurance status indicator, and disability days variables.

Last, the (nominal) hourly wage was restricted to be at least 2.00 dollars per hour. Admittedly, the 2-dollar per hour threshold is arbitrary (Olson (2002) also used 2 dollars per hour as an inclusion criterion), but it seems to be reasonable. Real wage is also defined by deflating the nominal wage using the consumer price index (CPI) for the period.

Applying these inclusion criteria to Sample1 yields a sample that will be used throughout this and the next two chapters (henceforth the **modeling sample**). A number of other sample indicators have also been created by applying other inclusion criteria.

5.3. Summary Statistics

The variables of interest include but are not limited to hourly wage, disability days variables (including the number of days missed work due to sickness and the number of days missed work to care for others), demographic variables, firm characteristics, and insurance-related variables. The firm size is re-defined as a binary dummy variable that equals 1 if the firm has at least 100 employees and 0 otherwise. The insurance offer indicator (called Offered in the tables) is a dummy variable that equal 1 if the person was offered health insurance by his or her employer or union and 0 otherwise. Insurance-related variables include insurance offer indicator, insurance status indicator, and insured-via-firm indicator. The insurance status indicator (Insured) is a binary dummy variable that equals 1 if the person was insured and 0 otherwise. The insured-via-firm variable (Insured-via-firm) is a binary dummy variable that equals 1 if the person was insured through his or her own employer and 0 otherwise.

Other variables such as industry codes, occupation codes, region dummies, and year dummies are also relevant for the analysis but are not included in the summary statistics tables because these variables are categorical and have too many attributes, and these attributes are not important to the analysis. The variables are grouped into the following categories: Dependent variables, Demographics (which may also be the independent variables), Independent Variables, and Instruments, as shown in tables that follow.

Table 5.1. Summary Statistics for the Full Sample

	Peron-years	Mean	Std. Dev	Minimum	**Maximum**
Dependent Variables					
Hourly Wage (Real)	59,442	17.29	11.45	1.52**	76.96*
#Days Missed Work Due to Sickness	59,442	2.80	6.64	0.00	32.00*
#Days Missed Work for Other Reasons	59,442	0.85	2.66	0.00	30.00*
Demographics					
Age	59,442	42.17	10.53	18.00	64.00
Years of Education	59,442	13.45	2.81	0.00	17.00
Male	59,442	0.482	0.500	0.00	1.00
Race	59,442	-	-	-	-
White	41,978	0.706	-	-	-
Black	10,618	0.179	-	-	-
Other	6,846	0.115	-	-	-
Independent Variables					
Offered Insurance	59,442	0.692	0.462	0.00	1.00
Insured	59,442	0.886	0.317	0.00	1.00
Insured via firm	59,442	0.559	0.496	0.00	1.00
Firm with >= 100 Employees	59,442	0.404	0.491	0.00	1.00
Unionized	59,442	0.133	0.340	0.00	1.00
Instruments	59,442				
Spouse Offered Insurance	59,442	0.627	0.483	0.00	1.00
Spousal Firm with >= 100 Employees	59,442	0.367	0.482	0.00	1.00
Spouse Unionized	59,442	0.121	0.326	0.00	1.00

Note: 1) Data: Sample where both the persons and their spouses were employed full-time and all the variables (except for the spousal variables) have non-missing values.
2) *Top-coded.
3) **Nominal is floored at 2 and CPI for some years is greater than 100% so the minimum real wage is smaller than 2.

Table 5.1 shows summary statistics for the entire sample (modeling sample). The statistics presented in the table show that people in the sample earn about 17.29 dollars per hour on average. The average number of days missed work due to illness or injury is 2.80, with a standard deviation 6.6 and the average number of days missed work for other reasons is 0.85, with a standard deviation of 2.66. Note that these two variables are top-coded at the 98.5 percentile. The average age of the persons in the sample is 42 years, with a standard deviation of 10 years.

On average, people in the sample have 13.4 years of schooling. 69.2% of the persons were offered health insurance; in contrast, the offer rate for the spouses is 62.7%. 88.6% of the workers were insured, and 55.9% of the insured workers were insured through their employers or unions. Roughly 40% of the persons worked in firms with at least 100 employees. The unionization rates for both the persons and the spouses are 13.3% and 12.1%, respectively, which are above the national average of 11% (BLS).

Statistics presented in Table 5.2 are for the Husband and Wife samples, each of which is a subset of the sample used for Table 5.1. Panel A presents statistics for the Husband sample and Panel 2B for the Wife sample. Husbands earned 19.29 dollars per hour and the wives earned 15.43 dollars. About 63.4% of the Husbands and 48.9% of the wives were health-insured through their employers or unions. The unionization rates for Husbands and Wives are 15.3% and 11.5%, respectively.

Table 5.2. Summary Statistics for Husband & Wife Samples

	Person-years	Mean	Std. Dev	Minimum	Maximum
Panel A: Husband					
Dependent Variables					
Hourly Wage	28,673	19.29	12.24	1.52	76.96*
#Days Missed Work Due to Sickness	28,673	2.22	5.81	0.00	32.00*
#Days Missed Work for Other Reasons	28,673	0.68	2.33	0.00	30.00*
Demographics					
Age	28,673	42.89	10.48	18.00	64.00
Years of Education	28,673	13.30	2.85	0.00	17.00
Race	28,673	-	-	-	-
White	20,173	0.703	-	-	-
Black	5,296	0.185	-	-	-
Other	3,204	0.112	-	-	-
Independent Variables					
Offered (=1)	28,673	0.746	0.435	0.00	1.00

Insured	28,673	0.879	0.326	0.00	1.00
Insured via firm	28,673	0.634	0.482	0.00	1.00
Firm with >= 100 Employees	28,673	0.421	0.494	0.00	1.00
Unionized	28,673	0.153	0.360	0.00	1.00
Instruments					
Spouse Offered Insurance	28,673	0.588	0.492	0.00	1.00
Spousal Firm with >=100 Employees	28,673	0.363	0.481	0.00	1.00
Spouse Unionized	28,673	0.106	0.308	0.00	1.00
Panel B: Wife					
Dependent Variables					
Hourly Wage	30,769	15.43	10.32	1.52	76.96
#Days Missed Work Due to Sickness	30,769	3.33	7.28	0.00	32.00
#Days Missed Work for Other Reasons	30,769	1.02	2.92	0.00	30.00
Demographics					
Age	30,769	41.51	10.53	18.00	64.00
Years of Education	30,769	13.58	2.77	0.00	17.00
Race	30,769	-	-	-	-
White	21,805	0.719		-	-
Black	5,322	0.173		-	-
Other	3,642	0.118	-	-	-
Independent Variables					
Offered Insurance	30,769	0.642	0.479	0.00	1.00
Insured	30,769	0.893	0.309	0.00	1.00
Insured via firm	30,769	0.489	0.500	0.00	1.00
Firm with >=100 Employees	30,769	0.387	0.487	0.00	1.00
Unionized	30,769	0.115	0.319	0.00	1.00
Instruments					
Spouse Offered Insurance	30,769	0.665	0.472	0.00	1.00
Spousal Firm with >= 100 Employees	30,769	0.370	0.483	0.00	1.00
Spouse Unionized	30,769	0.135	0.342	0.00	1.000

Note: 1) * Top-coded.
2) Only the real wage is presented.

These statistics demonstrate that Husbands earned 25% more per hour than wives, were more likely to be insured through their own employers, more likely to become union members, but less likely to be insured in general than Wives. Note that the statistics for Husbands and for Wives are not symmetric; for example, in the Husband data, 42.1% of the persons and 36.3% of their spouses worked in firms with at least 100 employees, but in the Wife data, the corresponding numbers are 38.7% and 37.0%, respectively. This asymmetry does not constitute an issue in and of itself because the samples are subsets of the entire dataset and some husbands may have been excluded from the samples because they are 65 years or older, but the wives are generally younger than their husbands and so they may still be included in the sample. In addition, sample sizes differ.

5.4. Correlations between Endogenous Variables and Instruments

The endogenous variables of interest include a through-firm insurance offer dummy and a through-firm insurance status indicator. The insurance offer indicator shows whether the person was offered health insurance through his or her employer or union (henceforth Offered). The through-firm insurance status indicator shows whether the person was insured through his or her employer or union (henceforth Insured via Firm). Variables that can potentially serve as instruments for the endogenous variables include a binary variable indicating whether the spouse was offered health insurance through his or her job, an indicator showing whether he or she worked at a firm with at least 100 employees, and a dummy indicating whether the spouse was unionized.

The spousal insurance offer variable can be used as an instrument because a person would become less likely to be insured through his or her own firm if his or spouse is employed at a firm that offers health insurance. Spousal firm size is a potential instrument because large firms are more likely than small firms to offer health insurance. The reason for using spousal union status as an instrument candidate is that ‘the greater eight is given to intra-marginal worker in the union firm relative to non-union firm which produces a greater preference among unionized workers for fringe benefits compared to wages” (Olson 2002). There is a reason for including the spousal insurance offer variable as an instrument even if the other variables are included: some firms (e.g., small firms) may offer individual insurance plans but some other firms (e.g., large firms) may offer family plans. So spousal firm size does not capture this but the spousal insurance offer variable does.

While it is arguable as to whether the above spousal variables are exogenous, using these variables as instruments can yield estimates that bound the true value of coefficient of interest. Olson 2002 suggests that using spousal firm size and union status generates an estimate that bounds the true value from below (lower bound) and using the spousal health insurance indicator generates an upper bound for the true value.

To find out whether the spousal variables can be used as instruments for the endogenous variables, I have created crosstabs using the data set used to generate the summary statistics tables above. In these crosstabs, the own variables are displayed on the rows, labeled as X’s, and the spousal variables are presented on the columns, labeled as Z’s. Since X is binary, the conditional expected value of X given Z, $E[X|Z]$, is the conditional probability of X given Z, and the difference between the conditional probabilities of X given Z, $E[X|Z=1]-E[X|Z=0]$, is the variation induced by Z in the probabilities of X, which measures the impact of shifting Z from 0 to 1 on the probability of X given Z. The Chi-square statistic, which measures the significance of association between two variables, and the Phi coefficient, which measures the degree of association between two binary variables, are also presented.

Table 5.3 provides statistics on the relationship between the variable indicating whether the person was offered health insurance through his or her job and the variable indicating whether their spouses was offered health insurance. The small Chi-Square statistic and the large corresponding probability (0.711) indicate that the relationship between the two variables Offered and Spouse Offered is insignificant. The Phi coefficient is also small and negative, implying a numerically tiny inverse relationship between the two variables.

Table 5.3. Person Offered Insurance vs Spouse Offered Insurance

Count (%)	Spouse Offered (Z)		
Offered (X)	No (=0)	Yes (=1)	Total
No (=0)	6,792	11,495	18,287
	11.43	19.34	30.76
Yes (=1)	15,351	25,804	41,155
	25.83	43.41	69.24
Total	22,143	37,299	59,442
	37.25	62.75	100.00
E[X|Z]	0.693	0.691	0.692
E[X|Z=1]-E[X|Z=0]		-0.002	-
Chi-Square (1)	0.138		Prob.=0.711
Phi Coefficient	-0.002		-

The probability that the person was offered health insurance is 0.691 if the spouse was also offered health insurance and is 0.693 otherwise. The difference of -0.002 between the two probabilities represents the variation induced by the variable Spouse Offered. This small number confirms the conclusion drawn in the preceding paragraph. Hence, whether the spouse was offered health insurance does not provide much information about whether the person was offered health insurance. Thus the Spouse Offered variable may not be a good instrument for the endogenous variable Offered.

Table 5.4. Person Offered Insurance vs Spousal Firm Size

Count (%)	Spousal Firm Size >=100 (Z)		
Offered (X)	No (=0)	Yes (=1)	Total
No (=0)	11,717	6,570	18,287
	19.71	11.05	30.76
Yes (=1)	25,925	15,230	41,155
	43.61	25.62	69.24
Total	37,642	21,800	59,442
	63.33	36.67	100.00
E[X|Z]	0.689	0.699	0.692
E[X|Z=1]-E[X|Z=0]	0.010		-
Chi-Square (1)	6.350		Prob.=0.012
Phi Coefficient	0.010		-

Table 5.4 displays statistics about the relationship between the variable indicating whether the person was offered health insurance through his or her employer or union and a spousal firm size dummy indicating whether the spouse worked at firm with at least 100 employees. The large Chi-Square statistic (or the small p-value) and a significant Phi coefficient of 0.010 tell that there is a small, positive, yet significant association between the two variables. The probability that the person was offered health insurance is 0.699 if the spouse worked at a firm with at least 100 employees and is 0.689 otherwise. The variation induced or explained by the Spousal Firm Size dummy in the probabilities that the person was offered health insurance is of the same magnitude as the Phi coefficient, which is 0.010.

Table 5.5 presents information on the relationship between the Offered variable and the spousal union status indicator. The statistics show a numerically small but statistically significant relationship between the two variables. The probability that the persons were offered health insurance by their firms or unions is 0.675 if their spouses were unionized and is 0.693 when they were not. The -0.020 difference reflects the variation induced or explained by the variable Spouse Unionized variable in the probabilities that the persons were offered health insurance.

Table 5.5. Person Offered Insurance vs Spousal Union Status

Count (%)	Spouse Unionized (Z)		
Offered (X)	No (=0)	Yes (=1)	Total
No (=0)	15,945	2,342	18,287
	26.82	3.94	30.76
Yes (=1)	36,300	4,855	41,155
	61.07	8.17	69.24
Total	52,245	7197	59,442
	87.89	12.11	100.00
E[X|Z]	0.695	0.675	0.692
E[X|Z=1]-E[X|Z=0]	-0.020		-
Chi-Square (1)	12.137		Prob.=0.0005
Phi Coefficient	-0.014		-

Table 5.6. Person Insured via Firm vs Spouse Offered Insurance

Count (%)	Spouse Offered (Z)		
Insured via Firm (X)	No (=0)	Yes (=1)	Total
No (=0)	8,168	18,038	26,206
	13.74	30.35	44.09
Yes (=1)	13,975	19,261	33,236
	23.51	32.40	55.91
Total	22,143	37,299	59,442
	37.25	62.75	100.00
E[X|Z]	0.631	0.516	0.559
E[X|Z=1]-E[X|Z=0]	-0.115		-
Chi-Square	741.950		Prob.=0.0001
Phi Coefficient	-0.112		-

Table 5.6 provides statistics on the relationship between the variable indicating whether the person was insured through his or her job and the dummy indicator showing whether the spouse was offered health insurance from his or her own employer. The Chi-Square statistic, with a probability of 0.0001, indicates a significant association between the two variables. The Phi coefficient (-0.112) indicates that the relationship is inverse and reasonably sizable. When the spouse was offered health insurance through his or her employer or union, the probability that the person was insured through his or her own job is about 51.6%; otherwise the probability is

63.1%. The difference of -0.115 reflects the variation induced by the variable Spouse Offered. The information shows that the correlation is sizable and significant, implying that the variable Spouse Offered may be a good instrument for the endogenous variable Insured via Firm.

Table 5.7 displays statistics on the association between the variable indicating whether the person was insured through his or her job and dummy showing whether the spouse worked at a firm with at least 100 employees. The probability that the person was insured through his or her own employer or union is 52.8% when the spouse was unionized and is 57.8% otherwise. Thus spousal unionization induces a -0.049 variation in the probability that the person was insured through his or her own job. Phi Coefficient is of similar magnitude. The Chi-Square statistic indicates a significant negative association between the Insured via Firm indicator and the Spousal firm size dummy. Though numerically the correlation is not large, its statistical significance may be a good reason for using the spousal variable as an instrument.

Table 5.7. Person Insured via Firm vs Spousal Firm Size

Count (%)	Spousal Firm Size>=100 (Z)		
Insured via Firm (X)	No (=0)	Yes (=1)	Total
No (=0)	15,914	10,292	26,206
	26.77	17.31	44.09
Yes (=1)	21,728	11,508	33,236
	36.55	19.36	55.91
Total	37,642	21,800	59,442
	63.33	36.67	100.00
E[X|Z]	0.578	0.528	0.559
E[X|Z=1]-E[X|Z=0]	-0.049		-
Chi-Square (1)	136.323		Prob.=0.0001
Phi Coefficient	-0.048		-

Table 5.8 presents statistics on the relationship between whether the persons were insured through their employers or unions and whether their spouses were unionized. Statistics show that 47.6% of the persons whose spouses were unionized and 57.0% of those whose spouses

were not unionized were insured through their own employers or unions and. This shows that the persons were less likely to be insured through their employers if their spouses were unionized than if they were not. The difference is the variation in the probabilities that is induced by the spousal status indicator is -0.094.

Table 5.8. Person Insured via Firm vs Spousal Union Status

Count (%)	Spouse Unionized (Z)		
Insured via Firm (X)	No (=0)	Yes (=1)	Total
No (=0)	22,440	3,766	26,206
	37.75	6.34	44.09
Yes (=1)	29,805	3,431	33,236
	50.14	5.77	55.91
Total	52,245	7197	59,442
	87.89	12.11	100.00
E[X|Z]	0.570	0.476	0.559
E[X|Z=1]-E[X|Z=0]	-0.094		-
Chi-Square (1)	225.682		Prob.=0.0001
Phi Coefficient	-0.062		-

The Chi-Square statistic, together with the corresponding probability, indicates that there is a significant relationship between the two variables Insured via Firm and Spouse unionized. The Phi coefficient is -0.062, which implies the relationship is inverse. The significant (inverse) relationship tells that the variable Spouse Unionized can also be used as an instrument for the endogenous variable.

5.5. Implications for Identifying Strategies

For the empirical models that will be analyzed in subsequent chapters, the dependent variables of interest include (real) hourly wage, the number of days missed work due to illness/injury, and the number of days missed work to care for others. The explanatory variable(s) of interest include the binary variable indicating whether the person was insured through his or her employer or union. This explanatory variable is likely to be endogenous because of selection. First, unhealthy people are more likely than others to accept job offers from firms that provide health insurance and thus are more likely to be insured through the firms (or unions). Second, unhealthy people are also more likely than others to become sick and thus miss work for more days due to illness and more likely to choose to be insured (via the firms). Third, wages and insurance policy tend to be offered together to persons with certain characteristics. The existence of endogeneity in the explanatory variable will not only bias the estimator; more damagingly, it may establish incorrect causal connections and thus lead to erroneous inference. Instrumental variables can be used to address this endogeneity issue.

The general idea about finding instruments is to look at spousal variables. Classic econometric theory dictates that the instruments meet three criteria. Conditional on the control variables, the instruments must be correlated with the endogenous variable(s) to be instrumented and must be orthogonal to the error terms. In addition, they should have a monotone relationship with the endogenous variable(s) (Angrist, Imbens, and Rubin 1996) when other controls are also present. Preliminary investigation seems to provide evidence that the spousal variables meet the first criterion. The intuition is that spouses make decisions jointly and/or they share something in common: people tend to marry those who possess characteristics similar to their own. Whether they also meet the second criterion requires further examination. However, the third criterion, monotonicity, may not be testable[7]. Because the potentially endogenous variable Insured-through-Firm is both employment- and health insurance-related, the strategy is to identify such employment and insurance related spousal variables. Thanks to

[7]A violation of monotonicity here would mean, for example, that because one's spouse is in a union, that one looks for a job with health insurance or takes up an offer of health insurance.

Olson (2002) that uses Husband's insurance coverage through his job, Husband's firm size, and Husband's union status as instruments for his wife's health insurance through her own employer, a few spousal variables that can serve as instruments have been found, including the binary variable indicating whether the spouse was offered health insurance through his or her own employer or union (Spouse Offered), the spousal firm size dummy showing whether the spouse worked at a firm with 100 or more employees (Spousal firm size), and the spousal union status indicator showing whether the spouse was a union member. In this section, the three variables have already been analyzed and discussed to certain extent. In the next chapter, they will be further investigated and validated.

5.6. Comments and Conclusions

The MEPS data files contain useful information for testing the relationship between insurance coverage and absenteeism, which I interpret as a marker of productivity. However, a lot of important information is still lacking; there is no complete information about total insurance premiums because employer contributions are not available, nor is there information on deductibles and on co-pay rates. Another limitation of the data is that some important variables have missing values and so some observations have to be excluded, thus reducing the sample size substantially. Despite these limitations, the MEPS data can meet the overall objectives of the research.

CHAPTER VI

VALIDATION OF INSTRUMENTS THROUGH THE WAGE REGRESSION

6.1. Introduction

In the preceding chapter, several instruments were proposed for the potentially endogenous explanatory variable of interest. These instruments include a binary variable indicating whether the spouse was offered health insurance through his or her employer or union, a binary firm size dummy that equals 1 if the firm the spouse worked for has at least 100 employees and 0 otherwise, and a binary variable indicating whether the spouse was unionized. This chapter presents a wage regression model, where the dependent variable is the logarithm of hourly wage and the explanatory variable of particular interest is a binary indicator telling whether the individuals was insured through his or her employer or union. In this regression equation, the explanatory variable is likely to be endogenous because of selection: firms with certain characteristics tend to offer wage and insurance packages to people with certain characteristics. For example, firm characteristics may include firm size and the industry in which they are operating and individual characteristics may include education, gender, and personality. Unobserved heterogeneity in the firms and in workers drives the firm-worker matching and selection process. Thus, this wage regression model can be used to validate the proposed instruments.

The reason for doing this validation exercise is to check whether the instruments work well in this model; if the instruments work well in the sense that they can yield estimates that are different than the OLS estimates but are consistent with results obtained by other researchers, then these instruments are likely to work in the disability days models to be presented in Chapter VII. On the other hand, if these instruments do not work well in the sense that the resulting estimates have signs that are not different from the signs of the OLS estimates, then they will not work in the models to be presented in the next Chapter. Moreover, if the instruments work well as expected, this would support the assumption that the instruments are valid; otherwise the assumption would not be supported.

Classic econometric theory dictates that valid instruments should be correlated with the instrumented variable but uncorrelated with the error term and that they should have a monotone relationship with the instrumented variable. Empirically, valid instruments should have some degree of strength as measured by both statistical significance and numerical magnitude of their coefficients, and should pass relevant statistical tests including but not limited to the specification tests.

These variables may be reasonably good instruments because they may be correlated with the instrumented variable, given that spouses normally make decisions jointly. In the preceding chapter, some preliminary statistical analysis has shown that each of the three potential instruments, including the spousal health insurance offer indicator, the spousal firm size dummy, and the spousal union status indicator, is to some extent correlated with the endogenous variable to be instrumented. However, whether these variables are correlated with the error term is yet to be investigated.

The theoretical justification for using the insurance and employment related spousal variables as instruments is detailed as follows. First, whether the individuals' spouses are offered health insurance affects the likelihood that the individuals would be insured through their own employers because the individuals' employers expect their employees to be insured through their spouses' employers (Dranove, Spier, and Baker 2000). Thus, either the employers would structure the compensation packages in such a way that their employees would become less likely to take the health insurance offered to them or the employers would become less likely to offer health insurance. Moreover, if the individuals' spouses are offered health insurance, then the individuals themselves would have a chance to obtain insurance as part of the family plans through their spouses' employers, thus becoming less likely to take up insurance offered through their own employers or unions. Whatever the specific causes are, the end result is that the individuals should become less likely to be insured through their own employers or unions if the spouses are offered health insurance, but would become more likely to be insured through their own employers or unions if the spouses are not.

Second, the spouse's firm size is likely to be correlated with the variable indicating whether the person is insured through their own jobs because firm size is associated with the likelihood that the firm offers health insurance. Generally speaking, small firms are less likely to offer health insurance than are larger firms (Brown, Hamilton, and Medoff 1996; Currie and Madrian 1999; Olson 2002). Hence, if the individuals' spouses are working in small firms, they are less likely to obtain health insurance through their own employers. On the other hand, if the individuals' spouses are working in large firms, they are more likely to be insured through their jobs. Therefore, there is an inverse relationship between firm size and the likelihood that the person is insured through his or her employer or union. The reason why small firms are less likely than large ones to offer health insurance may be that handling insurance provision and related issues requires a fixed investment (Cutler and Madrian 1998) and so the per worker average cost for small firms is larger than the per worker average cost for large firms.

Third, the spouse's union status should also be associated with the whether the person is insured through his or her own job because if the individual's spouse is a union member, they are more likely to have insurance through their jobs (or unions) as greater weight is given to the infra-marginal workers in firms with unions relative to workers in firms without unions and thus produces a greater preference among unionized workers for health benefits compared to wages (Olson 2002; Goldstein and Pauly 1976). As a consequence, unionized workers are more likely than non-union workers to be insured through their employers. This hypothesis suggests that there is an inverse relationship between the spousal union status variable and the through-firm insurance indicator. However, there might be other mechanism that drives the relationship between these two variables because the spousal union status variable seems to have a positive correlation with the through-firm insurance status indicator.

Job market selection as described above may affect the resulting estimates. For example, if an individual takes a job with health insurance, his or her spouse will likely take a job without health insurance, or vice versa. For any given individual, a job without health insurance should generally pay a higher wage than a job without health insurance. Thus, the selection would somehow affects the results negatively (biases the estimates away from zero).

While the relationship between the through-firm insurance status variable and one of the three spousal variables is inverse, the positive assortative mating theory, on the other hand, suggests that the correlation should be positive because of selection. Thus, positive assortative mating theory implies that using the spousal variables as instruments would bias the estimates toward zero. The overall net effect may be an empirical issue and entirely depends on the data used. The results from the first-stage regressions using each of the three instruments while controlling for all other factors that impact wages show that the correlation is negative for each of the spousal insurance offer indicator and the spousal firm size.

The remainder of this chapter is structured as follows. Section 2 specifies the econometric models. Section 3 briefly describes the data used for the regression. Section 4 presents and discusses the results obtained by using various approaches, including the ordinary least square regression, the two-stage least square regression, and the generalized methods of moments. Section 5 concludes the chapter.

6.2. Econometric Models

This section presents an econometric model that regresses the logarithm of hourly wage on a binary variable that equals 1 if the individual was insured through his or her own employer or union and 0 otherwise, as well as control variables such as demographics, region dummies, firm size, union status indicator, and the year dummies, whichever are appropriate. Demographic variables include gender, race, (own) education, potential experience, quadratic form of potential experience. Compensating wage theory suggests that workers who are insured through their employers should be paid lower wages than comparable workers who stay uninsured, holding human capital and other variables affecting wages constant (Olson 2002).

The wage regression model is first estimated using the ordinary least square regression (OLS). While the OLS estimate is expected to be biased upward (and positive) because of the endogeneity in the explanatory variable of interest that arises from selection in the labor market, it can serve as a benchmark with which results by using other methodologies can be compared. The OLS model simply takes the form of equation (6.2.1).

The two-stage least square (2SLS) method is then used to identify the model so that the true causal relationship between the dependent variable and the explanatory variable(s) of interest can be established. It is an econometric technique widely used to analyze structural equations. It consists of two stages (hence its name). In the first stage, the endogenous variable is regressed on the all the instruments, including the exogenous variables in the original OLS regression equation and the instruments not in the original OLS equation. In the second stage, the dependent variable of the original OLS regression equation is regressed on the fitted values of the endogenous variable and all other exogenous variables obtained using the estimated values for the parameters in the first stage. The 2SLS consists of equations (6.2.1) and (6.2.2).

The use of the fitted value of the endogenous variable distinguishes the 2SLS method from certain other approaches. This methodology has advantages over the OLS method in that it can identify models where at least one of the explanatory variables is endogenous while the OLS method is unable to do so. Compared to other approaches such as the (generic) generalized method of moments, the 2SLS method is efficient and has a transparent, explicit structure. When compared with instrumental variable (IV) method, it has the advantage that it does not require the number of orthogonality conditions be equal to the number of parameters to be estimated, while the IV approach does. Moreover, going through the two stage procedure can make it clear whether the instruments are valid or not.

The wage regression model using the ordinary least square regression (OLS) can be written as follows.

$$lnW_i = \phi_0 + \phi_1 HIE_i + \boldsymbol{X_i \phi_2} + \varepsilon_i \qquad (6.2.1)$$

The dependent variable $\boldsymbol{ln}W_i$ is the logarithm of Hourly Wage for individual i. ϕ_0 is a constant and ϕ_1 is the parameter of interest. HIE_i is a binary variable that equals 1 if the individual was insured through his or her employer or union and 0 otherwise. $\boldsymbol{X_i}$ is a vector of control variables, which, in the current context, include sex, potential experience, race, marital status, spousal education, region dummies, firm size, year dummies, and union status indicator. $\boldsymbol{\phi_2}$ is a (column) vector of parameters. The OLS estimate of ϕ_1 will be biased upward since HIE is endogenous. HIE is endogenous because: 1) more capable or better skilled workers receive better compensation packages that include health insurance and higher wages, while less capable or less skilled ones are paid lower wages and given no health insurance, and 2) the variables in $\boldsymbol{X_i}$ cannot fully capture the effect of the abilities or skills on compensation because these variables are inadequate or incomplete measures of abilities or skills. Some constituents of abilities or skills may be observable to employers but not to other outsiders such as data collectors because these constituents cannot be well-proxied or well-measured, and are thus not well-documented. Whatever is not captured by the independent variables is left in the error term ε_i, causing the variable HIE_i to be correlated with the error term, and so $\boldsymbol{E}[(HIE_i)\varepsilon_i] \neq \mathbf{0}.$ As a consequence, the estimate obtained from ordinary least square regression is not unbiased and even not consistent.

6.2.1. Identification

Endogeneity in the explanatory variable make the OLS estimator inconsistent and may, more damagingly, lead to erroneous inference. Resolving the issue arising from endogeneity requires the use of instrumental variables (henceforth instruments). For the instruments to work well, they should meet the following criteria. First, they are orthogonal to the error terms. Second, they are correlated with the endogenous variable. The relationship between the endogenous variable and the instruments can be established as follows.

$$HIE_i = \theta_0 + \boldsymbol{Z_{ex,i}\theta_1} + \boldsymbol{X_{c,i}\theta_2} + u_i \tag{6.2.2}$$

where $\boldsymbol{Z_{ex,i}}$ is a (row) vector of the (excluded) variables used as instruments for the endogenous variable HIE_i, $\boldsymbol{\theta_k}$ (where $k = 1,2$) is a column vector, and the set of control variables in $\boldsymbol{X_{c,i}}$ is the same as in Equation (6.2.1). Econometrically, the first criterion for valid instruments as stated in the preceding paragraph requires that

$$\boldsymbol{E}\left[\boldsymbol{Z'_{i,j}}\varepsilon_i\right] = 0 \text{ and } \boldsymbol{E}\left[\boldsymbol{Z'_{i,j}}u_i\right] = 0 \tag{6.2.3}$$

where $\boldsymbol{Z'_{i,j}} = \begin{pmatrix} \boldsymbol{Z'_{ex,i,j}} \\ \boldsymbol{X'_{c,i,j}} \end{pmatrix}$ and j indicates the jth element. Because $\boldsymbol{X_{c,i,j}}$ is assumed to be exogenous, these two conditions can be equivalently expressed as $\boldsymbol{E}\left[Z'_{ex,i,j}\varepsilon_i\right] = 0$ and $\boldsymbol{E}\left[Z'_{ex,i,j}u_i\right] = 0$, respectively, which state that each instrument has to be exogenous. The second criterion states that

$$\boldsymbol{COV}\left[\boldsymbol{Z'_{ex,i}}, HIE_i\right] \neq \boldsymbol{0} \tag{6.2.4}$$

Or equivalently, $\boldsymbol{\theta_1} \neq \boldsymbol{0}$**,** where $\boldsymbol{0}$ is a zero column vector with the same number of elements as in $\boldsymbol{Z'_{ex,i}}$.

To further elaborate on the criteria for determining valid instruments, define $\boldsymbol{D} = \boldsymbol{HIE}$ and $\boldsymbol{Z'_i} = \begin{pmatrix} \boldsymbol{Z'_{ex,i}} \\ \boldsymbol{X'_{c,i}} \end{pmatrix}$**.** Further, let $\boldsymbol{Y} = \begin{pmatrix} \boldsymbol{y_1} \\ \cdots \\ \boldsymbol{y_N} \end{pmatrix}$**,** $\boldsymbol{Z_{ex}} = \begin{pmatrix} \boldsymbol{Z_{ex,1}} \\ \cdots \\ \boldsymbol{Z_{ex,N}} \end{pmatrix}$**, and** $\boldsymbol{Z} = \begin{pmatrix} \boldsymbol{Z_1} \\ \cdots \\ \boldsymbol{Z_N} \end{pmatrix}$**.** Note that the set of instruments for the endogenous variable HIE enters Equation (6.2.2) and does not enter the main equation, Equation (6.2.1). This fact, coupled with the condition that the correlation between $\boldsymbol{Z_{ex}}$ and $\boldsymbol{\varepsilon}$ is zero, captures the notion of excludability, which can be more rigorously defined, in the spirit of Angrist, Imbens, and Rubin (1996), as follows.

The first is the exclusion criterion. The set of instruments $\boldsymbol{Z_{ex}}$ is excludable if

$$\boldsymbol{Y}(\boldsymbol{D}, \boldsymbol{Z_{ex}}) = \boldsymbol{Y}(\boldsymbol{D}, \boldsymbol{Z''_{ex}}) \quad \forall\, \boldsymbol{D} \text{ and } \forall\, \boldsymbol{Z_{ex}}, \boldsymbol{Z''_{ex}},$$

where $\boldsymbol{D}$ contains the endogenous variable and $\boldsymbol{Z''_{ex}}$ contains the same set of instruments as, but takes a different set of values than, does $\boldsymbol{Z_{ex}}$. Suppose $\boldsymbol{Z_{ex}}$ has only one element and so is a single instrument; suppose further it takes only two values, 0 and 1. Then this definition implies $\boldsymbol{Y}(\boldsymbol{D}, 1) = \boldsymbol{Y}(\boldsymbol{D}, 0)$**.** This states that the instruments have no direct effect on the outcome variable**.** Therefore, $\boldsymbol{Y}$ can be defined as a function of D alone:

$$\boldsymbol{Y}(\boldsymbol{D}) = \boldsymbol{Y}(\boldsymbol{D}, \boldsymbol{Z_{ex}}) = \boldsymbol{Y}(\boldsymbol{D}, \boldsymbol{Z''_{ex}}) \quad \forall\, \boldsymbol{D} \text{ and } \forall\, \boldsymbol{Z_{ex}}, \boldsymbol{Z''_{ex}}.$$

An instrument must meet this criterion in order to be valid, for otherwise the model represented by Equation (6.2.1) may not be fully identifiable because of the impact of the instruments on the outcome variable.

The second is the nonzero correlation criterion. This is simply what is stated earlier. A stronger version is that the set of instruments has a nonzero causal effect on the endogenous variable. This condition is nice to have but is not necessary because of the orthogonality or exclusion condition. The requirement that the instruments be orthogonal to the error terms rules out reverse causality whatsoever, for otherwise there would be endogeneity in Equation (6.2.2).

The exclusion condition, plus the nonzero correlation restriction, captures the notion that any effect of instrumental variables on the outcome variable must be through its effect on the endogenous variable instrumented by them.

While monotonicity is hard to define in econometric terms, Angrist, Imbens, and Rubin (1996), among others, have provided a way to do this.

The third is the monotonicity criterion. The endogenous variable is monotone in any of the instruments if

$$D_i(Z_{ex,i,j}, \boldsymbol{X_{c,i}}) \geq D_i(Z_{ex,k,j}, \boldsymbol{X_{c,i}}) \text{ for } Z_{ex,i,j} \geq Z_{ex,k,j} \text{ and for all } i, k = 1, \dots, N.$$

Monotonicity is well defined only for a single variable; this means all other variables should be held constant. Strong monotonicity requires the condition to hold with inequality. It is necessary to have strong monotonicity for at least some observations. If each of the instruments takes values 0 and 1, then the above condition is equivalent to

$$D_i(1) \geq D_i(0) \; \forall \; i$$

The last but not yet explicitly stated requirement for the instruments to be valid is stable unit treatment value assumption (Angrist, Imbens, and Rubin 1996), which essentially states that the choices by individuals do not interact with one another.

As proposed in the preceding chapter, potential variables that can serve as instruments for the potentially endogenous variable HIE_i include a binary variable indicating whether the spouses were offered insurance through their employers or unions (Spousal Insurance Offer Indicator), spousal firm size, and spousal union status. An important purpose of running this wage regression is to assess whether these potential instruments are valid and how strong they are.

The statistics presented in Chapter 5 show that the correlation between the through-firm insurance status indicator and any of the three spousal variables is negative. However, a positive correlation is not necessarily surprising because there might be other mechanism that drives the relationship between these two variables. For example, the positive assortative mating theory suggests that the correlation should be positive because people tend to marry those with similar characteristics and thus share similar preferences; this implies that they both may demand health insurance through their respective employers. The overall net effect may be an empirical issue and entirely depends on the data used.

If employer-sponsored health insurance does not alter worker productivity, then the estimate of the coefficient of the instrumented variable HIE_i should reflect a compensating wage differential: since the OLS estimate is normally positive, then the estimate obtained from the regression using the instruments will be smaller in magnitude if it is still positive, or it will even be negative. Some prior studies on the impact of employer-provided health insurance on wage have found evidence of wage offsets. Using CPS and BLS data, Gruber and Krueger (1991) find that a substantial portion of the cost of workers' compensation insurance is shifted to employees in the form of lower wages. Working with CPS and SIPP data, Cutler and Madrian (1998) find that employer-provided health insurance lowers wages, Olson (2002) found evidence that women with health insurance provided by their own employers accept a wage around 20% lower than they would while working full time in jobs without such benefits. Baicker and Chandra (2005) estimate that a 10 percent increase in health insurance premiums reduces wages by 2.3 percent. In all these cases, the coefficient of the through-firm insurance indicator is negative.

There are, however, instances where employer-sponsored health insurance does not affect wages paid to workers who accept such insurance coverage. Using CEX and MEPS data and regressing changes in wages on changes in health insurance, Levy and Feldman (2001) failed to find evidence confirming wage offsets: their estimates for the coefficients are either insignificant or wrong-signed, though they interpret the results not as evidence against what theory would predict, but rather as evidence that changes in health insurance are not exogenous. There are, of course, other reasons that the estimate is non-negative. One such reason is that employers do not necessarily lower wages for workers who accept employer-sponsored health insurance because they can lower other benefits if they offer such benefits. Another such reason is they can shift the premium costs to workers by increasing work intensity. Still another reason is that employer-sponsored health insurance may improve worker productivity. In all these three cases, a non-negative value does not necessarily invalidate the instruments used, because the estimate of the coefficient of the instrumented variable HIE_i could be positive or close to zero even if the instruments are truly valid and strong. But conversely, if the coefficient is close to zero or positive, it is not clear whether this is because the instruments are not good enough or because employer-sponsored health insurance improves productivity or because of other reasons.

Although a non-negative estimate for the health insurance variable is explainable, this dissertation favors the wage offsets hypothesis; that is, employer-sponsored health insurance tends to reduce wages paid to workers who accept employer-provided insurance. The resulting estimates obtained from the model presented in this chapter are unambiguously negative.

There may be potential issues with the spousal variables to be used as instruments. The argument by Olson (2002) suggests that positive assortative mating between spouses would generate a positive correlation between HIE_i and ε_i, and thus may also cause spousal firm size or spousal union status to be correlated with ε_i. Thus, positive assortative mating will generate an IV estimate of ϕ_1 that is biased away from zero. For example, individuals with low unobserved ability are likely to have spouses with low unobserved ability; thus both the husbands and the wives are likely to work for smaller or nonunion firms that pay lower wages and do not offer health insurance. On the other hand, the substitution effect of spousal variables on the through-firm insurance status variable HIE_i works in the opposite direction. So the net effect on HIE_i is ambiguous.

6.2.2. Hypotheses

The standard compensating wage theory predicts that the coefficient of the through-firm insurance status indicator is negative. On the basis of the analyses conducted in the preceding paragraphs, a natural hypothesis can be proposed that employer-sponsored health insurance does not affect wage. Expressed in other words, the null hypothesis is that the coefficient of the variable HIE is not significantly different from zero and the alternative hypothesis is that the coefficient of the variable is not zero, or symbolically**,**

$H_0: \phi_1 = 0$ VS $H_1: \phi_1 \neq 0$.

While the alternative hypothesis says that the coefficient may be positive or negative, a negative coefficient is desired because of the existence of wage offsets. Results will be presented in Section 4.

6.3. Data

The MEPS data will be used for the wage regression. The sample is obtained by imposing the restrictions that individuals worked full time and were at least 18 years old but younger than 65 years and that all variables of interest except for the spousal ones, have no missing values. The resulting sample contains 59,442 observations. The variables of interest include ones that will be used as the dependent, independent, and instrumental variables in the wage regression model. The dependent variable to be used in the wage regression equation is the logarithm of the (real) hourly wage. For details, see the summary statistics tables in Chapter 5.

6.4. Model Results

Each combination of the instruments is checked to see how it would work. Tables that follow show the results using each such combination. The OLS and 2SLS estimates (including first stage and second stage results) are displayed in the same table. Table 6.1 shows results for the wage regression models with one excluded instrument, Table 6.2 lists results obtained by using two excluded instruments, and results for models with all three excluded instruments are presented in Table A6.3. To further check how well the instruments would work for different data sets, the sample is then separated into two sub-sets, one for Husbands and the other for Wives. The results are shown in Table 6.3. The GMM results are listed in tables in the Appendix and will not be discussed in detail because they are almost identical to the 2SLS estimates.

The OLS estimate of the coefficient of the variable indicating the persons were insured through their own employers or unions is statistically significant and positive. This states that workers who were insured through their employers or unions were paid a higher wage on average. But standard compensating wage theory predicts that having health insurance through firms would result in an offsetting decrease in wages since normally firms would 'pay' part of its costs by reducing the wages supposed to be paid to its workers who have employer-sponsored health insurance. Hence, the coefficient should be negative. If the theory is applicable here, then this estimate is biased upward. The analysis in Section 6.2 tells that this biased estimate is a result of endogeneity in the variable. Using the two-stage least squares (2SLS) method can potentially resolve this issue.

Table 6.1. Results from the Wage Regression Model with One Excluded Instrument

VARIABLES	(1) OLS RSE	(2) 1st Stage	(3) 2SLS	(4) 1st Stage	(5) 2SLS	(6) 1st Stage	(7) 2SLS
Insured via firm	0.227***	-	-0.115***	-	-0.179***	-	-0.050
	(0.005)	-	(0.033)	-	(0.040)	-	(0.041)
Male	0.207***	0.124***	0.253***	0.132***	0.261***	0.128***	0.244***
	(0.005)	(0.004)	(0.006)	(0.004)	(0.007)	(0.004)	(0.007)
Firm with >= 100 employees	0.124***	0.208***	0.195***	0.225***	0.209***	0.208***	0.182***
	(0.005)	(0.004)	(0.008)	(0.004)	(0.009)	(0.00392)	(0.009)
Unionized	0.079***	0.226***	0.157***	0.230***	0.172***	0.250***	0.143***
	(0.007)	(0.006)	(0.010)	(0.006)	(0.011)	(0.00576)	(0.011)
<u>Excl. Instruments</u>	-	-	-	-	-	-	-
Spouse offered insurance	-	- 0.128***	-	-	-	-	-
	-	(0.004)	-	-	-	-	-
Spousal Firm w/ >= 100 employees	-	-	-	- 0.110***	-	-	-
	-	-	-	(0.004)	-	-	-
Spouse unionized	-	-	-	-	-	- 0.152***	-
	-	-	-	-	-	(0.006)	-
Constant	0.862***	- 0.0363**	0.830***	- 0.070***	0.824***	- 0.082***	0.836***
	(0.020)	(0.015)	(0.016)	(0.015)	(0.016)	(0.015)	(0.016)
F-stat (Excl. Instruments)	-	447.83	-	433.41	-	428.22	-
Sargan	-	-	-	-	-	-	-
	-	-	-	-	-	-	-
Hausman	-	-	61.71	-	62.81	-	25.13
	-	-	(p=0.0001)	-	(p=0.0001)	-	(p=0.0001)
Observations	59,442	59,442	59,442	59,442	59,442	59,442	59,442
R-squared	0.395	0.153	0.325	0.149	0.296	0.148	0.349

Note: 1) Robust standard errors in parentheses: *** p<0.01, ** p<0.05, * p<0.1
2) Control variables not listed in the table include Education, Race, Potential Experience, Region Dummies, and Year Dummies.
3) Sample: MEPS 2002-2014

The first stage estimates in **Table 6.1** are all significant at the 1% level. The partial F-statistic is relatively large for each of the excluded instruments entering the regression equation one at a time, confirming the significance of each instrument. The Hausman statistic for endogeneity test is significant as the associated p-value is extremely small, implying that the null hypothesis that the variables are exogenous can be rejected at nearly any significance level. These results are consistent with those shown in the cross tabulations presented in Tables 5.6 through **5.8** of Chapter 5.

The OLS estimate for the binary variable Insured via firm is 0.227. The second stage estimate resulting from using the variable Spouse offered and the spousal firm size dummy is -0.115 and -0.179, respectively, each statistically significant at the 1% level. However, the estimate obtained by using the spousal union status dummy as the (excluded) instrument is -0.050, but not statistically significant at the 10% level.

The difference between the OLS estimate and the 2SLS result using the spousal firm size dummy is the 0.404, which is the largest. This difference reflects the shift caused by the instrument Spousal firm size dummy. Using a sample containing wives only, Olson (2002) obtains a shift ranging from 0.40 to 0.50, depending on the specifications. The shift obtained from the model here is very close to the result in the previous work. However, the OLS estimate in Olson (2002) (between 0.1 and 0.2) is much smaller than the result obtained here. One possible explanation for the difference is that the previous uses a different sample. Benchmarked against results from previous work, the spousal firm size dummy works better as an instrument for the endogenous explanatory variable than either of the other two.

Table 6.2. Results from the Wage Regression Model with Two Excluded Instruments

VARIABLES	(1) OLS RSE	(2) 1st Stage	(3) 2SLS	(4) 1st Stage	(5) 2SLS	(6) 1st Stage	(7) 2SLS
Insured via firm	0.227***	-	-0.140***	-	-0.092***	-	-0.121***
	(0.005)	-	(0.029)	-	(0.028)	-	(0.031)
Male	0.207***	0.125***	0.256***	0.121***	0.249***	0.128***	0.253***
	(0.005)	(0.004)	(0.006)	(0.004)	(0.005)	(0.004)	(0.006)
Firm with >= 100 employees	0.124***	0.220***	0.201***	0.209***	0.191***	0.224***	0.197***
	(0.005)	(0.004)	(0.007)	(0.004)	(0.007)	(0.004)	(0.008)
Unionized	0.079***	0.227***	0.163***	0.242***	0.152***	0.248***	0.159***
	(0.007)	(0.006)	(0.009)	(0.006)	(0.009)	(0.006)	(0.009)
Excl. Instruments	-	-	-	-	-	-	-
Spouse offered	-	-	-	-	-	-	-
insurance		0.105***		0.113***			
	-	(0.004)	-	(0.004)	-	-	-
Spousal firm w/ >= 100	-	-	-	-	-	-	-
employees		0.079***				0.097***	
	-	(0.004)	-	-	-	(0.004)	-
Spouse unionized	-	-	-	-	-	-	-
				0.119***		0.129***	
	-	-	-	(0.006)	-	(0.006)	-
Constant	0.862***	-0.030**	0.828***	-0.033**	0.832***	-	0.829***
						0.063***	
	(0.020)	(0.015)	(0.016)	(0.015)	(0.016)	(0.015)	(0.016)
F-stat (Excl. Instruments)	-	446.86	-	448.69	-	438.26	-
Sargan	-	-	2.28	-	1.83	-	6.08
	-	-	(p=0.131)	-	(p=0.176)	-	(p=0.014)
Hausman	-	-	95.62	-	74.19	-	74.74
	-	-	(p=0.0001)	-	(p=0.0001)	-	(p=0.0001)
Observations	59,442	59,442	59,442	59,442	59,442	59,442	59,442
R-squared	0.395	0.158	0.314	0.159	0.334	0.156	0.322

Note: 1) Robust standard errors in parentheses: *** p<0.01, ** p<0.05, * p<0.1
2) Control variables not listed in the table include Education, Race, Potential Experience, Region Dummies, and Year Dummies.
3) Sample: MEPS 2002-2014

Results in Table 6.2 show that the first stage estimate using each pair of the (excluded) instruments is negative and significant. The partial F-statistic for each pair of the excluded instruments is large enough to be significant. For the first pair (the spouse offered insurance indicator and the firm size dummy) or the second (the spouse offered insurance indicator and spousal union status indicator), the Sargan statistic is small (2.28 or 1.83) and thus the corresponding p-value is relatively large (0.131 or 0.176), indicating that the orthogonality conditions hold in either case at any reasonable significance level (say 10% or less). This attests to the validity of each of these two pairs of instruments. However, the Sargan statistic for the last (third) pair (the spousal firm size dummy and spousal union status indicator) is large (6.02) and so the p-value is small (0.014). This large Sargan statistic indicates that, when these two spousal variables are included as instruments, the orthogonality conditions fail to hold and thus raise concern.

The second stage estimates from models with two excluded instruments are all negative and significant. Using the first pair of instruments, the spouse offered insurance indicator and the spousal firm size dummy, yields the largest estimate (in absolute term) (-0.140), and dominates the other two combinations in terms of test statistics and their effects on the estimate for the endogenous variable of interest, though it is smaller in absolute value than the estimate using only the Spouse offered insurance variable alone, as shown in Table 6.2.

Results presented in Table A6.3 in the Appendix show that the three instruments, if all included in the model, fail to pass the specification tests because the Sargan (or Hansen's J) statistic is very large and the associated p-value is small. Thus, the final choice should be made between the variable Spousal firm size alone and a combination of the Spouse offered variable and the spousal firm size dummy. While the one excluded instrument seems to perform better in terms of its ability to shift down the estimate relative to the OLS estimate, it is generally desirable to have more than one instrument, even at some cost, because with more excluded instruments than the instrumented variable, it is possible to test the over-identifying restrictions and obtain test statistic to measure the suitability of the instruments.

Table 6.3. Results from Wage Regression Models Using Husband and Wife Samples. h=Husband; w=Wife.

VARIABLES	(1) OLS RSE (h)	(2) 1st Stage (h)	(3) 2SLS (h)	(4) OLS RSE (w)	(5) 1st Stage (w)	(6) 2SLS (w)
Insured via firm	0.229***	-	-0.057	0.224***	-	- 0.195***
	(0.008)	-	(0.050)	(0.00711)	-	(0.034)
Firm with >= 100 employees	0.0953***	0.219***	0.155***	0.154***	0.221***	0.241***
	(0.008)	(0.006)	(0.012)	(0.007)	(0.006)	(0.009)
Unionized	0.059***	0.224***	0.123***	0.098***	0.230***	0.195***
	(0.009)	(0.008)	(0.014)	(0.010)	(0.009)	(0.012)
Excl. Instruments						
Spouse offered insurance	-	-0.080***	-	-	-0.134***	-
	-	(0.005)	-	-	(0.006)	-
Spousal firm w/ >= 100 employees	-	-0.067***	-	-	-0.081***	-
	-	(0.006)	-	-	(0.006)	-
Constant	1.104***	0.075***	1.112***	0.810***	-0.014	0.775***
	(0.027)	(0.021)	(0.022)	(0.028)	(0.021)	(0.022)
F-stat (Excl. Instruments)	-	207.96	-	-	199.81	-
Sargan	-	-	0.588	-	-	1.273
	-	-	(p=0.443)	-	-	(p=0.260)
Hausman	-	-	17.98	-	-	91.77
Sargan	-	-	(p=0.000)	-	-	(p=0.000)
Observations	28,673	28,673	28,673	30,769	30,769	30,769
R-squared	0.353	0.140	0.303	0.392	0.143	0.280

Note: 1) Robust standard errors in parentheses: *** p<0.01, ** p<0.05, * p<0.1
2) Control variables not listed in the table include Education, Race, Potential Experience, Region Dummies, and Year Dummies.
3) Sample: MEPS 2002-2014

The letter in parenthesis shown in the Table 6.3 indicates the sample used to estimate the models, with h=Husband and w=Wife. The first stage estimates for the two excluded instruments are negative and significant for both samples. Hausman statistics and the corresponding p-values tell that the instrumented variable is endogenous for both the Husband and Wife samples. The p-values associated with the Sargan statistics favor the null hypothesis that the over-identifying restrictions hold for both samples; thus the joint validity of the two excluded instruments is supported. The 2SLS estimates for the endogenous variable differ substantially across these two samples; the Husband sample yields an insignificant result whereas the Wife sample yields a significant, negative estimate which is even slightly larger in absolute value than the one obtained using spousal firm size alone as an (excluded) instrument. The estimate obtained from this model is closer to the results in Olson (2002).

However, the difference between the results using the Husband and Wife samples is too small to be considered significant. This says that when spouses were offered health insurance through their own employers or unions, married men are affected to the same extent by whether their spouses are offered insurance as are married women in terms of whether they are covered by employer-sponsored health insurance. This may be because married men are just as likely as married women to be insured through their spouses' employers or unions when their spouses are offered health insurance.

6.5. Conclusions

The OLS estimate of the through-firm insurance status variable is positive as expected because this variable is endogenous. The endogeneity is a result of labor market selection: workers with low ability and/or skills tend to end up with low wages and no insurance offers from their employers (or unions), whereas workers with higher ability and/or skills are more likely to end up with higher pay and insurance offers from their employers or unions. When the through-firm insurance status variable is instrumented using the spousal insurance offer indicator, spousal firm size, and spousal union status, the estimate changes its sign. This result is consistent with what standard compensating wage theory predicts; it is also confirmed by results in previous studies. The negative sign implies the existence of wage offsets or the substitution between wage and employer-sponsored health insurance. While the numerical magnitude of the estimate obtained when the endogenous variable is instrumented is arguably smaller than that presented in prior studies, this should not be a major concern because prior researchers might have used different samples.

The first stage results for each combination of the three excluded instrumental variables show that correlation exists between any combination of the spousal variables and the endogenous variable. These findings reassure the correlation shown in the cross tabulations presented in Chapter 5. Moreover, Hausman (or GMM C) statistics favor these spousal variables as instruments for the endogenous variable. Sargan (Hansen's J) statistics raise some concern about the validity of one of the three variables as an instrument, but some combinations of the spousal variables seem to be valid instruments.

While the spousal firm size dummy alone performs better than does the combination of two of the three instruments because the former yields an estimate closer to the one obtained in earlier studies such as Olson (2002), using two (or more) instruments in this case is more desirable than using a single one because with more instruments than endogenous variables, more statistics become available for measuring how good the instruments are. Furthermore, including two instruments does not seem to cause additional issues. Therefore, two instruments rather than one should be used in the analysis.

CHAPTER VII

DISABILITY DAYS REGRESSION

7.1. Introduction

The empirical part of the study is intended to explore whether, how, and to what extent employer-sponsored health insurance affects worker productivity. But because information on productivity or its proxies are not readily available, one alternative is to examine the empirical relationship between the disability days variables, especially the number of days a person missed work due to sickness, and a binary explanatory variable indicating whether the individuals was insured through his or her employer (or union). However, this binary explanatory variable is likely to be endogenous because of selection.

First, unobservable factors drive individuals both to obtain insurance through their employers (or unions) and to take more days off due to disability. For example, people with high ability may be able to obtain jobs that offer health insurance and jobs that offer health insurance generally pay higher wages; people who are paid high wages demand more health care and thus take more days off due to sickness. Hence the issue of selection arises. As a consequence, workers insured through their employers would, on average, take more days off due to sickness than workers either uninsured or insured through channels other than their employers.

For the variable measuring the number of days missed work to care for others, the endogeneity issue does not seem to be obvious. A worker does not seem to have an incentive to take days off to care for others solely because he or she has been insured through his or her employer, nor does a worker seem to possess certain characteristics that drive or enable him or her to end up being insured through his or her employer and to demand days off to care for others. If there is any association between how many days the worker would take off to care for others and whether he or she has been insured through his or her employer, then either this is coincidental or there is a factor other than the person's own characteristics that connects the two decisions. In the former case, the variable measuring the number of days missed work to care for others should not be endogenous, but in the latter case, it may be endogenous. For example, if a worker has young children who require care from someone else from time to time, then the worker may choose a firm that offers health insurance plans for which his or her children are eligible; then at some point, the worker takes days off to care for his or her children. This makes this variable endogenous.

The issue of endogeneity has to be addressed because it makes the estimator inconsistent, thus establishing misleading causal connections and leading to erroneous inference. The strategy to do so is to identify one or more variables that can serve as instruments for the endogenous explanatory variable. The econometric model for regressing the number of days missed work for other reasons will be used to do the falsification test on the instruments.

In Chapter 6, several instruments were tested and they seem to be promising. In this chapter, these instruments will be used to identify the econometric models regressing the disability days on the through-firm insurance status indicator. The model of primary interest is the one that regresses the number of days missed due to sickness on the employer-sponsored insurance status indicator. Multiple specifications and methodologies will be used to analyze the relationships between the two variables. The specifications include but are not limited to ordinary linear, Probit, and Poisson models, or a combination thereof. While the suitability of a specification depends on the distributions of the dependent variable of the econometric models under study, ordinary linear specifications generally have advantages over other forms in the sense that they are straightforward and intuitive. Methodologies used to estimate the models include the ordinary least square (OLS), the two-stage least square (2SLS), and the generalized method of moments (GMM). Each of these methodologies has its own advantages over and limitations relative to alternative approaches.

The remainder of this chapter is structured as follows. Section 2 specifies the econometric models. Section 3 briefs the data. Section 4 presents and discusses the results, including estimates obtained by using the ordinary least square regression, the two-stage least square regression, and the generalized methods of moments estimation approaches. Section 5 concludes the chapter.

7.2. Econometric Models

This section presents an econometric model that regresses the variable measuring the number of disability days due to sickness or for other reasons on the variable indicating having insurance through an employer or union. The model includes independent variables as controls, including but not limited to demographics, region, firm size, union status indicator, and the year dummy. Other variables may be added to the models as necessary; these variables may include the binary dummy indicating whether a person was offered paid sick leave days and the ones that measure the number of days the spouses missed work due to sickness or for other reasons.

Estimation methods include the ordinary least square (OLS) regression, two-stage least square, and the generalized methods of moments. The OLS method is the natural one to begin with because it can serve as a benchmark to compare results from other specifications. Specifications include linear and Poisson Dependent variables include the numbers of days a person missed work due to sickness or for other reasons. The explanatory variable of interest is a binary indicator that equals 1 if the individual was insured through his or her own employer (or union) and 0 otherwise. Other independent variables, which serve as controls, include demographics, region dummies, year dummies, etc..

The ordinary least square regression models can be written as

$$y_i = \beta_0 + \beta_1 HIE_i + \boldsymbol{X_i \beta_2} + \epsilon_i \tag{7.2.1}$$

where y_i is the number of *Disability Days (either due to sickness or for other reasons)* for individual i, β_0 constants, β_1 is the parameter of interest, HIE_i is the variable indicating whether the person is insured through his or her employer or union, just as i**n** Equation (6.2.1) in Chapter 6, that is, and $\boldsymbol{X_i}$ may be different from the corresponding term in Equation (6.2.1).

What Equation (7.2.1) is trying to do is to test whether and how having insurance through an employer or union affects absenteeism relative to not having insurance through such a channel. This, however, does not imply that having insurance through an employer or union has different effects on absenteeism than having insurance through from some other source or that having insurance from some other source does not affect absenteeism. While having insurance through an employer or union may affect absenteeism in the same way and to the same extent as does having insurance through other sources affects absenteeism, I choose to focus on what equation (7.2.1) is trying to do.

For the model regressing the number of days a person missed work due to sickness or injuries, the OLS estimate of β_1 is likely to be biased upward because the variable HIE is likely to be endogenous. Endogeneity arises because of (adverse) section. First of all, unhealthy workers are generally more likely than uninsured ones to have themselves insured against potential medical shocks and to demand more medical services. Second, certain workers possess unobservable characteristics that make them more likely than others to end up being insured through their employers or unions. Thus, workers insured through their employers will likely take more days off than other workers who were either uninsured or insured through channels other than their employers or unions. Whatever is not captured by the explanatory variables is left in the error term ϵ_i, causing the variable HIE_i to be correlated with the error term. Hence, HIE_i is endogenous, or $\boldsymbol{E}[(HIE_i)\epsilon_i] \neq \mathbf{0}.$ Therefore, the estimate obtained from ordinary least square regression is not unbiased and even not consistent.

For the model regressing the number of days a person missed work for other reasons (to care for others), the major focus of the analysis is on whether the estimate of β_1 obtained by using the instruments significantly differ from its OLS estimate. If the variable HIE is endogenous, then the estimate using the instruments is likely to be different than the estimate without using the instruments (OLS estimates); if it is not endogenous, then the estimate using instruments is unlikely to differ from the OLS estimate.

The first stage equation for the two stage least square regression is similar to Equation (6.2.2), which can written as

$$HIE_i = \theta_0 + \boldsymbol{Z_{ex,i}\theta_1} + \boldsymbol{X_i\theta_2} + u_i \tag{7.2.2}$$

As mentioned earlier, the control variables in $\boldsymbol{X_i}$ may not be the same as in Equation (6.2.2). If instruments are valid, then $\boldsymbol{\theta_1} \neq \mathbf{0}$, where $\mathbf{0}$ is a zero column vector with the same number of elements as in $\boldsymbol{Z_{ex,i}}$, $\boldsymbol{E}[\boldsymbol{Z_i'}\varepsilon_i] = 0$ and $\boldsymbol{E}[\boldsymbol{Z_i'}u_i] = 0$, where $\boldsymbol{Z_i'} = \begin{pmatrix} \boldsymbol{Z'_{ex,i}} \\ \boldsymbol{X_i'} \end{pmatrix} = \mathbf{0}$. The potential instruments are the spousal variables used in Chapter 6.

Equations (7.2.1) and (7.2.2) constitute the structural equations. Replacing the endogenous variable HIE_i in equation (7.2.1) with the right hand side of equation (7.2.2) yields the reduced form.

$$y_i = \alpha_0 + \alpha_1 \boldsymbol{Z_{sp,i}} + \boldsymbol{X_i \alpha_2} + \boldsymbol{v_i} \tag{7.2.3}$$

where $\alpha_0 = \beta_0 + \beta_1\pi_0$, $\alpha_1 = \beta_1\pi_1$, $\alpha_2 = \beta_2 + \beta_1\pi_2$, and $\boldsymbol{v_i} = \epsilon_i + \beta_1\xi_2$. An apparent limitation of the reduced form regression is that the estimate for the endogenous variable of interest is no longer available. Hence, the empirical results of the reduced form model will not be presented.

7.2.1. Hypotheses

There are two hypotheses to be tested, with one for the main model and the other for the falsification testing model. For the main model, the hypothesis to be tested is that employer-sponsored health insurance does not affect the number of days missed work due to sickness. Put it differently, the null hypothesis is that the coefficient of the variable HIE is not significantly different from zero and the alternative hypothesis is that the coefficient of the variable is not zero. Symbolically,

$$H_0: \beta_1 = 0 \text{ vs } H_1: \beta_1 \neq 0.$$

Although the coefficient can be negative or otherwise, a negative coefficient is desired because such a result is consistent with the theory that employer-provided health insurance boosts worker productivity. For the falsification test model, the major purpose is to check whether the instruments perform in the same way as they do in the main model; it is desired that the estimate of the coefficient of the instrumented variable does not differ from the OLS estimate. Results will be presented in Section 4.

7.3. Data

The dataset used in this chapter is the modeling sample presented in Chapter 5. Many of the independent variables used in the wage regression model will be used her and additional variables may be included as necessary and if appropriate. The dependent variables of interest include the number of days a person missed work due to sickness and the number of days a person missed work for other reasons (For details, refer to the summary statistics tables in Chapter 5). An additional sample containing only childless working adults will be created later to address the potential issue that spousal employment decisions may be endogenous. This sample is obtained by imposing an additional restriction that the individuals have no children[8] and hence it is a subset of the dataset presented in Chapter 5.

7.4. Model Results

Two types of empirical results are presented in this section: main and falsification test. Main results refer to results obtained from the model where the dependent variable is the number of days a person missed work due to sickness. These results are shown in Table 7.1, Table 7.2, Table 7.5, and Table A7.1. Falsification test results refer to results obtained from the model where the dependent variable is the number of days a person missed work for other reasons (to care for others). These results are shown in Table 7.3, Table 7.4, and Table 7.6.

Table 7.1 displays main results obtained using the whole sample**.** The variable of primary interest is shown as Insured Via Firm, which is a binary indicator that equals 1 the worker was insured through his or her own job and 0 otherwise. The OLS results are shown on Column (1).

---

[8]To obtain such a sample, I first include only those individuals who are not explicitly identified as someone else's mother or father; I then restrict the family size to 2 because people who may be someone else's mother or father are put in the same pool as people who have no children.

The estimate for the variable of interest is 0.571 (20% of the sample average of 2.8 days) and statistically significant at the 1% level. This significant positive number tells that a worker who was insured through his or her employer took, on average, about 0.6 (roughly 20% of the sample average) more day off due to sickness than did one who was not insured through his or her own job. However, the positive OLS estimate may reflect the effects of endogneity in the explanatory variable, because the endogeneity arising out of selection and/or moral hazards tend to bias the estimate upward.

Results for the two stages of the 2SLS are listed on Columns (2) and (2') in Table 7.1. First stage estimates for the two selected (excluded) instruments are -0.105 and -0.079, both significant at the 1% level; this provides evidence that these instruments are potentially good in this model. The small Sargan (or J in the case of GMM) statistic (and thus large p-value) favors the null hypothesis that the over-identifying restrictions hold and the Hausman statistic falsifies the null hypothesis.

The estimate for the through-firm insurance status indicator from the second stage of the 2SLS equals -1.456 and is significant at the 1% level. This means that the workers who were insured through their employers or unions took 1.5 fewer day off than those who were not. The GMM estimate for this variable is the same. The Poisson GMM estimate is significant and negative at -0.479, which means that the number of days is reduced by 1.34 days[9] (or 48% of the sample average) as a result of having insurance through the firm or union. The Poisson estimate may be more accurate than the 2SLS or GMM estimate because the relationship between the dependent variable and the independent variable of interest may be nonlinear (convex) and the Poisson

---

[9] This number is equal to the average of 2.80 days times the Poisson GMM estimate.

model can capture the nonlinearity, whereas the linear models cannot and may lead to overestimation.

Table 7.1. Results from Regression of #Days Missed Work Due to Sickness Using the Full Sample

Variables	(1) OLS RSE	(2) 1st Stage	(2') 2SLS	(3) GMM	(4) Pois GMM
Insured via firm	0.571***		-1.456***	-1.451***	-0.479***
	(0.063)		(0.386)	(0.385)	(0.123)
Male	-1.229***	0.125***	-0.959***	-0.959***	-0.357***
	(0.059)	(0.004)	(0.075)	(0.074)	(0.026)
Firm with >=100 employees	0.467***	0.220***	0.888***	0.886***	0.303***
	(0.063)	(0.004)	(0.098)	(0.099)	(0.032)
Unionized	0.817***	0.227***	1.282***	1.282***	0.421***
	(0.102)	(0.006)	(0.121)	(0.127)	(0.040)
Excl. Instruments					
Spouse offered insurance	-	-0.105***	-	-	-
		(0.004)			
Spousal firm with >=100 employees		-0.079***			
		(0.004)			
Constant	4.870***	-0.030**	4.682***	4.681***	1.638***
	(0.228)	(0.015)	(0.216)	(0.217)	(0.072)
F-stat (Excl. Instruments)	-	446.86	-	-	-
Sargan (2SLS)/J (GMM)	-	-	0.257	0.252	0.366
		-	(p=0.612)	(p=0.616)	(p=0.545)
Hausman (2SLS)/C (GMM)	-	-	14.39	29.08	-
	-	-	(p=0.0001)	(p=0.0001)	-
Observations	59,442	59,442	59,442	59,442	59,442
R-squared	0.016	0.158	-	-	-

Note: 1) Robust standard errors in parentheses: *** p<0.01, ** p<0.05, * p<0.1
2) Control variables not listed in the table include Education, Race, Potential Experience, Region Dummies, and Year Dummies.
3) Sample: MEPS 2002-2014.

These negative values seem to suggest that employer-sponsored health insurance helped reduce absenteeism overall. This result is consistent with the hypothesis that employer-provided health insurance improves worker productivity.

Table 7.2 shows results obtained using Husband and Wife Samples, where h=Husband and w=Wife. The dependent variable is again the number of days a worker missed due to sickness. The OLS estimate using the Husband sample (Columns with h) is numerically small at 0.099 and insignificant even at the 10% level. In contrast, the estimate obtained using the Wife sample (Columns with w in Table 7.2) is 0.977 and significant at the 1% level. The positive estimates reflect the effects of endogeneity.

Table 7.2. Results from Regression of #Days Missed Due to Sickness Using Husband and Wife Samples. h=Husband; w=Wife

Variables	(1) OLS RSE (h)	(2) 1st Stage (h)	(2') 2SLS (h)	(3) OLS RSE (w)	(4) 1st Stage (w)	(4') 2SLS (w)
Insured via firm	0.099		-1.405**	0.977***		-1.385***
	(0.080)		(0.604)	(0.094)		(0.499)
Firm with >=100 employees	0.237***	0.219***	0.548***	0.703***	0.221***	1.194***
	(0.080)	(0.006)	(0.144)	(0.097)	(0.006)	(0.135)
Unionized	1.003***	0.224***	1.340***	0.607***	0.230***	1.159***
	(0.130)	(0.008)	(0.167)	(0.159)	(0.009)	(0.178)
Excl. Instruments						
Spouse offered		-0.080***			-0.134***	
		(0.006)			(0.006)	
Spousal Firm w/ >=100 employees		-0.067***			-0.081***	
		(0.006)			(0.006)	
Constant	2.398***	0.075***	2.439***	5.665***	-0.014	5.470***
	(0.281)	(0.021)	(0.270)	(0.341)	(0.021)	(0.325)
F-stat (Excl. Instruments)	-	199.81	-	-	207.96	-
Sargan (2SLS)/J (GMM)	-	-	0.297	-	-	0.904
		-	(p=0.586)			(p=0.342)
Hausman (2SLS)/C (GMM)	-	-	3.19	-	-	11.85
		-	(p=0.041)			(p=0.0001)
Observations	28,673	28,673	28,673	30,769	30,769	30,769
R-squared	0.012	0.138	-	0.017	0.140	-

Note: 1) Robust standard errors in parentheses: *** p<0.01, ** p<0.05, * p<0.1
2) Control variables not listed in the table include Education, Race, Potential Experience, Region Dummies, and Year Dummies.
3) Sample: MEPS 2002-2014.

The instruments pass the over-identifying test and the endogeneity test using both the Husband and the Wife samples. This is evidence that the instruments are valid (or evidence against the contrary). The second stage estimate for the through-firm insurance status variable that is obtained by using the Husband sample is -1.405 (50% of the sample average of 2.8) and significantly different from zero at the 1% level. Compared with the OLS estimate, this represents a 1.4 day reduction. The estimate using the Wife sample is -1.385 (49% of the sample average) and significant at the 1% level. The estimates for both samples are similar in magnitude and so the difference does not seem to be significant.

Results from the model regressing the number of Days Missed Work for Other Reasons on the Insured via firm variable using the whole sample are displayed in Table 7.3. The OLS estimate is positive at 0.175 (20% of the sample average) and statistically significant. When the variable is instrumented, the coefficient becomes negative (at around -0.6 or 70% of the sample average of 0.85), and significant at the 1% level, as shown by 2SLS results listed in Column (2’). The change from the positive OLS to the negative 2SLS estimate may imply that the explanatory variable through-firm insurance status indicator used in this model is endogenous. Column (3) and (4) list the GMM and the Poisson GMM results, respectively. Suppose that the instruments are valid and so are the estimates using instruments. Then the 2SLS estimate says having insurance through the firm would reduce the number days off work to care others by 0.6 day (roughly 70% of the sample average of 0.85) (the Poisson estimate implies 62% or a 0.52 day reduction).

Table 7.3. Results from Regression of #Days Missed Work for Other Reasons Using the Full Sample

Variables	(1) OLS RSE	(2) 1st Stage	(2') 2SLS	(3) GMM	(4) Pois GMM
Insured via firm	0.175***	-	-0.592***	-0.601***	-0.616***
	(0.025)	-	(0.154)	(0.153)	(0.148)
Male	-0.344***	0.125	-0.241***	-0.239***	-0.292***
	(0.024)	(0.004)	(0.030)	(0.030)	(0.034)
Firm with >=100 employees	0.0595**	0.220***	0.219***	0.223***	0.235***
	(0.025)	(0.004)	(0.0393)	(0.039)	(0.039)
Unionized	0.138***	0.227***	0.313***	0.318***	0.342***
	(0.039)	(0.006)	(0.048)	(0.050)	(0.051)
Excl. Instruments					
Spouse offered	-	-0.108***	-	-	-
	-	(0.004)	-	-	-
Spousal firm with >=100 employees	-	-0.074***	-	-	-
	-	(0.004)	-	-	-
Constant	0.570***	-0.030**	0.498***	0.493***	-0.872***
	(0.082)	(0.015)	(0.086)	(0.079)	(0.103)
F-stat (Excl. Instruments)	-	446.86	-	-	-
Sargan (2SLS)/J (GMM)	-	-	7.43	7.27	5.74
	-	-	(0.006)	(P=0.007)	(p=0.017)
Hausman (2SLS)/C (GMM)	-	-	12.84	26.77	-
	-	-	(p=0.0001)	(p=0.0001)	-
Observations	59,442	59,442	59,442	59,442	59,442
R-squared	0.014	0.158	-	-	-

Note: 1) Robust standard errors in parentheses: *** p<0.01, ** p<0.05, * p<0.1
2) Control variables not listed in the table include Education, Race, Potential Experience, Region Dummies, and Year Dummies.
3) Sample: MEPS 2002-2014.

However, the Sargan (Hansen's J in the case of GMM) statistic is too large (so the associated p-value is too small). This raises issues with the instruments: the endogeneity problem is not completely solved by using the instruments.

Table 7.4 displays results obtained using the Husband and Wife samples. For both the Husband and wife samples, the OLS estimates are positive and significant: having health insurance through firms or unions encourages the insurance holders to take days off to care for others. The 2SLS results are also significant but negative for both samples: after the selection issue has been addressed, having health insurance through firms or unions actually discourages the insurance holders to take days off to care for others.

Table 7.4. Results from Regression #Days Missed Work for Other Reasons Using Husband and Wife Samples. h=Husband; w=Wife

Variables	(1) OLS RSE (h)	(2) 1st Stage (h)	(2') 2SLS (h)	(3) OLS RSE (w)	(4) 1st Stage (w)	(4') 2SLS (w)
Insured via firm (yes=1)	0.158***	-	-0.711***	0.194***	-	-0.518***
	(0.030)	-	(0.243)	(0.038)	-	(0.200)
Firm with >=100 employees	0.114***	0.219***	0.294***	0.007	0.228***	0.155***
	(0.032)	(0.006)	(0.058)	(0.0384)	(0.006)	(0.054)
Unionized	0.100**	0.224***	0.295***	0.196***	0.234***	0.362***
	(0.048)	(0.008)	(0.067)	(0.064)	(0.009)	(0.071)
Excl. Instruments						
Spouse offered	-	-0.0804***	-	-	-0.081***	-
	-	(0.006)	-	-	(0.006)	-
Spousal Firm >=100 employees	-	-0.067***	-	-	-0.134***	-
	-	(0.006)	-	-	(0.006)	-
Constant	0.583***	0.075***	0.606***	0.347***	-0.059***	0.288**
	(0.101)	(0.021)	(0.109)	(0.126)	(0.021)	(0.130)
F-stat (Excl. Instruments)		199.81		-	207.96	-
Sargan	-	-	3.29	-	4.21	-
	-	-	(p=0.074)	-	(p=0.040)	-
Hausman	-	-	6.66	-	6.64	-
	-	-	(p=0.001)	-	(p=0.013)	-
Observations	28,673	28,673	28,673	30,769	30,769	30,769
R-squared	0.016	0.157	-	0.010	0.140	-

Note: 1) Robust standard errors in parentheses: *** p<0.01, ** p<0.05, * p<0.1
2) Control variables not listed in the table include Education, Race, Potential Experience, Region Dummies, and Year Dummies.
3) Sample: MEPS 2002-2014.

But again, the Sargan statistics for the two samples suggest that the instruments do not pass the over-identifying restrictions at the 10% level and so they may not be valid in these models. Hence, the estimates using the instruments may not be accurate. If the significance level is limited at 5%, the instruments can pass the test for the over-identifying restrictions when the Husband sample is used, but cannot when the wife sample is used.

7.4.1. Further Testing Using A Sample with Childless Adults

The relatively large Sargan (or Hansen's J) statistics presented in Table 7.3 show that the model fails to pass the over-identifying restrictions. A possible cause for this failure is that the employment decisions by people with children are endogenous. To get around this problem, the analysis presented below utilizes a sample containing married individuals without children. Results obtained using this sample are shown in the tables below.

Table 7.5 presents results from the model where the dependent variable is the number of days a person missed work due to illness or injuries. The OLS estimate for the employer-provided health insurance variable is still positive but no longer significant. This means that, for married people who do not have children, insuring through firms or unions does not affect the number of days they would take due to sickness or injuries. The 2SLS and GMM estimates tell that married workers without children would take 3 fewer days off if insured through firms or unions than otherwise. Given an average of 2.80 days off, the Poisson GMM estimate of -0.898 implies a slightly 2.50 fewer days ((89.8% of the sample average of 2.80). As before, the test statistics such as Sargan (or Hansen's J), Hausman, and the first stage estimates show that there are no problems with the instruments. Compared with the results shown in Table 7.1, these estimates seem to be a bit too large. One possible explanation is that childless people are a special group and may not be the same as people who have children.

Table 7.5. Results from Regression of #Days Missed Work Due to Sickness Using Sample Containing Childless Married Adults

Variables	(1) OLS RSE	(2) 1st Stage	(2') 2SLS	(3) GMM	(4) Pois GMM
Insured via firm	0.157		-2.968***	-2.954***	-0.898***
	(0.133)		(1.062)	(1.068)	(0.324)
Male	-0.911***	0.090***	-0.629***	-0.636***	-0.228***
	(0.136)	(0.009)	(0.157)	(0.162)	(0.055)
Firm with >=100 employees	0.613***	0.208***	1.228***	1.208***	0.400***
	(0.132)	(0.008)	(0.241)	(0.244)	(0.079)
Unionized	1.161***	0.247***	1.941***	1.958***	0.630***
	(0.225)	(0.012)	(0.317)	(0.326)	(0.112)
Excl. Instruments					
Spouse offered	-	-0.065***	-	-	-
	-	(0.009)	-	-	-
Spousal firm with >=100 employees	-	-0.080***	-	-	-
	-	(0.009)	-	-	-
Constant	4.351***	0.168***	4.696***	4.662***	1.464***
	(0.537)	(0.034)	(0.500)	(0.511)	(0.161)
F-stat (Excl. Instruments)	-	69.77	-	-	-
Sargan (2SLS)/J (GMM)	-	-	1.79	1.70	1.84
	-	-	(0.181)	(P=0.192)	(p=0.147)
Hausman (2SLS)/C (GMM)	-	-	4.61	9.03	-
	-	-	(p=0.010)	(p=0.011)	-
Observations	12,161	12,161	12,161	12,161	12,161
R-squared	0.022	0.128	-	-	-

Note: 1) Robust standard errors in parentheses: *** p<0.01, ** p<0.05, * p<0.1
2) Control variables not listed in the table include Education, Race, Potential Experience, Region Dummies,
and Year Dummies.
3) Sample: MEPS 2002-2014.

Table 7.6 displays results from the model where the dependent variable is the number of days a person missed work for other reasons. Most notably, the OLS, 2SLS, GMM, and Poisson GMM estimates are all statistically insignificant! The Sargan (and Hansen's J) statistic shows that the instruments pass the test for the over-identifying restrictions. The Hausman (or C) statistic is small; this can be interpreted as saying that the instrumental variables estimate is consistent with the OLS estimate.

Table 7.6. Results from Regression of #Days Missed Work for Other Reasons Using Sample Containing Childless Married Adults

Variables	(1) OLS RSE	(2) 1st Stage	(2') 2SLS	(3) GMM	(4) Pois GMM
Insured via firm	0.071		-0.123	-0.129	-0.211
	(0.047)		(0.363)	(0.368)	(0.578)
Male	-0.213***	0.090***	-0.196***	-0.193***	-0.344***
	(0.049)	(0.009)	(0.054)	(0.058)	(0.098)
Firm with >=100 employees	0.036	0.208***	0.074	0.071	0.123
	(0.047)	(0.008)	(0.082)	(0.083)	(0.139)
Unionized	0.107	0.247***	0.156	0.163	0.276
	(0.073)	(0.012)	(0.108)	(0.113)	(0.180)
Excl. Instruments					
Spouse offered	-	-0.065***	-	-	-
	-	(0.009)	-	-	-
Spousal firm with >=100 employees	-	-0.081***	-	-	-
	-	(0.009)	-	-	-
Constant	-0.165	0.168***	-0.143	-0.130	-2.088***
	(0.161)	(0.034)	(0.171)	(0.161)	(0.324)
F-stat (Excl. Instruments)	-	69.77	-	-	-
Sargan (2SLS)/J (GMM)	-	-	1.29	1.20	1.06
	-	-	(0.255)	(P=0.273)	(p=0.302)
Hausman (2SLS)/C (GMM)	-	-	0.144	0.278	-
	-	-	(p=0.866)	(p=0.870)	-
Observations	12,161	12,161	12,161	12,161	12,161
R-squared	0.010	0.126	0.008	0.008	-

Note: 1) Robust standard errors in parentheses: *** p<0.01, ** p<0.05, * p<0.1
2) Control variables not listed in the table include Education, Race, Potential Experience, Region Dummies, and Year Dummies.
3) Sample: MEPS 2002-2014.

It is evident by looking at results in Table 7.5 and Table 7.6 that the instruments meet these criteria. Alternatively interpreted, the instruments perform as intended. Thus, the hypothesis that they work universally (across the two models) has been falsified. Passing this falsification test reinforces the belief that the instruments are indeed valid.

7.4.2. Alternative Specifications

Appropriate alternative specifications include the two-part model and the Poisson model. The two-part model is appropriate for situations where two types of decisions are involved; one is a binary decision, which decides whether or not to takes days off, and the other is a quantitative decision, which determines how many days to take off. The context involved in taking days off does not quite fit the requirements for the two-part model, but the distribution of the disability days variables is very similar to the distribution require by the two-part model. In this sense, the two-part model is applicable. To work out the two-part model, define a binary decision variable that equals 1 if the worker took days off and 0 otherwise. Conceptually, after this decision has been made, the worker would decide to take how many days off. Hence, then the **two-part model** can be written as

$$Pr(D = 1) = \mu_0 + \mu_1 HIE_i + \boldsymbol{X_i \mu_2} + \zeta_i \qquad (7.2.1'')$$

$$ln(SicD_i | D = 1) = \gamma_0 + \gamma_1 HIE_i + \boldsymbol{X_i \gamma_2} + \eta_i \qquad (7.2.1''')$$

$D = 1$ is equivalent to stating that the number of days the worker took off is greater than zero. For the variable measuring the number of days missed work for other reasons, the model has a similar form. According to the distribution of the dependent variable, another suitable specification is the zero-inflated generalized Poisson model. Both the first and the second parts of these two-part models are estimated using the same instruments as those used to estimate the models presented earlier.

For the convenience of exposition, the results of the general Poisson (not the two-part) models are displayed in the same tables for results from linear specifications.

Table A7.1 in the Appendix presents Two-Part Model results for the days missed work due to sickness. Results for the variable measuring the number of days a person missed worked for other reasons are not presented. The estimate for the Insured via Firm variable from the first part of the ordinary Probit model is positive (the corresponding result obtained from the first part of the IV Probit model is for reference only because the IV Probit procedure is not appropriate for the model with a binary endogenous variable). For the bivariate model (Bi Prob.), the estimate for the Insured via Firm variable is negative and significant. This result says that having health insurance through the firm or union decreases the likelihood a worker would take days off due to sickness. However, for all other models, including OLS, 2SLS, GMM and Poisson GMM, none of the results for the second part are significant**.** These results (excluding the OLS estimate which does not tell the true causal relationship) together state that having insurance through firms or unions only make workers less likely than otherwise to take days off but does not affect how many days the person would take once the decision has been made. In the present framework where having insurance through firms or unions is assumed to affect productivity, these results say that having health insurance through firms or unions influences productivity only by making workers less likely than otherwise to take sick leave.

7.5. Conclusions

The nonnegative OLS estimate of the coefficient of the Through-firm insurance status indicator is consistent with the assumption that this explanatory variable is endogenous. The negative 2SLS and GMM results show that employer-sponsored health insurance actually reduces the number of days a person missed work due to sickness. This result is consistent with the theoretical prediction that employer-sponsored health insurance may enhance worker productivity because reducing the number of days a person missed work due to sickness is equivalent to increasing productivity, everything else held constant. This conclusion is reinforced by the result from the falsification test that employer-sponsored health insurance also reduces the number of days a person missed work to care for others. These results provide evidence that employer-sponsored health insurance tends to reduce absenteeism in the workplace that would happen either because workers are sick or because they want to take care of others. An important implication of the results from the Husband and Wife samples is that married men do not utilize employer-provided health insurance as much as married women do or, alternatively, men may internalize health benefits to a much larger extent than women.

However, the above conclusions rest on the assumptions that the instruments are valid in the models. The first stage estimates for the instruments and test statistics from the second stage seem to provide evidence that the instruments are valid in certain cases. The first stage estimates for the two instrumental variables are all negative and significant, confirming the theoretical prediction of an inverse relationship. Moreover, the instruments pass the specification tests for some of the specifications. These results provide evidence that the instruments are valid.

The falsification test results using the main sample fail to falsify the assumptions about the instruments because when the two instruments are used in the regression for the variable measuring the number of days a worker missed work for other reasons, the estimate for the target variable differs from the OLS estimate. However, failure to falsify does not invalidate the instruments either, because it is not quite certain whether or not the instrumented variable is endogenous in this equation. Therefore, the instruments and the results obtained using them may still be valid. Further investigation using a sample with childless working adults shows that the instruments perform well in the main model where the dependent variable is the number of days a person missed work due to sickness, but does not do so in the model where the dependent variable is the number of days a person missed work for other reasons. This provides additional evidence for the validity of the instruments.

CHAPTER VIII

SUMMARY AND CONCLUSIONS

8.1. A Brief Review

In Chapter 1, I provided some basic facts about employer-sponsored health insurance in the United States. I then raised the questions of why firms offer health insurance, how firms' insurance offering behavior affects the choice of their employees, and what are the fundamental forces behind firms' behavior. The desire to answer these questions motivates this dissertation. Specifically, I put forward the productivity effect hypothesis as an alternative explanation about why firms offer health insurance. In Chapter2, I reviewed prior theoretical work on this topic; this is the fundamental framework in which the models presented in later chapters are built. I also explored the technical conditions necessary for the desired equilibrium outcomes through numerical analysis. Finally, I pointed out the limitations of this theoretical framework and proposed potential extensions. In Chapter 3, I first built two models within the framework of prior research to study how productivity affects the firm's decision, and then modified this framework to allow workers to remain uninsured. In Chapter 4, I constructed an alternative model to study cases where people face credit constraints.

In the empirical part of the dissertation, I showed in Chapter 5 how the data is constructed and proposed and examined possible instruments for health insurance coverage through one's employer. In Chapter 6, I reviewed prior studies on instrumental variables, validated the proposed instruments through the wage regression, and selected the instrumental variables according to classic criteria. In Chapter 7, I first tested the main models without using instruments and obtained

OLS estimates for the variable representing employer-provided health insurance. Then I included the instruments in the model and obtained results that are different than those obtained without using instruments. I also split the data into the Husband and Wife samples and used these two samples to test the models. Finally, I explored alternative specifications such the two-part model.

8.2. Theory

The theory part of the dissertation has attempted to provide (alternative) answers or at least partial answers to the questions why firms offer health insurance to their employees, why firms of different sizes offer different types or numbers of insurance plans, and why insurance offer rates and/or take-up rates vary across industries and/or occupations. If health insurance is just a cost to the firm or the firm merely serves as a pass-through, then the firm would not have enough incentive to take the trouble of offering health insurance in the first place, let alone the variety of types of insurance plans.

Prior studies, including Miller (2005), have provided explanations. One such explanation is that health insurance premiums are tax-deductible and therefore encourage firms to offer insurance and pay part of the premiums thereof, rather than pay higher wages to those employees who are willing to accept the insurance offers as long as they are not thus made worse off than they otherwise would. There are other explanations as well. For example, the issue associated with informational asymmetry is made less severe if insurance applications are handled through employers than directly by the insurance companies themselves and so an insurer can extract some information rent and share it with the firm; on the other hand, there would be an efficiency gain if the firm handles the insurance application process for all its employees who take up health insurance offered through the firm, rather than let each of these employees do the same individually, and this efficiency gain can somehow be shared between the firm and its employees. While such explanations are reasonable and, theoretically speaking, cannot be refuted entirely, they do not seem to be adequate, because there is still variability in firms' insurance offering behavior even after effects of the aforesaid factors have been accounted for.

The theory presented in the dissertation proposes an alternative explanation. It does so by assuming that health insurance influences worker productivity. This productivity impact hypothesis explains firm's behavior for the part other factors may fail to explain: a firm may enroll more employees than old theories suggest. However, this theory is not intended to refute previous hypotheses, nor should it be interpreted as such. Rather, it is intended to address the same issue from a different, alternative perspective and to complement previous theories. Despite their differences, this theory and others generate consistent predictions, however.

To explore how health insurance may affect productivity, the dissertation proposes two models, as discussed in chapter 3. The first model is one where worker productivity depends on the type of health insurance only and does not change with the number of workers covered under each type of health insurance. The second assumes that worker productivity changes with one of the two insurance types and with the number of workers covered under that type, holding the productivity of the other insurance type fixed. The first can be loosely interpreted as depicting situations where health insurance alters individual worker productivity only. In contrast, the second is more about team productivity than it is about individual productivity. Nevertheless, their similarities outweigh their disparities when each is benchmarked against the base model, that is, Miller's.

In both the two models, the firm's behavior is examined and equilibrium conditions are analyzed. In the first model, productivity plays the role of shifting the extra benefit function (or equivalently, shifting the premium cost curve in the opposite direction) relative to the case without productivity, thus changing the equilibrium outcome. In the second model, on the other hand, productivity not only shifts the extra benefit curve, but also tilts it; yet the equilibrium outcome is affected more by the shifting than by the tilting. As a consequence of the productivity effect, more (fewer) workers would be enrolled in the generous plan than the base model predicts if the effect is positive (negative) and if the resulting equilibrium is separating. Moreover, positive (negative) productivity effect may also increase the likelihood of a high-end (low-end) pooling if the base model equilibrium is separating.

Despite their advantages, all the models, including Miller's, presumes one thing: all workers end up choosing either this or that health insurance plan and thus no one is uninsured. Assuming that every worker is insured is sometimes desirable, but not always so; nor is this assumption realistic. Realizing the inadequacy of these models, the dissertation moves one step forward to relax such a restriction, and proposes an alternative model where workers can choose to be uninsured to study the take-up decisions of workers. What differentiates this model from either of the models described earlier is not the number of plans the firm offers, but the choice that certain (risk-averse) workers may make to remain uninsured. Allowing for the possibility that individuals stay uninsured makes the model more closely mimic the real world; but this is not because people do not like health insurance, but rather because the profit-maximizing decision by the firm pushes them to the point where they would rather leave themselves unprotected against potential risk than be exploited by the firm. This suggests that the welfare distribution may not be desirable, because for risk averse individuals, facing risk is a hard choice, especially when an alternative is available.

The productivity effect that health insurance may bring about and the possibility that people choose to stay uninsured even given an alternative choice are all studied in the framework where wages and health benefits are treated as equivalents for any given type of people. But realistically speaking, they may not always be equivalents. To explore the situations where wage benefits and health benefits are not equivalents, at least to some extent, the dissertation proposes a somewhat different modeling framework where workers are assumed to prefer wage benefits to some degree by requiring a reservation wage. This same framework can, not accidentally, be used to study the case where people face credit or liquidity constraints, for wages can be used to meet alternative needs but health benefits cannot. When the reservation wage constraints are imposed, the equilibrium conditions and so optimal outcomes potentially change; because optimality requires the firm to strike a tradeoff between wage and health insurance offers and because wage offer is bounded by the reservation wage from below, the health insurance offer has to be bounded from above. Hence, the optimal choice for the cutoff level of worker type is bounded above. However, the fundamental results remain the same that productivity effect shifts the optimal decision.

The dissertation has also analyzed the distribution of rents. In the two-plan models, the information rents accrue to workers except for the lowest-cost ones in all types of equilibria, be it separating, low-end pooling, or high-end pooling; so the firm extracts no rents. But efficiency differs between pooling and separating equilibria; any pooling equilibrium, either low-end or high-end, is Pareto-efficient, whereas a separating one is not. A pooling equilibrium is Pareto-efficient because there is no (informational) distortion or no deadweight loss in the allocation process;

when the equilibrium is pooling, the only difference between the full-information and the asymmetric-information cases is in the distribution of rents: the firm would extract all rents and so workers get nothing if information is symmetric or perfect, and the opposite would be true if information is asymmetric. A separating equilibrium is Pareto-inefficient because "separating" or "cutting in the middle" involves a deadweight loss, which results from informational distortion: when the equilibrium is separating, the firm has to pay something similar to a "transfer" cost, which no one receives.

In the one-plan model without reservation constraint, all workers except the lowest-cost ones would take all information rents if the equilibrium is high-end pooling (where all workers are covered) and so the firm extracts nothing; the firm would extract all rents if the equilibrium is low-end pooling, but if the equilibrium is separating, the firm and workers with health cost above the cutoff level would receive rents and all other workers wouldn't. In the model with reservation wage constraint, workers would receive rents even if the equilibrium is low-end pooling because their reservation wage is above their reservation utility, but there is no difference regarding distribution of rents for other types of equilibria. As for efficiency, there is no difference between the one-plan models and the two-plan ones: any pooling equilibrium is Pareto-efficient and any separating is not.

The dissertation also briefly explains how health insurance can improve productivity. It does so by citing arguments and evidence from prior research. First, health insurance may somehow help improve human capital and thus enhance productivity; second, having health insurance ultimately enhances a worker's well-being and morale, thus boosting productivity. Third, having health insurance is associated with improving worker rights and thus, through some mechanism, enhances productivity. Prior empirical studies imply that health insurance has a (dynamic) positive effect on worker productivity.

In the appendices, the dissertation presents an efficiency model to analyze how the firm's optimal decision would be affected and what the equilibrium outcomes would become when wage influences worker productivity. In addition, the dissertation also presents another one-plan model to examine the firm's alternative decision making mechanism. In this model, the firm is assumed to manipulate the proportion of premium cost it bears, rather than wage, to affect the choice of its employees. This model provides a simple but different way to look at firm's decision and to investigate the technical requirements for the benefit function of health insurance. Such technical requirements include but are not limited to the convexity and slope of the health benefit curve. Sensible equilibrium outcomes require that the convexity and slope meet certain conditions. Furthermore, numerical analyses for the Miller model and for the one-plan model as presented in the appendixes have shown that equilibrium outcomes depend on the specific forms of the (excess) benefit of the (generous) insurance plan and on the distribution functions for the worker type.

8.3. Empirics

If health insurance affects productivity, how does it do so empirically? The dissertation has, in its empirics part, provided an answer to this question by using absenteeism at the workplace as a proxy for (negative) productivity. If health insurance influences productivity, it can do so by affecting absenteeism. If it improves productivity, then it would discourage absenteeism; the opposite would hold true if it hinders productivity. Alternatively, if it does have any impact on productivity, then its association with absenteeism would be weak or even nonexistent. While this variable is not a perfect measure of productivity, it should serve the purpose to a considerable extent, for otherwise firms wouldn't care about or would even encourage absenteeism. But then, if health insurance can manifest its impact on productivity in data by influencing absenteeism at the workplace, what does the data tell about their relationship? Finding an answer to this question is the major task of the empirics part of the dissertation.

The dissertation utilizes the MEPS data for its empirical analysis. Why MEPS? First and foremost, this data contains the necessary information for analyzing the empirical relationship between absenteeism at the workplace and employer-sponsored health insurance, but any other publicly available data does not have all the required information. Second, the MEPS data is readily available. Like most other public data sources, the MEPS does not require authorization or approval and so is easy to access at any time. Third, the data is rather clean and well structured, and so does not require much additional work.

To identify the causal relationship between the productivity proxy variable and the employer-sponsored health insurance variable, I have taken advantage of a few relevant spousal variables as instruments for the potentially endogenous variable representing employer-sponsored health insurance. To get a sense of whether and how well each spousal variable can serve as an instrument, the dissertation has, in its data chapter, shown the statistical association between each pair of the relevant variables and relevant statistics that measure the strength of such association. Then, by using a more rigorous method, the dissertation has investigated the validity and strength of instruments by looking at relevant statistics. It turns out that two of the three instruments are valid and the strongest. These two instruments include the variable indicating whether the spouse was offered health insurance through his or her own employer and the variable measuring the size of the firm the spouse worked.

With instruments chosen, the dissertation moves forward to analyze the model testing the relationship between the productivity proxy variable absenteeism and the variable representing employer-sponsored health insurance. But before jumping to the analysis using instrumental variables, the dissertation has first examined the association between the absenteeism due to sickness or injury and the variable representing employer-sponsored health insurance, with all controls included but without the instruments. The result shows a positive association, as expected. This implies that employer-sponsored health insurance harms productivity. However, this positive number may result from labor market selection or moral hazard or both. Labor market selection means that people with certain unobservable characteristics are more likely both to choose jobs offered with health insurance and to take sick days off. Moral hazard means that once insured through their employers, workers will over-utilize the insurance. Each of the two factors pushes up the measure of the association between the two variables and makes it positive in the end.

Then, by utilizing instruments with advanced econometric techniques, the dissertation has found that the measure of the association between absenteeism due to sickness or injury and the employer-provided insurance status indicator variable becomes negative but is still significant. This result supports the hypothesis that employer-sponsored health insurance improves worker productivity because the relationship between absenteeism due to sickness or injury and productivity is negative. By using different specifications and different data sets, the dissertation has shown that this negative relationship is robust. However, results obtained from the two-part model seem to suggest that employer-provided health insurance only affects the decision about whether to take days off or not, and does not influence the decision about how many days to take.

This, however weird it may seem, is coincidentally or otherwise consistent with one of the theoretical assumptions that health insurance shifts productivity.

Separately using the Husband and the Wife data, the dissertation has found another couple of several results. First health insurance does not seem to affect absenteeism due to sickness or injury for married male workers, but deteriorates it for married female workers before selection or moral hazard effect is removed. Second, health insurance reduces absenteeism due to sickness or injury for both married men and married women after selection or moral hazard effect is purged. Third, the overall impact of health insurance on absenteeism due to sickness or injury is more or less the same for both married men and married women. To further check the validity of the instruments, the dissertation has conducted a “falsification” test on the instruments by checking whether using the instruments changes the relationship between the variable absenteeism for other reasons and the same endogenous variable representing employer-sponsored health insurance relative to the relationship without using the instruments. The dissertation has found that, when a sample containing individuals both with and without children is used, using the instruments does not significantly change the association between the two variables relative, but when a sample containing only childless people, using the instruments does change the association. While causing a significant change in the association does not disprove the validity of the instruments, an insignificant change is expected so that the instruments can be falsified as invalid in this case.

REFERENCES

Abraham, Jean Marie and Feldman, Roger. ”Taking Up or Turning Down: New Estimates of Household Demand for Employer-Sponsored Health Insurance”, Inquiry, Vol. 47, No. 1 (Spring 2010), pp. 17-32

Aizawa, Naoki and Fang, Hanming. “EQUILIBRIUM LABOR MARKET SEARCH AND HEALTH INSURANCE REFORM”, Working Paper 18698

Angrist, Joshua D., Imbens, Guido W.. and Rubin, Donald B., “Identification of Causal Effects Using Instrumental Variables”, *Journal of the American Statistical Association*, Vol. 91, No. 434 (Jun., 1996), pp. 444-455

Baker, Tom. “HEALTH INSURANCE, RISK, AND RESPONSIBILITY AFTER THE PATIENT PROTECTION AND AFFORDABLE CARE ACT”, *University of Pennsylvania Law Review*, Vol. 159, No. 6, Symposium: THE NEW AMERICAN HEALTH CARE SYSTEM: REFORM, REFORMATION, OR MISSED OPPORTUNITY? (June 2011), pp. 1577-1622

Becker, Gary S. and Murphy, Kevin M.. “The Division of Labor, Coordination Costs, and Knowledge”, *The Quarterly Journal of Economics*, Vol. 107, No. 4 (Nov., 1992), pp. 1137-1160

Bloom, David, and Canning, David. “Health as Human Capital and its Impact on Economic Performance”, *The Geneva Papers on Risk and Insurance. Issues and Practice*, Vol. 28, No. 2, SPECIAL ISSUE ON HEALTH (April 2003), pp. 304-315

Boer, Willem de and Boo, Anne de. “Consumer Price Sensitivity in Dutch Health Insurance,” *International Journal of Health Care Finance and Economics*, Vol. 8, No. 4 (Dec., 2008),

Buchele, Robert and Christiansen, Jens. “Worker Rights Promote Productivity Growth”, *Challenge*, Vol. 38, No. 5 (SEPTEMBER-OCTOBER 1995), pp. 32-37

Koç, Çağatay. “Health-Specific Moral Hazard Effects”’ *Southern Economic Journal*, Vol. 72, No. 1 (Jul., 2005), pp. 98-118

Card, David, and Shore-Sheppard, Lara D.. “Using Discontinuous Eligibility Rules to Identify the Effects of the Federal Medicaid Expansions on Low-Income Children.” *The Review of Economics and Statistics* 86, no. 3 (March 13, 2006): 752–66.

Cardon, James H. and Hendel, Igal. “Asymmetric Information in Health Insurance: Evidence from the National Medical Expenditure”, *The RAND Journal of Economics*, Vol. 32, No. 3 (Autumn, 2001), pp. 408-427

Chandra, Amitabh, and Staiger, Douglas. “Productivity Spillovers in Health Care: Evidence from the Treatment of Heart Attacks.” *Journal of Political Economy* 115, no. 1 (2007): 103–40.

Sims, Christopher A.. “But Economics Is Not an Experimental Science, “The Journal of Economic Perspectives”, Vol. 24, No. 2 (Spring 2010), pp. 59-68

Costa-Font, Joan, and Font-Vilalta, Montserrat. “Preference for National Health Service Use and the Demand for Private Health Insurance in Spain”, *The Geneva Papers on Risk and Insurance*. Issues and Practice, Vol. 29, No. 4, SPECIAL ISSUE ON HEALTH (October 2004), pp. 705-718

Cutler, David M. and Madrian, Brigitte C. “Labor Market Responses to Rising Health Insurance Costs: Evidence on Hours Worked”, *The RAND Journal of Economics*, Vol. 29, No. 3 (Autumn, 1998), pp. 509-530Published

Dennis, Jr., William J.. “Wages, Health Insurance and Pension Plans: The Relationship between Employee Compensation and Small Business Owner Income”, *Small Business Economics*, Vol. 15, No. 4 (Dec., 2000), pp. 247-263

Dranove, D., K., Spier, E., and Baker, L. (2000). “`Competition’ among employers offering health insurance”. *Journal of Health Economics,* 19(1), 121–140. https://doi.org/10.1016/S0167-6296(99)00007-7

Dusansky, Richard and Koç, Çağatay. “Implications of the Interaction Between Insurance Choice and Medical Care Demand”, *The Journal of Risk and Insurance*, Vol. 77, No. 1, Special Issue on Health Insurance (March 2010), pp. 129-144

Einav, Liran, and Finkelstein, Amy. “Selection in Insurance Markets: Theory and Empirics in Pictures.” *Journal of Economic Perspectives* 25, no. 1 (February 2011): 115–38. doi:10.1257/jep.25.1.115.

Einav, Liran, Finkelstein, Amy, and Cullen, Mark R.. “Estimating Welfare in Insurance Markets Using Variation in Prices*.” *Quarterly Journal of Economics* 125, no. 3

Emile, Tompa. “The Impact of Health on Productivity: Empirical Evidence and Policy Implications.” *The Review of Economic Performance and Social Progress*, 2, (2002). pp. 93–116.

Fang, Hanming and Gavazza, Alessandro. “Dynamic Inefficiencies in an Employment-Based Health Insurance System: Theory and Evidence”, *The American Economic Review*, Vol. 101, No. 7 (DECEMBER 2011), pp. 3047-3077

Grossman, Michael. “On the Concept of Health Capital and the Demand for Health”, *The Journal of Political Economy*, Vol. 80, No. 2. (Mar. - Apr., 1972), pp. 223-255.

Gruber, Jonathan and Hanratty, Maria. “The Labor-Market Effects of Introducing National Health Insurance: Evidence from Canada”, *Journal of Business & Economic Statistics*, Vol. 13, No. 2, JBES Symposium on Program and Policy Evaluation (Apr., 1995), pp. 163-173

Gruber, Jonathan. “The Incidence of Mandated Maternity Benefits”, *The American Economic Review*, Vol. 84, No. 3 (Jun., 1994), pp. 622-641

Gruber, Jonathan, and Poterba, James. “Tax Incentives and the Decision to Purchase Health Insurance: Evidence from the Self-Employed.” *The Quarterly Journal of Economics* 109, no. 3 (August 1994): 701–33.

Huang, Feng and Gan, Li. “IMPACT OF CHINA'S URBAN EMPLOYEE BASIC MEDICAL INSURANCE ON HEALTH CARE EXPENDITURE AND HEALTH OUTCOMES”, *Working Paper* 20873

Kapur, Kanika, Escarce, José J., Marquis, M. Susan, and Simon, Kosali I.. “Where Do the Sick Go? Health Insurance and Employment in Small and Large Firms”, *Southern Economic Journal,* Vol. 74, No. 3 (Jan., 2008), pp. 644-664

Levy, Helen, and Feldman, Roger. “Does the Incidence of Group Health Insurance Fall on Individual Workers?”, *International Journal of Health Care Finance and Economics*, Vol. 1, No. 3/4, Special Issue: Why Do Employers Do What They Do? Studies of Employer Sponsored Health Insurance (Sep. - Dec., 2001), pp. 227-247

Miller, Nolan H.. “Pricing health benefits: A cost-minimization Approach”, *Journal of Health Economics* 24 (2005), pp. 931–949

Monheit, Alan C.. “Thoughts on Health Insurance Expansions and the Value of Coverage,” *Inquiry*, Vol. 44, No. 2 (Summer 2007), pp. 133-136

Monheit, Alan C.. and Vistnes, Jessica Primoff. “Health Insurance Enrollment Decisions: Preferences for Coverage, Worker Sorting, and Insurance Take-Up”, *Inquiry*, Vol. 45, No. 2 (Summer 2008), pp. 153-167

Monheit, Alan C. and Philip F. Cooper, “Health Insurance and Job Mobility: Theory and Evidence”, *Industrial and Labor Relations Review,* Vol. 48, No. 1 (Oct., 1994), pp. 68-85

Robertson, Ann and Tracy, C Shawn. “Health and productivity of older workers”, *Scandinavian Journal of Work, Environment & Health*, Vol. 24, No. 2 (April 1998), pp. 85 -97

O'Brien, Ellen. “Employers' Benefits from Workers' Health Insurance,” *The Milbank Quarterly*, Vol. 81, No. 1 (2003), pp. 5-43

Olson, Craig A. “Do Workers Accept Lower Wages in Exchange for Health Benefits?”, *Journal of Labor Economics*, Vol. 20, No. 2, Part 2: Compensation Strategy and Design (Apr., 2002), pp. S91-S114

Schulte, Paul and Vainio, Harri. “Well-being at work – overview and perspective”, *Scandinavian Journal of Work, Environment & Health*, Vol. 36, No. 5 (September 2010), pp. 422-429

Smart, Michael and Stabile, Mark. “Tax Credits, Insurance, and the Use of Medical Care”, *The Canadian Journal of Economics / Revue canadienne d'Economique,* Vol. 38, No. 2 (May, 2005), pp. 345-365

The Kaiser Family Foundation and Health Research & Educational Trust, “Employer Health Benefits Annual Survey Report”, 2015

Zuvekas and Olin, Gary L.. “Accuracy of Medicare Expenditures in the Medical Expenditure Panel Survey”, *Inquiry*, Vol. 46, No. 1 (Spring 2009), pp. 92-108

WHO on Health and Economic Productivity, Population and Development Review, Vol. 25, No. 2 (Jun., 1999), pp. 396-401

APPENDIX A

APPENDICES FOR CHAPTER II

A2.1. Tables

Table A2.1. Values of Model Parameters

Parameters	Values	Remarks
$\boldsymbol{\alpha}$	0.75	All cases
$\boldsymbol{\gamma}$	100	All cases
Tax rate t	0.25	All cases

Table A2.2. Distributions of Worker Types

Probability Density Function f(c)	Values	Remarks
$\frac{1}{b-a}$	$b=\gamma; b=0$	Uniform
$\frac{Ae^{Bc}}{1-e^{B\gamma}}$	$A=0.1; B=-0.1$	Truncated Exponential
$N(\mu,\sigma^2)$	$\mu=25,50,75; \sigma=20$	Truncated Normal

Table A2.3. Excess Benefit Functions, Distributions, and Equilibria

g(c)	**f(c)**	c_g^*	**Equilibrium**	**Remarks**
$exp(0.04c)$ $- exp(0.02c)$	$\frac{1}{\gamma - 0}$	89.90	Separating	Increasing, Convex $g'(0) < 1 - \alpha$ $g(\gamma) > (1-\alpha)\gamma$
	$\frac{0.1e^{-0.1c}}{1 - e^{-0.1\gamma}}$	84.90	Separating	
	$N(25, 20^2)$	74.70	Separating	
	$N(50, 20^2)$	76.90	Separating	
	$N(75, 20^2)$	83.30	Separating	
$exp(0.0375c)$ $- exp(0.0350c)$	$\frac{1}{\gamma - 0}$	100	Low-end Pooling	Increasing, Convex $g'(0) < 1 - \alpha$ $g(\gamma) < (1-\alpha)\gamma$
	$\frac{0.1e^{-0.1c}}{1 - e^{-0.1\gamma}}$	100	Low-end Pooling	
	$N(25, 20^2)$	100	Low-end Pooling	
	$N(50, 20^2)$			
	$N(75, 20^2)$	100	Low-end Pooling	
$0.5c^2 + 0.8c$	$\frac{1}{\gamma - 0}$	66.50	Separating	Increasing, Convex $g'(0) > 1 - \alpha$ $g(\gamma) < (1-\alpha)\gamma$
	$\frac{0.1e^{-0.1c}}{1 - e^{-0.1\gamma}}$	19.60	Separating	
	$N(25, 20^2)$	34.80	Separating	
	$N(50, 20^2)$	43.80	Separating	
	$N(75, 20^2)$	65.20	Separating	
$8 + 0.05c$	$\frac{1}{\gamma - 0}$	100	Low-end Pooling	Increasing, Concave $g'(0) < 1 - \alpha$ $g(\gamma) > (1-\alpha)\gamma$
	$\frac{0.1e^{-0.1c}}{1 - e^{-0.1\gamma}}$	0	High-end Pooling	
	$N(25, 20^2)$	0	High-end Pooling	
	$N(50, 20^2)$	36.10	Separating	
	$N(75, 20^2)$	100	Low-end Pooling	
$4\ln(1 + 0.15c)$	$\frac{1}{r - 0}$	100	Low-end Pooling	Increasing, Concave $g'(0) > 1 - \alpha$ $g(\gamma) > (1-\alpha)\gamma$
	$\frac{0.1e^{-0.1c}}{1 - e^{-0.1\gamma}}$	14.40	Separating	
	$N(25, 20^2)$	25.30	Separating	
	$N(50, 20^2)$	100	Low-end Pooling	
	$N(75, 20^2)$	100	Low-end Pooling	
$10 - 0.05c$	$\frac{1}{r - 0}$	27.80	Reversed Separating	Decreasing, Concave $g'(0) < 0$ $g(\gamma) < (1-\alpha)\gamma$
	$\frac{0.1e^{-0.1c}}{1 - e^{-0.1\gamma}}$	21.10	Reversed Separating	
	$N(25, 20^2)$	29.70	Reversed Separating	
	$N(50, 20^2)$	32.70	Reversed Separating	
	$N(75, 20^2)$	33.20	Reversed Separating	

Note: 1) $(\mu, 20^2)$ represents a (truncated) normal density with mean μ and standard deviation of 20.

Table A2.3 shows the relationships among the excess benefit functions, the distributions of worker types, and the resulting equilibria. It can be concluded from the results listed in the table that when the excess benefit functions are increasing and convex, the types of equilibria are robust to the distributional assumptions, but when the excess benefit functions are increasing and concave, the types of equilibria may change with the distributions. Another noticeable point is that when the excess benefit function is decreasing and concave, the separating equilibrium is reversed: workers with costs higher than the cut-off level are enrolled in the moderate plan, whereas those with costs lower than the cut-off level are covered under the generous plan.

A2.2. Matlab Codes

See A3.2.1. Matlab Code Excluduing the Part for the Productivity Model

APPENDIX B

APPENDICES FOR CHAPTER III

A3.1. Graphical Analysis of Productivity Effects on Wages and Equilibria

This graphical analysis has two major purposes. The first IS to show how the productivity effect of health insurance affects the wage associated with the insuranc eplan; the second is to illustrate how hthe productivity effect of health insurance affects a equilibirum. For simplicity, the one-plan model is used. For the examples below, the firm is assumed to bear 75% of the expected health insurance cost ($\alpha = 0.75$), the tax rate is 25% ($t = 0.25$, and) and γ is equal to 100.

Let $w_{mK} = \frac{w_m}{1-t} = \frac{w_u}{1-t} - \delta[1 - F(c)] - \frac{m(c_m)}{1-t}[1 - F(c)]$ denote the pre-tax wage paid to insured workers given productivity differential δ, where K indicates whether δ is positive, zero, or negative. Suppose the health benefit funciton $m(c)$ used here takes a quadratic form as follows:

$$m(c) = Ac^2 + Bc$$

where $A \geq 0$ and $B \geq 0$. The coefficients affect the decisions of the firm. “Wage”, “Wage-“, and “Wage+“, as shown in **Figure A3.1** and **Figure A3.3**, represent the equilbirum wages paid to insured workers in the cases where employer-provided health insurance has no productivity, negative productivity and positive productivity effects, respectively. It is evident that when health insurance has an effect on producitivty, the wage required for all insured workers except the most risky ones is lower than otherwise. Moreover, the wage curves decline intially and rise after they reach a minimum lelel; the three curves converge at the terminal $\boldsymbol{c} = \gamma = 100$.

Figure A3.1. Wages Paid to Insured Workers when $\mathbf{m(c) = 0.01c^2 + 0.01c}$

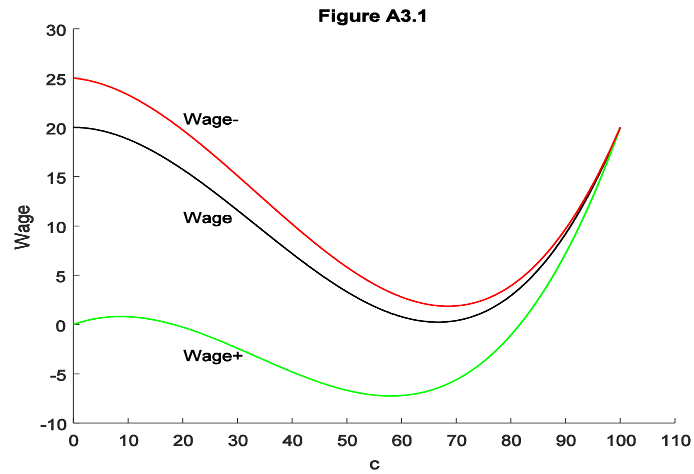
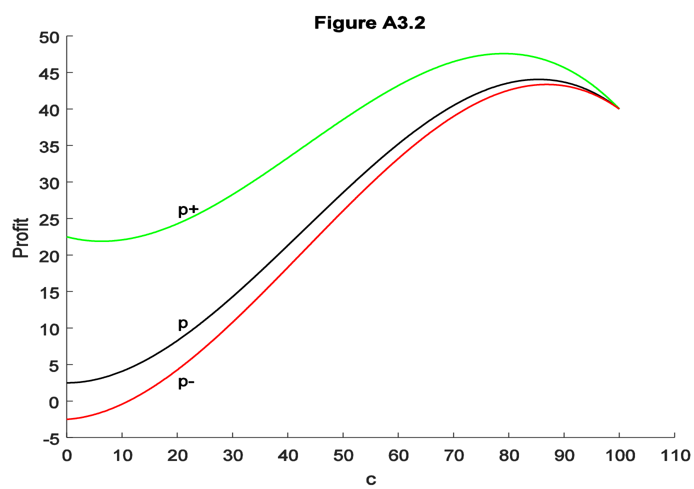


Figure A3.2 and **Figure A3.4** illustrate the profit curves under the three scenarios, where "p+" representsthe profit with positive productivity effect, "p" the profit with no productivity effect, and "p-" the profit with negative productivity effect.

Figure A3.2. Profits Earned by Firm when $\mathbf{m(c) = 0.01c^2 + 0.01c}$



In **Figure 3.2**, the profit curve under any of the three scenarios reaches its maximum before the limit of the expected health insurance cost c ($\boldsymbol{c} = \gamma = 100$) is reached. Hence a separating equilibrium eixsts where the optimal cut-off satisfies $0 < \boldsymbol{c} < \gamma = 100$. This illustrates that, even if health insurance has a postive effect on worker productivty, the nature of the equilibirum does not change, though the equilibrium (optimal cut-off) point is shifted leftward, increasing the number of insured workers.

Figure A3.3. Wages Paid to Insured Workers for $\mathbf{m(c) = 0.005c^2 + 0.005c}$

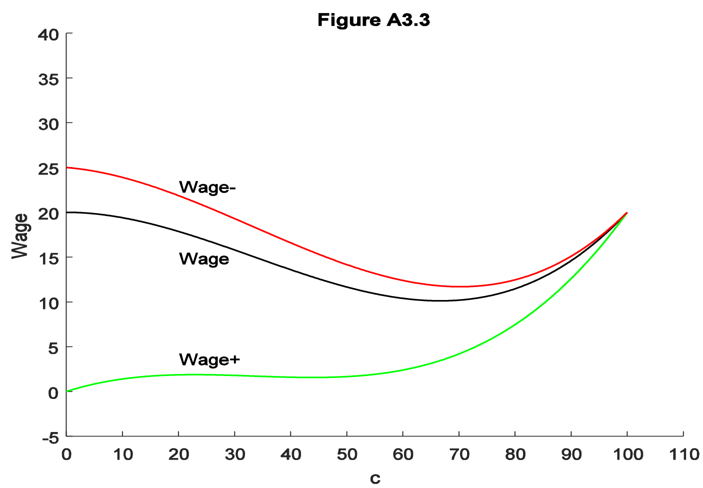
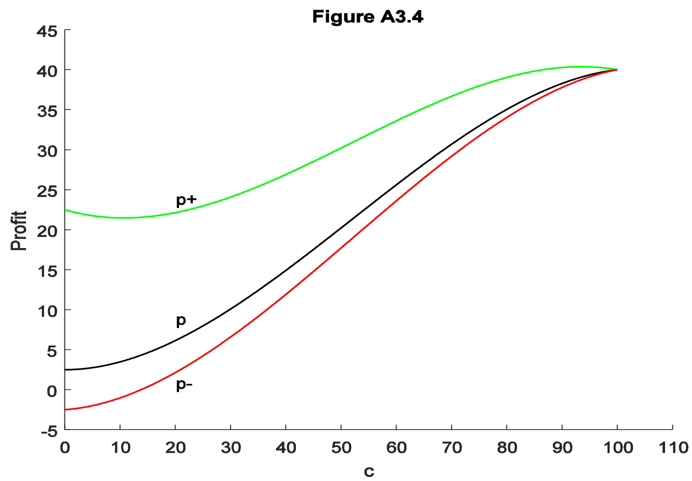


Figure A3.4 illustrates a case where the euilibrium low-end pooling under the scenario where health insurance does not have an effect on worker productivity: the profit curve under this scenario always rise and reach its maximum at the terminal where $\boldsymbol{c} = \gamma = 100$. Under the scenario where health boosts productivty, profit function, the profit curve reaches its maximum at $c = 93.5$, so the equlibrium is separating. This shows that positive productivrty effect may change the equilibirum from low-end pooling to separating. The negativ productivity does not change the nature of the equilibirum, though it raises the optimal cutoff, thereby redducing the numbe rof insured workers.

Figure A3.4. Profits Earned by Firm when $\mathbf{m(c) = 0.005c^2 + 0.005c}$



The above figures illustrate that the productivity of health insurance and the health benefit functions affect wages paid to insured workers as well the nature of equilibrium.

The following table summarizes the effects of worker productivity resulting from employer-provided health insurance on the type of equilibrium.

Table A3.1. Effects of Worker Productivity of Health Insurance on Equilibria

m(c)	f(c)	δ	c_m^*	Equilibrium	Remarks
$0.01c^2 + 0.01c$	$\frac{1}{\gamma - 0}$	0	85.4	Separating	Nature of equilibrium is not affected
		20	79.1	Separating	
		-5	86.	Separating	
$0.005c^2 + 0.005c$	$\frac{1}{\gamma - 0}$	0	100	Low-end pooling	Nature of equilibrium is affected
		20	93.5	Separating	
		-5	100	Low-end pooling	

A3.2. Matlab Codes

A3.2.1. Matlab Code for Section 3.2 and Section 3.3

```
clear

gamma=100;
alpha=0.75;
delta0=0;
deltap=2.25;
%deltan=-5;
e=20;
t=0.25;
Wm=15;
syms x c

%Define a uniform pdf;

fc=unifpdf(c,0,gamma);
F =int(fc,0,c); %CDF;
gc=exp(0.035*c)-exp(0.018*c);
dgc=diff(gc,c,1);
gcc=gc-dgc*(1-F)/fc;
dgcf=dgc*(1-F)/fc;
ic=(1-alpha)*c;
delta=deltap+0.00035*c^2;
delta2=0+sqrt((1-F)/fc/gamma)+0.5*1/((1-F)/fc/gamma)*(1-
c/gamma)+0.00035*c^2; %productivity changes with c;

cg=0:0.1:gamma; %define and assign values to cg;
gc_=subs(gc,c,cg); % replace c with cg;
dgc_=subs(dgc,c,cg);
dgcf_=subs(dgcf,c,cg);
ic_=subs(ic,c,cg);
gcc_=subs(gcc,c,cg);
delta_=subs(delta,c,cg);
delta2_=subs(delta2,c,cg);
gct=gc_/(1-t);;
gcta=gcc_/(1-t);
gctap=gcc_/(1-t)+delta_;
gctap2=gcc_/(1-t)+delta2_;
hc=gcta; %redefine a new function;
hold on

figure
plot(cg,ic_,'r',cg,gc_,'g','LineWidth',1.0);
box off
xlim([0 110]);
ylim([-10 40]);
xlabel('c');
ylabel('function of c');
```

```
title('Figure 2.1')
text(101,25,'(1-a)c','FontSize',10) %add text to the graph
text(101,28,'g(c)','FontSize',10)
text(97,23.5,'A','FontSize',10)

figure
plot(cg,ic_,'r',cg,gc_,'g',cg,gct,'b','LineWidth',1.0);
box off
xlim([0 110]);
ylim([-10 40]);
xlabel('c');
ylabel('function of c');
title('Figure 2.2')
text(101,25,'(1-a)c','FontSize',10)
text(101,28,'g(c)','FontSize',10)
text(84,34,'g(c)/(1-t)','FontSize',10)
text(97,23.5,'A','FontSize',10)
text(85,23,'B','FontSize',10)

figure
plot(cg,ic_,'r',cg,gc_,'g',cg,gct,'b',cg,gcta,'k','LineWidth',1.0);
box off
xlim([0 110]);
ylim([-10 40]);
xlabel('c');
ylabel('function of c');
title('Figure 2.3')
text(101,25,'(1-a)c','FontSize',10)
text(101,28,'g(c)','FontSize',10)
text(84,34,'g(c)/(1-t)','FontSize',10)
text(101,34,'h(c)','FontSize',10)
text(97,23.5,'A','FontSize',10)
text(85,23,'B','FontSize',10)
text(92,25,'C','FontSize',10)

%Productivity model
figure
plot(cg,ic_,'r',cg,gcta,'k','LineWidth',1.0);
box off
xlim([0 110]);
ylim([-10 40]);
xlabel('c');
ylabel('function of c');

title('Figure 3.1')
text(101,25,'(1-a)c','FontSize',10)
text(101,34,'h(c)','FontSize',10)
text(95.5,23,'C','FontSize',10)

figure
plot(cg,ic_,'r',cg,gcta,'k',cg,gctap,'m','LineWidth',1.0);
box off
xlim([0 110]);
```

```
ylim([-10 40]);
xlabel('c');
ylabel('function of c');
title('Figure 3.2')
text(101,25,'(1-a)c','FontSize',10)
text(101,34,'h(c)','FontSize',10)
text(90,34,'p(c)','FontSize',10)
text(95.5,23,'C','FontSize',10)
text(89,24,'D','FontSize',10)

figure
plot(cg,ic_,'r',cg,gcta,'k',cg,gctap2,'m','LineWidth',1.0);
box off
xlim([0 110]);
ylim([-10 40]);
xlabel('c');
ylabel('function of c');
title('Figure 3.3')
text(101,25,'(1-a)c','FontSize',10)
text(101,34,'h(c)','FontSize',10)
text(90,34,'q(c)','FontSize',10)
text(95.5,23,'C','FontSize',10)
text(90,24.25,'E','FontSize',10)

hold off
```

A3.2.2. Matlab Code for Appendix A3.1

```
clear
gamma=100;
alpha=0.75;
delta0=0;
deltap=20;
deltan=-5;
e=60;
MaxIter=50;
t=0.25;
Wu=15;
syms x
c=0:0.1:gamma;
[rc,dimc]=size(c);
    cm=0:0.1:gamma;
    mc=zeros(rc,dimc)+NaN;
    Wm0=zeros(rc,dimc)+NaN;
    Wmp=zeros(rc,dimc)+NaN;
    Wmn=zeros(rc,dimc)+NaN;
    ic=zeros(rc,dimc)+NaN;
    %Define a uniform pdf;
    p = unifcdf(cm,0,gamma);

for ik=1:dimc
    mc(:,ik)=0.01*(cm(:,ik)^2+0.01*cm(:,ik));
    %mc(:,ik) =0.005*(cm(:,ik)^2+0.005*cm(:,ik));
```

```
    Wm0(:,ik)=Wu/(1-t)-(1-p(:,ik))*delta0-(1-p(:,ik))*mc(:,ik)/(1-t);
    Wmp(:,ik)=Wu/(1-t)-(1-p(:,ik))*deltap-(1-p(:,ik))*mc(:,ik)/(1-t);
    Wmn(:,ik)=Wu/(1-t)-(1-p(:,ik))*deltan-(1-p(:,ik))*mc(:,ik)/(1-t);
end;

profit0=e-Wm0-ic;
profitp=e-Wmp-ic;
profitn=e-Wmn-ic;

hold on

figure
plot(cm,Wm0,cm,Wmp,cm,Wmn,'LineWidth',1.0);
box off
xlim([0 110]);
ylim([-5 40]);
xlabel('c');
ylabel('Wage');
title('Figure A3.1');
text(20,23,'Wage-','FontSize',10)
text(20,15,'Wage','FontSize',10)
text(20,3.75,'Wage+','FontSize',10)

figure
plot(cm,profit0,'k',cm,profitp,'g',cm,profitn,'r','LineWidth',1.0);
box off
xlim([0 110]);
%ylim([-5 50]);
xlabel('c');
ylabel('Profit');
title('Figure A3.2')
text(20,24,'p+','FontSize',10)
text(20,9,'p','FontSize',10)
text(20,1,'p-','FontSize',10)

hold off

%maximize profit function with no productivity;
[maxpi0,index]=max(profit0);
maxpi_0=double(maxpi0)
display('Maximizing cm_0');
display(cm(index))
%maximize profit function with positive productivity;
[maxpip,index]=max(profitp);
maxpi_p=double(maxpip)
display('Maximizing cm_p');
display(cm(index))
%maximize profit function with negative productivity;
[maxpin,index]=max(profitn);
maxpi_p=double(maxpin)
display('Maximizing cm_n');
display(cm(index))
```

A3.3. Calculation of Profits for the One-plan Model

The total profit function is

$$e-\frac{w_u}{1-t}+\delta\int_{c_m^*}^{\gamma}f(c)dc-\alpha\int_{c_m^*}^{\gamma}cf(c)dc+\frac{m(c_m^*)}{1-t}\int_{c_m^*}^{\gamma}f(c)dc$$

which can be expressed as

$$e-\frac{w_u}{1-t}+\delta\big(1-F(c_m^*)\big)-\alpha\int_{c_m^*}^{\gamma}cf(c)dc+\frac{m(c_m^*)}{1-t}\big(1-F(c_m^*)\big)$$

The term $\int_{c_m^*}^{\gamma}cf(c)dc$ can be calculated through integration by parts as follows:

$$\int_{c_m^*}^{\gamma}cf(c)dc=[c^2f(c)]_{c_m^*}^{\gamma}-\int_{c_m^*}^{\gamma}cd[cf(c)]$$

$$\int_{c_m^*}^{\gamma}cf(c)dc=[c^2f(c)]_{c_m^*}^{\gamma}-\int_{c_m^*}^{\gamma}c[f(c)dc+cdf(c)]$$

$$\int_{c_m^*}^{\gamma}cf(c)dc=[c^2f(c)]_{c_m^*}^{\gamma}-\int_{c_m^*}^{\gamma}cf(c)dc-\int_{c_m^*}^{\gamma}cdf(c)$$

$$2\int_{c_m^*}^{\gamma}cf(c)dc=[c^2f(c)]_{c_m^*}^{\gamma}-\int_{c_m^*}^{\gamma}cdf(c)$$

$$2\int_{c_m^*}^{\gamma}cf(c)dc=[c^2f(c)]_{c_m^*}^{\gamma}-[cf(c)]_{c_m^*}^{\gamma}+\int_{c_m^*}^{\gamma}f(c)d\,c$$

$$2\int_{c_m^*}^{\gamma}cf(c)dc=[c^2f(c)]_{c_m^*}^{\gamma}-[cf(c)]_{c_m^*}^{\gamma}+[F(c)]_{c_m^*}^{\gamma}$$

$$2\int_{c_m^*}^{\gamma}cf(c)dc=\gamma^2f(\gamma)-c_m^{*2}f(c_m^*)-[\gamma f(\gamma)-c_m^*f(c_m^*)]+[1-F(c_m^*)]$$

$$2\int_{c_m^*}^{\gamma}cf(c)dc=(\gamma-1)\gamma f(\gamma)-(c_m^*-1)c_m^*f(c_m^*)+[1-F(c_m^*)]$$

Hence,

$$\int_{c_m^*}^{\gamma} cf(c)dc = \frac{1}{2}\{(\gamma-1)\gamma f(\gamma) - (c_m^* - 1)c_m^* f(c_m^*) + [1 - F(c_m^*)]\}$$

A3.4. An Efficiency Wage Model

Suppose productivity changes continuously with the wages offered to workers, rather than varies with health insurance plans. Specifically, assume worker productivity is an increasing, continuous function of wages. Let $e(w_k)$(where $k = m, g$) denote the productivity as a function of the relevant wage, and assume $e'(w_k) > 0$. Informational asymmetry or anti-discriminatory policy forces the firm to offer an equal wage to all workers who choose the same health insurance. Then, using the same notations and functions as before, the firm's problem can be expressed as follows.

$$\max_{0 \le w_m, 0 \le w_g} \int_0^{c_g} \left(e(w_m) - \frac{w_m}{1-t} - \alpha c\right) f(c)dc + \int_{c_g}^{\gamma} \left(e(w_g) - \frac{w_g}{1-t} - c\right) f(c)dc$$

$$w_m + m(c) \ge \overline{u}$$

$$w_g + m(c_g) + g(c_g) \ge \overline{u}$$

$$w_g + m(c_g) + g(c_g) \ge w_m + m(c_g)$$

The first is the participation constraint for workers who would elect the moderate plan, the second is the participation constraint for those who would choose the generous plan, and the third is the incentive-compatibility constraint. These constraints do not need to hold simultaneously. The first is required only if at least some workers are to be enrolled in the moderate plan and the second is needed only if at least some workers choose the generous plan. If the first holds, then the second constraint is redundant because it is implied by the first and the last. But if the second holds, the first does not necessarily hold. Suppose the first holds now. So the second is redundant and can be removed. If the first constraint holds for all workers then it must hold for the lowest-cost ones, who have zero cost. On the other hand, as long as it holds for the lowest-cost workers, it must hold for all, because $m(c)$ is non-decreasing in c. The last constraint can be written as $w_g + g(c_g) \geq w_m$. Hence the firm's problem can be re-written as

$$\max_{0 \leq w_m, 0 \leq w_g} \int_0^{c_g} \left(e(w_m) - \frac{w_m}{1-t} - \alpha c\right) f(c)dc + \int_{c_g}^{\gamma} \left(e(w_g) - \frac{w_g}{1-t} - c\right) f(c)dc$$

$$w_m + m(0) \geq \overline{u}$$

$$w_g + g(c_g) \geq w_m$$

The difference between this model and the any of the two plan models presented in Chapter 3 is that the optimal wage associated with the generous plan may not be lower than the optimal wage associated with the moderate plan, because here wage boosts productivity whereas in the models presented earlier, wage does not.

By continuity, there exists a smallest c that can make the incentive-compatibility constraint hold with equality, i.e., $w_g + m(c_g) + g(c_g) = w_m + m(c_g)$. This implies $g(c_g) = w_m - w_g$, so $c_g = g^{-1}(w_m - w_g)$. Hence the firm's problem can be re-written as

$$\max_{0 \leq w_g < w_m} \int_0^{g^{-1}(w_m - w_g)} \left(e(w_m) - \frac{w_m}{1-t} - \alpha c\right) f(c)dc + \int_{g^{-1}(w_m - w_g)}^{\gamma} \left(e(w_g) - \frac{w_g}{1-t} - c\right) f(c)dc$$

$$w_m + m(0) \geq \overline{u}$$

It is important to note that both the two wage variables are endogenous. Moreover, the (participation) constraint may or may not be both binding.

Suppose first that the constraint is binding (for lowest-cost workers), i.e., $w_m + m(0) = \overline{u}$, so $w_m = \overline{u} - m(0)$, which says w_m is fixed; given $g(c_g) = w_m - w_g$ and w_m is constant, there is one to one mapping between c_g and w_g.The problem is therefore the same as what is presented in Chapter 3 but the optimality conditions are the same as well. There are three possible equilibria, including high-end pooling, in which case $c_g = 0$, so $w_m = w_g = \overline{u} - m(0)$; low-end pooling, in which case $c_g = \gamma$, so $w_m - w_g \geq g(\gamma)$, and $w_m = \overline{u} - m(0)$; and separating, in which case $0 < c_g < \gamma$, so $0 < w_m - w_g < g(\gamma)$. The profit that the firm would obtain under each type of equilibrium can be respectively calculated as follows.

$$\pi^{hp(w_m=\overline{u}-m(0))} = e(\overline{u} - m(0)) - \frac{\overline{u} - m(0)}{1-t} - \int_0^\gamma cf(c)dc$$

$$\pi^{lp(w_m=\overline{u}-m(0))} = e(\overline{u} - m(0)) - \frac{\overline{u} - m(0)}{1-t} - \alpha \int_0^\gamma cf(c)dc$$

$$\pi^{sp(w_m=\overline{u}-m(0))} =$$

$$\left[e(\overline{u} - m(0)) - \frac{\overline{u}-m(0)}{1-t}\right] F(c_g) - \alpha \int_0^{c_g} cf(c)dc + \left[e(w_g) - \frac{w_g}{1-t}\right][1 - F(c_g)] - \int_{c_g}^\gamma cf(c)dc\ .$$

where hp stands for high-end pooling, lp for low-end pooling, and sp for separating equilibrium. Comparing the three profit expressions pairwise reveals that, while it is not certain whether the profit under separating equilibrium is greater than the profit under high-end pooling equilibrium, it is obvious that the profit under low-end pooling equilibrium exceeds the profit under any of the other two equilibria, given that the productivity function is rising in wage. Therefore, the firm's optimal decision is to set the wage for the moderate plan at $\overline{u} - m(0)$ and the wage for the generous plan as low as $\overline{u} - m(0) - g(\gamma)$ or even lower, thus pooling all its workers into the moderate plan.

The above analysis assumes the participation constraint is binding, but that does not need to be the case.

Suppose, alternatively, the constraint is not binding, i.e., $w_m + m(c) > \overline{u}$. Then the problem becomes one of unconstrained maximization. The first order conditions (FOC) can therefore be derived as

$$\left(e'(w_m) - \frac{1}{1-t}\right) F\left(g^{-1}\left(w_m - w_g\right)\right) + \left(e(w_m) - \frac{w_m}{1-t} - \alpha g^{-1}\left(w_m - w_g\right)\right)\frac{f(w_m - w_g)}{g'(w_m - w_g)} - \left(e\left(w_g\right) - \frac{w_g}{1-t} - g^{-1}\left(w_m - w_g\right)\right)\frac{f(w_m - w_g)}{g'(w_m - w_g)} \leq 0 \quad (1)$$

$$\left(e'\left(w_g\right) - \frac{1}{1-t}\right)\left(1 - F\left(g^{-1}\left(w_m - w_g\right)\right)\right) + \left(e\left(w_g\right) - \frac{w_g}{1-t} - g^{-1}\left(w_m - w_g\right)\right)\frac{f(w_m - w_g)}{g'(w_m - w_g)} - \left(e(w_m) - \frac{w_m}{1-t} - \alpha g^{-1}\left(w_m - w_g\right)\right)\frac{f(w_m - w_g)}{g'(w_m - w_g)} \leq 0 \quad (2)$$

The inequalities hold if and only if the respective wage is zero and the equalities hold if and only the respective wage is positive.

Adding the two inequalities and simplifying yields

$$\left(e'(w_m) - e'\left(w_g\right)\right) F\left(c_g\right) + \left(e'\left(w_g\right) - \frac{1}{1-t}\right) \leq 0 \quad (3)$$

where $c_g = g^{-1}\left(w_m - w_g\right)$. Note that if $c_g = 0$, then equation (3) becomes $e'\left(w_g\right) - \frac{1}{1-t} \leq 0$, and if $c_g = \gamma$, it becomes $e'(w_m) - \frac{1}{1-t} \leq 0$. Under reasonable assumptions, under reasonable assumptions about the reservation utility and the health benefit of the moderate plan for the lowest-cost workers, because the participation constraint would be violated; even if no such assumptions are made, the solution still is trivial and is not worth analyzing, the solution $w_m =$

$w_g = 0$ can be ruled out. Therefore, equations (1), (2) and (3) hold with equalities. Equation (3) holding with equality can be solved for special cases.

(1) If $w_m \leq w_g$, then $w_m - w_g \leq 0$, so $g^{-1}(w_m - w_g) \leq g^{-1}(0)$; hence $c_g = 0$ because $c_g = 0$ is the smallest value c_g that can take. This means the resulting equilibrium outcome is high-end pooling. Then, because $F\left(g^{-1}(w_m - w_g)\right) = F(0) = 0$ and because equation (3) therefore becomes (assuming a positive wage)

$$e'(w_g) - \frac{1}{1-t} = 0$$

The condition above has to hold with equality because it is assumed that $0 < w_m < w_g$. Optimal wage for the generous plan is determined by the above equation. The firm's profit is therefore

$$\pi^{hp(w_m > \bar{u} - m(0))} = e(w_g^*) - \frac{w_g^*}{1-t} - \int_0^{\gamma} cf(c)dc$$

where $\int_0^{\gamma} \mathrm{cf(c)dc} = \frac{1}{2}\{(\gamma - 1)\gamma\mathrm{f}(\gamma) + 1\}$ and hp stands for high-end pooling because $c_g = 0$.

(2) If $w_m > w_g$, then there are two possibilities: 1) $w_m - w_g \geq g(\gamma)$; and 2) $0 < w_m - w_g < g(\gamma)$.

1) If $w_m - w_g \geq g(\gamma)$, then $c_g = \gamma$, and so the equilibrium is low-end pooling. Moreover, because $F\left(g^{-1}(w_m - w_g)\right) = F(\gamma) = 1$, Equation (3) can be simplified as (assuming a positive wage)

$$e'(w_m) - \frac{1}{1-t} = 0$$

The optimal wage is determined by the above equation, and is the same as in (1). The profit is

$$\pi^{lp(w_m > \overline{u} - m(0))} = e\left(w_g^*\right) - \frac{w_g^*}{1-t} - \alpha \int_0^{\gamma} cf(c)dc$$

Obviously, $\pi^{lp(w_m > \overline{u} - m(0))} > \pi^{hp(w_m > \overline{u} - m(0))}$because wage is the same in both cases. This tells that the low-end pooling equilibrium is superior to the high-end pooling one when wage boosts productivity.

2) If $0 < w_m - w_g < g(\gamma)$, then $g^{-1}(0) < g^{-1}(w_m - w_g) < g^{-1}(\gamma)$ so $0 < c_g < \gamma$. The potential equilibrium is separating. The optimal solutions can be derived by solving (1) and (2), assuming they hold with equalities. The profit is therefore

$$\pi^{sp(w_m > \overline{u} - m(0))} = \left(e(\, w_m^*) - \frac{w_m^*}{1-t}\right) F\left(c_g^*\right) - \alpha \int_0^{c_g^*} cf(c)dc + \left(e\left(w_g^*\right) - \frac{w_g^*}{1-t}\right)\Big(1 -$$

$$F\left(c_g^*\right)\Big) - \int_{c_g^*}^{\gamma} cf(c)dc$$

where $\int_{c_1}^{c_2} \mathrm{cf(c)dc} = \frac{1}{2}\{(c_2 - 1)c_2 \mathrm{f}(c_2) - (c_1 - 1)c_1 \mathrm{f}(c_1) + [\mathrm{F}(c_2) - \mathrm{F}(c_1)]\}$.

It is not clear whether $\pi^{lp(w_m - w_g \geq g(\gamma))} \geq \pi^{sp(w_m > \overline{u} - m(0))}$ or $\pi^{lp(w_m - w_g \geq g(\gamma))} < \pi^{sp(w_m > \overline{u} - m(0))}$. Hence, the firm chooses the maximum of the two.

$$\pi^{max} = \max(\pi^{lp(w_m > \overline{u} - m(0))}, \pi^{sp(w_m > \overline{u} - m(0))})$$

Then the firm decides the corresponding optimal scheme, which leads to either a low-end pooing or a separating equilibrium.

Conclusions

When wage boosts productivity, high-end pooling equilibrium is no longer possible because it is dominated by the low-pooling equilibrium from the firm's perspective: pooling leads to the same optimal wage, but given the same wage, enrolling all workers in the moderate plan generates the same amount of output but incurs less expenses than enrolling them all in the generous plan. But it is not quite clear whether low-end pooling is super to separating equilibrium.

APPENDIX C

APPENDICES FOR CHAPTER IV

A4.1. A Model in Which the Employer Pays a Variable Share of the Premium

The model presented in Section 3.4 of Chapter assumes that the firm bears a fixed portion of the expected health insurance cost and attains its goal of profit maximization by manipulating wages. However, there may be firms that use something other than wages as instruments to affect its employees' insurance choice decisions. This section presents a model where the firm utilizes such an alternative tool to induce its employees to either accept or decline the insurance policy.

In principle, the firm can utilize two wages and one insurance cost sharing tool as instruments to influence the decisions of its workers. However, if the firm is assumed to offer two wages, the model would be different from the one presented in the previous section. Thus, it is reasonable and desirable to assume that the firm offers only one wage. The wage is primarily used to recruit and retain workers, and the cost-sharing tool is used to induce workers to take up or turn down the insurance policy offer. If workers accept the insurance plan, they would receive a compensation package comprising a wage and the health insurance policy. Otherwise, they will receive a package consisting of the wage only.

Let α denote the proportion of the expected health insurance cost to be borne by the firm. So the cost that the firm would need to assume is αc if a worker of type c chooses to be insured, and the portion of the expected health insurance cost worker would need to bear is $1-\alpha$. Let w denote the wage offered to workers. If a worker is enrolled in the plan, his total utility is $\mathrm{w} + m(\mathrm{c}) - (1-\alpha)\mathrm{c}$; otherwise, his total utility is simply the wage w.

Assume productivity differs between insured and uninsured workers. If a worker is uninsured, his productivity would be e. If the worker is insured, his productivity would be $e + \delta$. Under the distributional assumption about worker types (or workers' expected health care cost) made earlier, the profit the firm obtains from insured worker $\int_{c_m}^{\gamma}\left(e+\delta-\frac{w}{1-t}-\alpha c\right)f(c)dc$ and the profit from each uninsured worker is $\int_{0}^{c_m}\left(e-\frac{w}{1-t}\right)f(c)dc$. The firm's objective is to maximize its expected profit, given that the compensation packages are acceptable to the workers. Mathematically, the firm's problem can be written as

$$\max_{0\le w, \mathbf{0}\le\alpha\le 1}\int_{0}^{c_m}\left(e-\frac{w}{1-t}\right)f(c)dc\ +\int_{c_m}^{\gamma}\left(e+\delta-\frac{w}{1-t}-\alpha c\right)f(c)dc$$

s.t.:

$$w \ge \bar{u}$$

$$w + m(c) - (1-\alpha)c \ge \bar{u}$$

$$m(c) \ge (1-\alpha)c$$

The constraints are not required to hold simultaneously. Only if the firm would like to separate its workers into insured and uninsured groups (separating equilibrium) are all the three constraints necessary. Otherwise, only the first constraint or the last two are needed. The last one is the incentive compatibility constraint, which always holds with equality for $c = 0$. This means there might be situations where only the lowest cost-workers choose to be insured. To prevent this kind of situations from happening, a restriction has to be imposed: if only one type of workers chooses to be insured and that type is indifferent between insured and uninsured, then insurance will not be offered. The case where $w < \bar{u}$ is not considered, because the firm can use the other tool to achieve the same goal.

Optimality requires $w = \bar{u}$. So the second (participation) constraint holds with equality for $c = 0$. In situations where $m(c) \geq (1-\alpha)c$ for all $\alpha \in [0,1]$ and for all $c \in [0,\gamma]$. For example, if $m(c) > c$ for $c \in (0,\gamma)$, then all workers self-select to be insured regardless of the firm's decisions on α. Therefore, the firm would not have enough instruments to induce workers' choice and thus only one possible type of equilibrium exists: pooling into the insurance plan. To avoid this kind of situations, it is necessary to impose restrictions on the function $m(c)$. So assume

$$m'(0) < 1 \qquad (A1)$$

With this assumption, the firm can choose α to affect workers' decisions, because then $m(c) < c$ for some c and thus by adjusting α, the firm can have $m(c) \geq (1-\alpha)c$ for its desired level of c.

For the last two constraints in the firm's profit-maximization problem, there exists $c_m \in [0,\gamma]$ such that they both are binding, i.e., $w + m(c_m) - (1-\alpha)c_m = \bar{u}$ and $m(c_m) - (1-\alpha)c_m = 0$. Hence only one of the last two constraints is necessary. If some workers choose to be insured, then the lowest-cost workers who opt to be insured are indifferent between insured and uninsured. With $m(c) \geq (1-\alpha)c$ and $w = \bar{u}$, the firm's objective function can be re-written as

$$\max_{\mathbf{0 \leq \alpha \leq 1}} e - \frac{\bar{u}}{1-t} + \int_{c_m}^{\gamma} (\delta - \alpha c) f(c) dc$$

s.t.:

$$m(c) \geq (1-\alpha)c$$

If $m(c) < c$, there exists $c_m \in [0,\gamma]$ such that the incentive compatibility constraint can be made binding, i.e., $m(c_m) = (1-\alpha)c_m$, which implies $\alpha c_m = c_m - m(c_m)$. Solving for α in terms of c_m yields $\alpha = 1 - \frac{m(c_m)}{c_m}$. There is a one-to-one mapping between α and c_m. But $\mathbf{0} \leq \alpha \leq 1$ implies $0 \leq \frac{m(c_m)}{c_m} \leq 1$, or $0 \leq m(c_m) \leq c_m$. This may not always be the case, because it is possible that $m(c_m) > c_m$; to accommodate this possibility, let's define $\alpha = \max(1 - \frac{m(c_m)}{c_m}, 0)$. In addition, using l'Hôpital's rule, I define $\frac{m(c_m \to 0)}{c_m \to 0} = \frac{m'(0)}{1} = m'(0)$. Moreover, $w = \bar{u}$ at the optimum. The firm's problem then becomes

$$\max_{c_m \in [0,\gamma]} \delta \int_{c_m}^{\gamma} f(c)dc - max\left(1 - \frac{m(c_m)}{c_m}, 0\right) c_m \int_{c_m}^{\gamma} cf(c)dc$$

s.t.:

$$m(c_m) \geq \left(1 - max\left(1 - \frac{m(c_m)}{c_m}, 0\right)\right) c_m$$

The Langrangian is skipped. The way to find the solution to the firm's profit maximization problem is write down the Langrangian and then differentiate the Langrangian with respect to c_m. A simpler way to solve the problem is to take advantage of the relationship between c_m and the function $m(c_m)$: if $m(c_m) \geq c_m$, then $max\left(1 - \frac{m(c_m)}{c_m}, 0\right) = 0$. So the constraint becomes $m(c_m) \geq c_m$, which is consistent with the assumption. Thus, only the first term is left in the objective function. If $m(c_m) < c_m$, then $max\left(1 - \frac{m(c_m)}{c_m}, 0\right)$=$1 - \frac{m(c_m)}{c_m}$. The constraint becomes $m(c_m) \geq m(c_m)$, which still holds but is not constraining and so can be disregarded, and the objective function is simplified. In either case, the problem is easy to solve.

Solving the problem using the approach stated at the end of the last paragraph yields the following first order conditions, which are presented as Proposition A4.1.

Proposition A4.1. *The Equilibrium is characterized by the first order conditions for the firm's profit maximization problem:*

$$\frac{1}{f(c_m^*)}\left(\frac{m(c_m^*)}{c_m^{*2}} - \frac{m'(c_m^*)}{c_m^*}\right) I_A \int_{c_m^*}^{\gamma} cf(c)dc + \max(c_m^* - m(c_m^*), 0) \begin{cases} < \delta, if\ m(c_m^*) \geq c_m^* \geq 0 \\ = \delta, \text{if } m(c_m^*) < c_m^* < \gamma \\ > \delta, if\ m(\gamma) < c_m^* = \gamma \end{cases}$$

where $I_A = I_{m(c_m^*)<c_m^*} \begin{cases} = 1 \text{ if } m(c_m^*) < c_m^* \\ = 0 \text{ otherwise} \end{cases}$.

Proof. Differentiating the firm's objective function with respect to c_m^* yields the results. See Appendix for the derivation.

The term $\int_{c_m^*}^{\gamma} cf(c)dc$ can be computed using integration by parts as follows (see the Appendix for the derivation)

$$\int_{c_m^*}^{\gamma} cf(c)dc = \frac{1}{2}\{(\gamma - 1)\gamma f(\gamma) - (c_m^* - 1)c_m^* f(c_m^*) + 1 - F(c_m^*)\}$$

The firm's decisions depend on productivity differential δ. Suppose δ is non-positive. Then if $m(c_m^*) \geq c_m^* \geq 0$, the first order conditions would be violated. In this case, $\alpha^* = 0$, which is the only instrument under the firm's control (the wage instrument is for induce workers to join the firm only).

In other situations, firm's desired optimal solutions to the problem would be $\alpha^* = 0$ and $c_m^* = \gamma$. However, the firm can only make choices on wages and the share of insurance cost to induce its workers to choose to be insured or uninsured and cannot directly make decisions for its workers. Hence the optimal solutions are $w = \bar{u}$ and $\alpha = 0$. Then the decision on whether or not to be insured is up to the workers.

Workers' decisions rest entirely on the insurance benefit function $\mathrm{m}(\mathrm{c})$. If $\mathrm{m}(\mathrm{c}) < c$ holds for all $\mathrm{c} \in (0, \gamma]$, then all workers would choose to stay uninsured because the incentive compatibility constraint is violated. A sufficient (but not necessary) condition for this conclusion to hold is as follows. If the marginal benefit of health insurance is lower than its marginal cost or $\mathrm{m}'(\mathrm{c}) < 1$ for all $\mathrm{c} \in [0, \gamma]$, because in this case the benefit from insurance is always lower than its cost. If, on the other hand, $\mathrm{m}(\mathrm{c}) \geq c$ holds for all $\mathrm{c} \in [0, \gamma]$ (or more restrictively, the marginal benefit of health insurance is greater than its marginal cost or $\mathrm{m}'(\mathrm{c}) > 1$ for all $\mathrm{c} \in [0, \gamma]$), then all workers would choose to be insured, since the benefit from health insurance always outweighs its cost. In the case where insurance has adverse effects on worker productivity, the firm may suffer because it is unable to prevent its workers from being insured if they choose to do so. This is negative externality of health insurance. If there exists some $\mathrm{c} \in (0, \gamma)$ such that $\mathrm{m}(\mathrm{c}) \geq \mathrm{c}$, then there must exists $\mathrm{c} \in (0, \gamma)$, denoted as $\mathrm{c_m}$, that satisfies the condition $\mathrm{m}(\mathrm{c_m}) = \mathrm{c_m}$. Hence, some workers would opt to be insured and others to be uninsured. These results are summarized below.

Proposition A4.2. *Suppose* $m'(c) > 0$*,* $m''(\mathrm{c}\,) > 0$, and $m(c) = 0$ *for* $c \in [0, \gamma]$*; suppose further that the health insurance does not boost worker productivity. Then a separating equilibrium exists if and only if the insurance benefit function* $m(c)$ *has one and only one fixed point, i.e.,* $m(\mathrm{c_f}) = \mathrm{c_f}$ *for* $\mathrm{c_f} \in (0, \gamma)$*.*

Proof. Let $\mathrm{c_f} \in (0, \gamma)$ denote the fixed point. So $\mathrm{m}(\mathrm{c_f}\,) = \mathrm{c_f}$. This implies $\mathrm{m}(\mathrm{c_f}\,) - \mathrm{c_f} = 0$. Moreover, since health insurance does not boost worker productivity, the productivity differential between potentially insured and uninsured workers is non-positive, or $\delta \leq 0$. Hence, the profit-maximizing firm optimally sets $\mathrm{w} = \overline{\mathrm{u}}$ and $\alpha = 0$. Let $\mathrm{c_m} = \mathrm{c_f}$. Then $\mathrm{m}(\mathrm{c_f}\,) - \mathrm{c_f} = 0$ implies

$m(c_m) - c_m = 0$, which, together with $\alpha = 0$, implies $m(c_m) - (1-\alpha)c_m = 0$ for $c_m \in (0, \gamma)$. Let $h(c) = m(c) - c$. Then according to the mean-value theorem, there exists $c \in (0, c_f)$ such that $h(c_f) - h(0) = h'(c)(c_f - 0)$, implying $h'(c) = 0$; $h''(c) = m''(c) > 0$, which implies $h'(c_f) > h'(c) = 0$ for $c \in (0, c_f)$,and $h'(c) > h'(c_f) > 0$ for $c \in (c_f, \gamma)$; Thus either $h(c) = m(c) - c \geq h(c_f) = h(c_m) = m(c_m) - c_m = 0$ for $c \in (c_m, \gamma)$. Therefore, workers with expected health care cost that is equal to or greater than $c_m = c_f$ would choose to be insured and workers whose expected health care cost is less than $c_m = c_f$ would opt to be uninsured.

Conversely, the existence of a separating equilibrium implies there is $c \in (0, \gamma)$ that makes the incentive compatibility constraint $m(c) - (1-\alpha)c \geq 0$ holds; then for the lowest-cost workers, this constraint is binding, i.e., $m(c_m) - (1-\alpha)c_m = 0$. Because $\alpha = 0$, $m(c_m) - c_m = 0$,or $m(c_m) = c_m$. Hence c_m is the fixed point. **Q.E.D.**

This separating equilibrium, if it exists, is solely the result of workers' own choice because the firm does not have the incentive to enroll any of its workers but, on the other hand, it is unable to induce them to reject the insurance policy offer (withdrawing the policy is not an option in this model) since it does not have enough instruments. This may be called the worker-opted separating equilibrium, as compared to the firm-induced separating equilibrium that will be presented later.

If Proposition A4.2. is assumed to hold, except that the second derivative of the insurance benefit function is changed to $m''(c) < 0$, then there would be a different separating equilibrium where workers with expected health care cost that is equal to or less than $c_m = c_f$ would choose to be insured and workers whose expected health care cost is greater than $c_m = c_f$ would opt to be uninsured. This says that workers are risk-loving: they value the insurance coverage less than its cost beyond some point and their utility as a function of their health cost is concave.

Proposition A4.2 and the analysis in the preceding paragraph show that the type of the separating equilibrium depends on the shape of the health benefit function. There could be multiple separating equilibria if the insurance benefit function is not monotonic. *The existence of a fixed point is critical to the existence of a separating equilibrium in the case where insurance does not enhance productivity and wage is equal across insured and uninsured workers. If no fixed point exists, then only a pooling equilibrium may exist.*

Now suppose $\delta > 0$. If $m(c_m^*) = c_m^* = 0$, define $\frac{m(c_m^* \to 0)}{(c_m^* \to 0)^2} = \frac{m''(0)}{2}$ and $\frac{m'(c_m^* \to 0)}{c_m^* \to 0} = \frac{m''(0)}{1}$. Hence $\max\left(\frac{m(c_m^*)}{c_m^{*2}} - \frac{m'(c_m^*)}{c_m^*}, 0\right) = 0$; moreover, $\max(c_m^* - m(c_m^*), 0) = 0$; the first order conditions then becomes $0 < \delta$; this holds given a positive productivity differential. The solutions $\alpha^* = 0$.

If $m(c_m^*) = c_m^* > 0$, then $\frac{m(c_m^*)}{c_m^{*2}} - \frac{m'(c_m^*)}{c_m^*} = \frac{1}{c_m^*}\left(1 - m'(c_m^*)\right)$. Recall from discussions in the preceding paragraphs on negative productivity differential that $m'(c_m^*) > 1$ at the fixed point given $m''(c_m^*) > 0$. Thus the left-hand side of the first order condition is 0, or $0 < \delta$. This condition holds if and only if $\delta > 0$. Note that when $c_m^* = m(c_m^*) > 0$, $\alpha^* = 0$. This means that if there exists a cost type $c \in (0, \gamma]$ that solves the firm profit maximization profit and the insurance benefit function $m(c)$ has a fixed point at c_m^*, then the firm would not bear any share of insurance cost for its workers even if insurance coverage boosts worker productivity. This is because workers with expected health care at or above the optimal cutoff level would choose to be insured without being incentivized by the firm.

If $\mathrm{m}(c_m^*) > c_m^* > 0$, then $\alpha^* = 0$ and the first order condition becomes $0 < \delta$. [Note that for an increasing convex function $\mathrm{m}(\mathrm{c})$ that passes through the origin, $\frac{\mathrm{m}(c_m^*)}{c_m^{*2}} - \frac{\mathrm{m}'(c_m^*)}{c_m^*} < 0$].

In all cases above, including cases where productivity differential is negative, the firm's (excess) profit from insured workers can be generally expressed as $\int_{c_m}^{\gamma} \delta f(c)dc = \delta\left(1 - F(c_m^*)\right)$, because $\alpha^* = 0$. Note that in this case, it is also true that $\frac{\mathrm{m}(c_m^*)}{c_m^{*2}} - \frac{\mathrm{m}'(c_m^*)}{c_m^*} < 0$.

If $\mathrm{m}(\gamma) < c_m^* = \gamma$, the first order condition becomes $\frac{1}{f(c_m^*)}\left(\frac{m(\gamma)}{\gamma^2} - \frac{m'(\gamma)}{\gamma}\right)\int_{\gamma}^{\gamma} cf(c)dc +$ $\gamma - \mathrm{m}(\gamma) > \delta$, i.e.,

$$\gamma - \mathrm{m}(\gamma) > \delta.$$

which is possible, and the solution is $c_m^* = \gamma$ and $\alpha^* = 1 - \frac{m(\gamma)}{\gamma} > 0$. In the above two cases, the firm's (excess) profit from insured workers is $\int_{\gamma}^{\gamma} \delta(1 - \frac{m(\gamma)}{\gamma})\gamma f(c)dc = 0$.

If $\mathrm{m}(c_m^*) < c_m^* < \gamma$, then the first order condition can be written as

$$\frac{1}{f(c_m^*)}\left(\frac{m(c_m^*)}{c_m^{*2}} - \frac{m'(c_m^*)}{c_m^*}\right)\int_{c_m^*}^{\gamma} cf(c)dc + c_m^* - m(c_m^*) = \delta$$

This can hold because although the part in the parenthesis is negative, the last two term together are positive. The optimal c_m^* is determined implicitly by this equation. $\alpha^* = 1 - \frac{m(c_m^*)}{c_m^*} > 0$. This is a separating equilibrium that occurs below the 45-degree line since $\mathrm{m}(c_m^*) < c_m^*$. In this case, the firm's (excess) profit obtained from insured workers is

$$\int_{c_m^*}^{\gamma}\left(\delta - \left(1 - \frac{m(c_m^*)}{c_m^*}\right)c_m^*\right)f(c)dc = \left(\delta - \left(c_m^* - m(c_m^*)\right)\right)\left(1 - F(c_m^*)\right)$$

which depends on δ. This constitutes a separating equilibrium, which may be called firm-induced separating equilibrium because it results primarily from the firm's decision.

The following proposition summaries and generalizes an important result: conditions for the existence of a separating equilibrium.

Proposition A4.3. *Suppose the health benefit function is increasing and convex over the domain of its argument. Then an induced separating equilibrium exists if and only if health insurance boosts productivity and* $m'(0) < 1$**.**

Proof. If $m'(0) < 1$, then with a small $c > 0$ in the neighborhood of zero, $m(c) < c$; with $0 < \delta$, the firm can choose $\alpha > 0$ so that $\delta \geq \alpha c$,where c can also be chosen in such a way that $m(c) = (1 - \alpha)c$ is satisfied. This α solves the firm's profit maximization problem, as shown earlier. Conversely, if an induced separating equilibrium exists, then $\alpha > 0$ and the health insurance must boost productivity for otherwise the firm would set $\alpha = 0$; moreover, the existence of an induced separating equilibrium also implies $m(c) < c$ for some c, for otherwise all workers would choose to be insured and no incentive-compatible scheme the firm may utilize to induce worker to act otherwise, and thus the equilibrium would be pooling.

If the benefit function has a fixed point, i.e., $m(c_f) = c_f$, then $m(c) > c$ for $c > c_f$, given that the function is increasing and convex. Then all workers with a cost higher the fixed point would choose to be insured without being incentive by the firm, and because $m(c) > (1 - \alpha)c$ for all $\alpha \in [0,1]$, there are no incentive-compatible schemes the firm may utilize to induce worker opt out of insurance; thus a worker-opted separating equilibrium has to exist. Conversely, if a work-opted separating equilibrium, then $m(c) \geq c$ for some $c \in [0, \gamma]$; the minimum level of c,

c_m, must satisfy this condition with equality, i.e., $\mathrm{m}(\mathrm{c_m}) = \mathrm{c_m}$. Thus a fixed point exists, which is $\mathrm{c_m}$. **Q.E.D.**

To sum up, when the firm offers only one insurance plan and it is required to pay equal wages to all its workers, then some workers may choose to be insured, even if the firm does not share the cost of health insurance. The reason is that some workers value the benefit of the health insurance plan more than its cost. In such a case, the firm may suffer a loss in excess of the wage cost because insurance has negative externality. However, if the externality of health insurance is positive, the firm can benefit from all the health insurance without doing anything under certain circumstances. If the workers do not value the health insurance plan as much as it costs, but health insurance may lead to higher productivity, then the firm will be willing to share the cost of insurance. The important conclusion is that the health insurance benefit function affects the firm's decisions and profitability. It also determines, in many cases, the type of equilibrium. For example, if this function rises too fast relative to the cost, then a large portion of workers would choose to be insured, even if no particular incentives are provided by the firm. In this situation, the firm essentially loses its ability to affect workers' health insurance decisions, because the instruments available to it no longer work. The firm can influence its workers' choice only when the benefit from health insurance is low relative to its cost for a substantial portion of workers. If the one-wage restriction is relaxed by assuming the firm offers two (potentially) different wages, then the problem just described can be solved: the firm can adjust both the wage associated with the insurance policy and the share of cost it bears. However, this would be very similar to the one-plan model analyzed in Section 3.4 of Chapter 3.

For rents distribution, the firm extracts all rents when the equilibrium is low-end pooling (all workers are uninsured), but receives no workers receive rents because the individual rationality constraint would be binding for all workers. If the equilibrium is separating, then the firm extracts rents (only) from uninsured workers who receive no rents, but the insured workers receive rents, since in this case, the participation constraint for uninsured workers is binding, whereas the participation constraint for the insured ones is not. If the equilibrium is high-end pooling, then the firm extracts no rents and all workers except for the lowest-cost ones receive rents, because the participation constraint is binding only for the lowest-cost workers but the lowest-cost workers have no risk and thus cannot generate rents.

APPENDIX D

APPENDICES FOR CHAPTER VI

A6.1. Tables

Table A6.1. Results from the Wage Regression Model with Two Excluded Instruments

VARIABLES	(1) OLS RSE	(2) 1st Stage	(3) 2SLS	(4) 1st Stage	(5) 2SLS	(6) 1st Stage	(7) 2SLS
Insured via firm	0.235***		-0.278***		-0.186***		-0.201***
	(0.005)		(0.032)		(0.030)		(0.033)
Male	0.285***	0.148***	0.364***	0.145***	0.350***	0.151***	0.352***
	(0.006)	(0.004)	(0.007)	(0.004)	(0.007)	(0.004)	(0.007)
Firm with >= 100 Employees	0.130***	0.218***	0.235***	0.207***	0.216***	0.221***	0.219***
	(0.005)	(0.004)	(0.008)	(0.004)	(0.008)	(0.004)	(0.008)
Unionized (yes = 1)	0.125***	0.249***	0.253***	0.264***	0.230***	0.269***	0.233***
	(0.007)	(0.006)	(0.010)	(0.006)	(0.010)	(0.006)	(0.011)
Spouse offered insurance		-0.105***		-0.112***			
		(0.004)		(0.004)			
Spousal Firm w/ >= 100 employees		-0.072***				-0.093***	
		(0.004)				(0.004)	
Spouse unionized				-0.115***		-0.129***	
				(0.006)		(0.006)	
Constant	2.302***	0.502***	2.508***	0.496***	2.471***	0.451***	2.477***
	(0.014)	(0.012)	(0.018)	(0.012)	(0.018)	(0.011)	(0.018)
F-stat (Excl. Instruments)							
Sargan (2SLS)/J (GMM)			1.794				
			(p=0.181)				
Hausman (2SLS)/C (GMM)			166.687				
			(p=0.0001)				
Observations	59,442	59,442	59,442	59,442	59,442	59,442	59,442
R-squared	0.358	-	0.201	-	0.252	-	0.245

Note: 1) Robust standard errors in parentheses: *** p<0.01, ** p<0.05, * p<0.1
2) Control variables not listed in the table include Occupation, Race, Potential Experience, Region Dummies, and Year Dummies.
3) Sample: MEPS 2002-2014.

Table A6.2. Results from Wage Regression Models Using Husband and Wife Samples; h=Husband; w=Wife

VARIABLES	(1) OLS RSE (h)	(2) 1st Stage (h)	(3) 2SLS (h)	(4) OLS RSE (w)	(5) 1st Stage (w)	(6) 2SLS (w)
Insured via firm	0.273***		0.117**	0.279***		-0.121***
	(0.008)		(0.0491)	(0.007)		(0.036)
Firm with >= 100 employees	0.118***	0.231***	0.153***	0.165***	0.232***	0.252***
	(0.008)	(0.006)	(0.012)	(0.008)	(0.006)	(0.010)
Unionized	0.035***	0.218***	0.069***	0.143***	0.252***	0.245***
	(0.010)	(0.008)	(0.014)	(0.0116)	(0.009)	(0.013)
<u>Excl. Instruments</u>		-			-0.131***	
		0.086***				
Spouse offered insurance		(0.006)			(0.006)	
		-			-0.084***	
		0.069***				
Spousal firm w/ >= 100 employees		(0.006)			(0.006)	
	1.317***	0.168***	1.339***	1.220***	0.205***	1.283***
Constant	(0.028)	(0.021)	(0.024)	(0.026)	(0.020)	(0.023)
	0.273***		0.117**	0.279***		-0.121***
F-stat (Excl. Instruments)	-	176.94	-	-	174.10	-
Sargan (2SLS)/J (GMM)	-	-	7.10	-	-	0.0009
	-	-	(p=0.008)	-	-	(p=0.992)
Hausman (2SLS)/C (GMM)	-	-	5.28	-	-	71.86
Sargan (2SLS)/J (GMM)	-	-	(p=0.005)	-	-	(p=0.0001)
Observations	28,568	28,568	28,568	30,590	30,590	30,590
R-squared	0.286	0.130	0.271	0.319	0.120	0.215

Note: 1) Robust standard errors in parentheses: *** p<0.01, ** p<0.05, * p<0.1
2) Control variables not listed in the table include Spousal Education, Race, Potential Experience, Region Dummies, and Year Dummies.
3) Sample: MEPS 2002-2014

Table A6.3. Results from the Wage Regression Model with Three Excluded Instruments

VARIABLES	OLS RSE	1st Stage	2SLS	GMM
Insured via firm	0.227***		-0.116***	-0.114***
	(0.005)		(0.026)	(0.026)
Male	0.207***	0.122***	0.253***	0.252***
	(0.005)	(0.004)	(0.005)	(0.005)
Firm with >= 100 employees	0.124***	0.219***	0.195***	0.195***
	(0.005)	(0.004)	(0.007)	(0.007)
Unionized	0.079***	0.242***	0.158***	0.157***
	(0.007)	(0.006)	(0.008)	(0.008)
Excl. Instruments				
Spouse offered insurance		-0.097***		
		(0.004)		
Spousal Firm w/ >= 100 employees		-0.069***		
		(0.004)		
Spouse unionized		-0.106***		
		(0.006)		
Constant	0.862***	-0.028*	0.830***	0.830***
	(0.020)	(0.015)	(0.016)	(0.017)
F-stat (Excl. Instruments)				
Sargan (2SLS)/J (GMM)			6.20	6.39
			(p=0.045)	(p=0.041)
Hausman (2SLS)/C (GMM)			103.53	205.34
			(p=0.0001)	(p=0.0001)
Observations	59,442	59,442	59,442	59,442
R-squared	0.395	0.167	0.324	0.325

Note: 1) Robust standard errors in parentheses: *** p<0.01, ** p<0.05, * p<0.1
2) Control variables not listed in the table include Education, Race, Potential Experience, Region Dummies, and Year Dummies.
3) Sample: MEPS 2002-2014.

APPENDIX E

APPENDICES FOR CHAPTER VII

A7.1. Tables

Table A7.1. Results from Two-part Model for #Days Missed Work Due to Sickness Using the Full Sample

VARIABLES	1st part Prob.	2nd part OLS	1st eqn. Bi Prob.	2nd eqn. Bi Prob.	1st part IV Prob.	2nd part 2SLS	2nd part GMM	2nd part Pois GMM
Insured via firm	0.218***	-0.016		-0.245***	-0.403***	-0.581	-0.564	-0.087
	(0.011)	(0.122)		(0.0618)	(0.0695)	(0.665)	(0.664)	(0.101)
Male	-0.257***	-1.359***	0.358***	-0.190***	-0.165***	-1.301***	-1.301***	-0.197***
	(0.011)	(0.115)	(0.011)	(0.014)	(0.0154)	(0.133)	(0.128)	(0.020)
Firm with >=100 employees	0.060***	0.723***	0.624***	0.154***	0.185***	0.830***	0.815***	0.120***
	(0.011)	(0.117)	(0.012)	(0.016)	(0.017)	(0.171)	(0.171)	(0.026)
Unionized	0.0747***	1.409***	0.715***	0.178***	0.212***	1.531***	1.539***	0.219***
	(0.016)	(0.165)	(0.018)	(0.021)	(0.021)	(0.217)	(0.224)	(0.032)
Spouse offered insurance			-0.327***					
			(0.012)					
Spousal Firm w/ >= 100 employees			-0.223***					
			(0.0121)					
Constant	-0.159***	11.90***	-1.493***	-0.201***	-0.210***	11.90***	11.87***	2.604***
	(0.041)	(0.448)	(0.043)	(0.041)	(0.041)	(0.448)	(0.457)	(0.063)
Observations	59,442	24,810	59,442	59,442	59,442	24,810	24,810	24,810
R-squared		0.022	-			0.021	0.021	

Note: 1) Robust standard errors in parentheses: *** p<0.01, ** p<0.05, * p<0.1
2) For the bivariate Probit (Bi Prob.) models, rho=0.285***.
3) Control variables not listed in the table include Education, Race, Potential Experience, Region Dummies, and Year Dummies.
4) Sample: MEPS 2002-2014.

Table A7.2. Results from Regression of #Days Missed Work Due to Sickness Using the Full Sample, where Spousal Education Replaces Education.

Variables	(1) OLS RSE	(2) 1st Stage	(2') 2SLS	(3) GMM	(4) Pois GMM
Insured via firm	0.533***	-	-1.498***	-1.494***	-0.496***
	(0.062)	-	(0.384)	(0.384)	(0.124)
Male	-1.199***	0.111***	-0.951***	-0.951***	-0.353***
	(0.059)	(0.004)	(0.072)	(0.071)	(0.025)
Firm with >=100 employees	0.453***	0.231***	0.898***	0.897***	0.307***
	(0.063)	(0.004)	(0.102)	(0.102)	(0.033)
Unionized	0.809***	0.234***	1.287***	1.287***	0.424***
	(0.102)	(0.006)	(0.123)	(0.128)	(0.041)
Excl. Instruments					
Spouse offered insurance	-	-0.109***	-	-	-
	-	(0.004)	-	-	-
Spousal firm with >=100 employees	-	-0.077***	-	-	-
	-	(0.004)	-	-	-
Constant	4.609***	0.139***	4.803***	4.801***	1.669***
	(0.223)	(0.014)	(0.211)	(0.216)	(0.069)
F-stat (Excl. Instruments)	-	389.58	-	-	-
Sargan (2SLS)/J (GMM)	-	-	0.125	0.122	0.201
		-	(p=0.724)	(P=0.727)	(P=0.653)
Hausman (2SLS)/C (GMM)	-	-	14.60	29.37	-
	-	-	(p=0.0001)	(p=0.0001)	-
Observations	59,158	59,158	59,158	59,158	59,158
R-squared	0.016	0.141	-	-	-

Note: 1) Robust standard errors in parentheses: *** p<0.01, ** p<0.05, * p<0.1
2) Control variables not listed in the table include **Spousal Education**, Race, Potential Experience, Region Dummies, and Year Dummies.
3) Sample: MEPS 2002-2014.

Table A7.3. Results from Regression of #Days Missed Work for Other Reasons Using the Full Sample, where Spousal Education Replaces Education.

Variables	(1) OLS RSE	(2) 1st Stage	(2') 2SLS	(3) GMM	(4) Pois GMM
Insured via firm	0.193***	-	-0.545***	-0.552***	-0.576***
	(0.025)	-	(0.154)	(0.152)	(0.155)
Male	-0.363***	0.111***	-0.273***	-0.271***	-0.333***
	(0.024)	(0.004)	(0.029)	(0.028)	(0.033)
Firm with >=100 employees	0.064**	0.231***	0.226***	0.230***	0.248***
	(0.025)	(0.004)	(0.041)	(0.040)	(0.042)
Unionized	0.139***	0.234***	0.313***	0.317***	0.352***
	(0.039)	(0.006)	(0.049)	(0.050)	(0.054)
Excl. Instruments					
Spouse offered	-	-0.110***	-	-	-
	-	(0.004)	-	-	-
Spousal firm with >=100 employees	-	-0.076***	-	-	-
	-	(0.004)	-	-	-
Constant	0.682***	0.139***	0.752***	0.743***	-0.522***
	(0.084)	(0.014)	(0.085)	(0.082)	(0.099)
F-stat (Excl. Instruments)	-	446.86	-	-	-
Sargan (2SLS)/J (GMM)	-	-	7.36	7.18	5.74
	-	-	(0.007)	(P=0.007)	(p=0.017)
Hausman (2SLS)/C (GMM)	-	-	11.99	25.17	-
	-	-	(p=0.0001)	(p=0.0001)	-
Observations	59,158	59,158	59,158	59,158	59,158
R-squared	0.014	0.141	-	-	-

Note: 1) Robust standard errors in parentheses: *** $p<0.01$, ** $p<0.05$, * $p<0.1$

2) Control variables not listed in the table include **Spousal Education**, Race, Potential Experience, Region Dummies, and Year Dummies.

3) Sample: MEPS 2002-2014.

Table A7.4. Results for Regression of #Days Missed Work for Other Reasons for the Husband and Wife Samples, where Spousal Education replaces Own Education. h=Husband, w=Wife.

Variables	(1) OLS RSE (h)	(2) 1st Stage (h)	(2') 2SLS (h)	(3) OLS RSE (w)	(4) 1st Stage (w)	(4') 2SLS (w)
Insured via firm	0.163***		-0.524**	0.218***		-0.568***
	(0.030)		(0.233)	(0.038)		(0.204)
Firm with >=100 employees	0.116***	0.231***	0.266***	0.013	0.232***	0.185***
	(0.033)	(0.006)	(0.059)	(0.039)	(0.006)	(0.057)
Unionized	0.090*	0.218***	0.241***	0.217***	0.252***	0.417***
	(0.048)	(0.008)	(0.065)	(0.065)	(0.009)	(0.075)
Excl. Instruments						
Spouse offered	-	-0.0804***	-	-	-0.081***	-
	-	(0.006)	-	-	(0.006)	-
Spousal Firm w/ >=100 employees	-	-0.067***	-	-	-0.134***	-
	-	(0.006)	-	-	(0.006)	-
Constant	0.583***	0.075***	0.606***	0.347***	-0.059***	0.288**
	(0.101)	(0.021)	(0.109)	(0.126)	(0.021)	(0.130)
F-stat (Excl. Instruments)		176.94		-	174.10	-
Sargan (2SLS)/J (GMM)	-	-	1.98	-	5.14	-
	-	-	(p=0.160)	-	(p=0.023)	-
Hausman (2SLS)/C (GMM)	-	-	4.49	-	7.81	-
	-	-	(p=0.011)	-	(p=0.0004)	-
Observations	28,568	28,568	28,568	30,590	30,590	30,590
R-squared	0.017	0.130	-	0.009	0.120	-

Note: 1) Robust standard errors in parentheses: *** p<0.01, ** p<0.05, * p<0.1
2) Control variables not listed in the table include **Spousal Education**, Race, Potential Experience, Region Dummies, and Year Dummies.
3) Sample: MEPS 2002-2014.