Essays on Distortionary Effects of Employer-Sponsored Health Insurance

by

Nastassia Krukava

A Dissertation Presented in Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy

Approved April 2017 by the
Graduate Supervisory Committee:

Galina Vereshchagina, Chair
Berthold Herrendorf
Gustavo Ventura

ARIZONA STATE UNIVERSITY

May 2017

ABSTRACT

This dissertation consists of two chapters. Chapter one studies distortionary effects of tax exemption of employer-sponsored health insurance (ESHI) premiums. First, I argue that, in the competitive labor market, tax deductibility of ESHI premiums generates an implicit labor cost subsidy to the employers sponsoring health insurance (HI) which distorts the allocation of labor across employers. Second, I quantify the extent of this misallocation measured as output loss in a general equilibrium model of firm dynamics extended to incorporate tax exemption of ESHI premiums and endogenous provision of HI by the employers. The calibrated model shows that elimination of tax exemption increases aggregate output by 1.73%. About two-thirds of this effect comes from removing the misallocation of labor across existing establishments, and the remaining one-third comes from the increase in the number of operating establishments. Third, I use the model to analyze how tax exemption interacts with the employer mandate of the Affordable Care Act imposing a tax on large employers not sponsoring HI. Quantitative results show that implementing the employer mandate when the tax exemption is present reduces output by 0.13%.

Chapter two studies macroeconomic implications of a higher cost of health services faced by the unemployed which arise because 1) workers lose access to ESHI when they leave their jobs and 2) the uninsured face inflated health care prices. First, I provide evidence suggesting that the cost of health services for the privately insured is about 50% lower than for the uninsured. Second, I quantify the effects of higher cost of health services for the unemployed in the Lucas and Prescott (1974) island model extended to allow the workers to pay an extra cost of health services contingent on their employment status. Calibration procedure uses the differences between costs of health services for the privately insured and uninsured inferred from the data as a gap between costs of health services for the employed and unemployed. Quantitative results show that equalizing these costs across workers increases labor productivity by 1.2% and unemployment rate by 1.5 percentage points. The increased unemployment dominates quantitatively leading to a decrease in aggregate output by 0.26%.

ACKNOWLEDGMENTS

I wish to express my warmest gratitude to the members of my dissertation committee, Galina Vereshchagina, Berthold Herrendorf and Gustavo Ventura, and all the faculty at the Department of Economics at Arizona State University whose comments, questions, and ability to critically evaluate economic research, support and encouragement have served as a great input into this work.

First and foremost, I want to thank my advisor Galina Vereshchagina. I very much appreciate her patience and contributions of time and effort in helping me develop as a researcher. I also appreciate her encouragement, both personal and academic. Our conversations were always inspirational, pushing me forward to better understand the forces behind my findings and to ask new questions. I am very thankful for the guidance, constant feedback on my work that she provided during the years of my Ph.D. studies, and optimism she kindly shared with me.

I would like to thank two other members of my dissertation committee, Berthold Herrendorf and Gustavo Ventura, for their guidance, incredibly motivational questions, helpful comments and invaluable suggestions as to how to communicate my ideas and results to the audience. I appreciate that they continually shared their vision of the questions which I raised in my research and provided a supportive and stimulating work environment.

My sincerest thanks and warmest regards go to Alexander Bick, Domenico Ferraro, Natalia Kovrijnykh, Nicolai Kuminoff, Edward Schlee, Todd Schoellman, Daniel Silverman and Matthew Wiswall for insightful discussions which helped me move forward.

I would also like to thank all my classmates and friends at Arizona State University for helpful conversations and comments on my work, as well as personal support through ups and downs during the years of my Ph.D. I am especially grateful to Xican Xi, Jaehan Cho, and Zhizhuang Ge for sharing interest in my work.

Lastly, but most importantly, I would like to thank my family, especially my parents and my brother, for their love and encouragement. Though we have been far away from each other, I have always felt their faithful support. This was the most valued help in the long journey of pursing my Ph.D.

TABLE OF CONTENTS

CHAPTER Page

LIST OF TABLES

Table Page

LIST OF FIGURES

Chapter 1

DISTORTIONARY EFFECTS OF TAX EXEMPTION FOR EMPLOYER-SPONSORED HEALTH INSURANCE PREMIUMS

1.1 Introduction

Under the current U.S. tax code, health insurance (HI) premiums are exempted from determination of the workers' income and payroll taxes [1] if workers obtain HI through the employer. In contrast, if workers buy HI on their own through the individual insurance market, HI premiums are paid from the after-tax income. In the competitive labor market, this implies that the employers sponsoring HI face a lower price of labor than those not doing it. Even though tax deductibitility of ESHI is one of the largest tax breaks in the tax code [2] , not all employers take advantage of it. For example, according to the Employer Health Benefits Survey (2015) conducted by the Kaiser Family Foundation, only 57% of firms offered health benefits to their employees in 2015. Such heterogeneity in HI provision coupled with the implicit subsidy to the employers sponsoring HI leads to a misallocation of labor across firms and a reduction of aggregate output. The objective of this paper is to quantify the extent of this misallocation. More specifically, I ask what would be the effect of eliminating the tax exemption of the ESHI premiums on the aggregate output.

To address this question, I extend a general equilibrium model of firm dynamics as in Hopenhayn (1992) by incorporating the tax exemption of the premiums paid by

[1] These include federal and state income taxes, Social Security and Medicare taxes.

[2] Total federal tax revenue foregone because of the exclusion of the ESHI was $250 billion in 2013 which accounted for 1.5% of GDP (estimate is from Options for Reducing The Deficit: 2014 to 2023 (2013)).

the employers and allowing the employer to choose whether or not sponsor HI. In the model, provision of HI is associated with the following trade-off faced by the employer. On the one hand, the employer must incur a fixed cost of setting up insurance plans, bargaining with insurance companies, etc. On the other hand, provision of HI allows the employer to lower the gross wage bill because ESHI premiums are deducted from taxes. Put it differently, if the worker does not receive HI through the workplace and instead buys it on the individual insurance market, his gross wage must be higher to compensate for additional taxes paid on the HI premium. Thus, in the competitive labor market, where the workers are indifferent between employers of different HI provision status, the tax exemption of ESHI translates into an implicit subsidy to the employers sponsoring HI [3] . Using the 1997 Robert Wood Johnson Foundation Employer Health Insurance Survey data (Long and Marquis (1997)), I estimate that such a subsidy is equivalent to a wage reduction of almost 10 percent. I then use widely documented facts on establishment dynamics, as well as the data from the aforementioned Survey to calibrate the model to match the key empirical patterns of the establishment's size distribution and HI provision among employers in the U.S. To answer the main question of this paper, I completely eliminate the tax exemption of the ESHI premiums within the calibrated model and find that the aggregate output increases by 1.73%.

Two features of the equilibrium firm dynamics model are crucial for accurately evaluating the distortionary effect of the tax exemption of the ESHI premiums: (i) establishments' heterogeneity; (ii) a life cycle of the establishments. The importance

[3]How much the gross wage can be lowered also depends on the utility that the worker attaches to the HI obtained from different sources. To concentrate only on the distortionary effects of the tax exemption of the ESHI premiums, I assume that the cost and the quality of HI do not depend on whether it is provided by the employer or purchased individually. Typically, HI in the individual insurance market would be more expensive which, as I argue in Section 1.4.2, further increases the size of distortion.

of these features is rooted in the empirical pattern of HI provision among employers. In particular, I document that small employers are less likely to sponsor HI than the large ones. For example, only about 34% of establishments with up to 5 employees sponsor HI, while 98% of those with more than 100 employees do so. This positive correlation between HI provision and the size implies that the tax exemption is, in effect, a size-dependent policy reallocating labor resources across existing establishments (intensive margin). Previous studies [4] analyzing size-dependent policies in other contexts suggest that an impact of the policy on the output depends on the extent of the correlation between introduced taxes/subsidies and the employer size. Therefore, the ability of the model to capture HI provision across establishments of different sizes is essential. At the same time, establishment's life cycle in conjunction with the described pattern has important implications for the number of establishments in operation determined by the entry and exit (extensive margin). On the one hand, because the large establishments are more likely to sponsor HI, the subsidy increases profits of the large establishments relative to the small ones. On the other hand, establishments enter the economy small and grow large over time. Because of discounting, relatively larger profits received when the establishment grows large vanish in the entrant's expected life-time profit. As a result, the entrant's value decreases relative to the average incumbent which discourages entry. Similar argument applies to the small establishments experiencing bad shocks which implies that the subsidy affects the incentives to exit. Quantitatively, I find that this channel plays a substantial role: about one-third of 1.73% gain in the output is obtained on the extensive margin through the changes in the entry and exit.

In this paper, I also study the effects of the employer mandate, a component of the ongoing Patient Protection and Affordable Care Act of 2010 (ACA) reform, when

[4] For example, Restuccia and Rogerson (2008), Guner *et al.* (2008), Hopenhayn (2012).

the tax exemption of the ESHI premiums is in place. In practice, it introduces a tax on the large employers not sponsoring HI. On its own, such tax is obviously distortionary because it imposes a wedge between labor costs faced by the employers. But in interaction with the tax exemption of the ESHI premiums, the effects may be non-trivial. On the one hand, since the tax applies to only the large employers, it reallocates profits in the direction opposite to that induced by the tax exemption. Specifically, it increases profits of the small establishments relative to the large ones and can offset, at least partially, a negative effect of the tax exemption on the extensive margin. On the other hand, the tax increases the gap between the marginal costs of labor faced by the large establishments sponsoring and not sponsoring HI. This intensifies the extent of the misallocation across establishments within the same productivity class and can enforce more distortion on the intensive margin. I quantify these two forces in the calibrated model and find that the latter one dominates. As a result, the aggregate output decreases by 0.13% compared to the economy featuring the tax exemption only.

While this paper isolates an uncontroversial implication of the tax exemption of the ESHI premiums resulting in the distortion on the output, caution must be taken in interpreting the results. For the most part, because I analyze the distortion on the production side of the economy, the focus is to model the establishment's decisions and to replicate a complex relationship between the employer size and HI provision observed in the data. In contrast, the consumer's side is kept very parsimonious yet allowing me to realistically represent the U.S. economy. Because of that, many benefits of ESHI to the individuals are not taken into account which prevents me from drawing any welfare comparisons. For example, one could argue that by decreasing the number of individuals receiving HI through the workplace, elimination of the ESHI system reduces the welfare. Given this concern, my findings should be understood that any alternative system able to cover the individuals covered under ESHI without distorting the marginal costs of labor may result in the output increase by as much as 1.73%. More meaningful welfare comparisons accounting for the changes in the coverage would require to expand the consumer's side in the model.

This paper is related to several strands of literature. On the health insurance literature side, my paper relates to the group of papers studying implications of the tax exemption of the ESHI premiums. For example, Jeske and Kitao (2009) focus on the regressive nature of the ESHI effectively providing larger subsidies to the individuals with higher income and the implications of that for the cost of the premiums, HI coverage and welfare. Huang and Huffman (2014) examine how a combination of the shifts in the consumption of medial services relative to other goods and changes in the worker's movements in and out of employment/unemployment due to the tax exemption affects welfare. These papers concentrate exclusively on the implications of the tax exemption of the ESHI premiums but they construct models which fully abstract from the distribution of HI provision across employers of different sizes. Thus, they also abstract from the key mechanisms evaluated in this paper. A contrasting example is a paper by Brugemann and Manovskii (2009). They design a model with endogenous choice of HI provision by the employers and directly address heterogeneity in the employer's HI provision and size. Though they incorporate tax deductibility of the ESHI premiums as a component of the regulatory environment of the model, the focus is on the effects of the ACA in the quantitative analysis. To sum up, my contribution to this literature is that I draw the attention and quantitatively evaluate the distortion created by the differential tax treatment of the HI premiums on the production side of the economy which was previously overlooked in the literature.

Methodologically, the paper can be considered as a case study complementary to the strand of the literature on the misallocation of resources across heterogeneous firms. In these studies, regulations take many forms, from generic family of distortions reflected in different taxes/subsidies faced by firms to more specific policies. In the first case, the rationale for considering generic policies, as in Restuccia and Rogerson (2008), Bhattacharya *et al.* (2013), Gabler and Poschke (2013), is often an observation that smaller firms face lighter regulations as compared with large firms. These papers usually examine a very wide range of the taxation schemes corresponding to different extent of distortions and, hence, often find substantial negative effects on the aggregate output. For example, Restuccia and Rogerson (2008) calculate that taxing 50% of the establishments at 40% reduces aggregate output by 31% (case of a correlated distortion). In the second case, the studies concentrate on measuring distortions of empirically-plausible policies. Among these studies are, for example, Hopenhayn and Rogerson (1993) and Veracierto (2000), who study the policies making more costly for the firms to adjust the employment level over time; Vereshchagina (2005), who studies the effects of targeted employment subsidy programs under which firms receive subsidies for hiring disadvantaged workers; Buera *et al.* (2013), who study the effects of the policy providing individual-specific subsidies to productive entrepreneurs that remain fixed over time, and many others. These papers usually find milder disruptive effects on the aggregate output. For example, Roys and Gourio (2013) evaluate that eliminating the exemption of the small firms with less than 50 employees from the regulations in France leads to an increase in the output by less than 0.3% when labor supply and number of firms is fixed [5] . In this respect, my work finds a distortionary effect which, quantitatively, is not an outlier. All in all, my contribution to this lit-

[5] An environment comparable to this paper to a larger extent includes inelastic labor supply and elastic entry. Roys and Gourio (2013) find that elimination of the exemption of the small firms from the regulation in this environment yields a steady-state output loss of 0.02%. The output loss from removing a distortionary policy should not, however, come as a surprise because this result does not account for the changes in the output along transition to the new steady state.

erature is that I study the distortionary effects of the policy not directly comparable to those studied in the mentioned papers.

The rest of the paper is organized as follows. In Section 2.4, a model with heterogeneous employers and endogenous HI provision is set up. Section 1.3 covers a calibration procedure and explores quantitative results from contractual experiments. The experiments cover elimination of the tax exemption of the ESHI premiums and introduction of the ACA employer mandate in the presence of the subsidy generated by the tax exemption. Section 1.4 provides the discussions omitted in the previous sections which, however, are important remarks to the main findings. Section 2.6 concludes.

1.2 The Model

In this section, I describe a quantitative model based on Hopenhayn (1992) framework. To characterize the responses of the economy to changes in the tax regulation of the ESHI premiums, the model departs from Hopenhayn (1992) in letting establishments choose whether or not to sponsor HI for their employees. In equilibrium, establishments sponsoring insurance benefit from an implicit labor cost subsidy that generates a distortion to the allocation of labor resources across production units and the number of operating establishment which I quantify in the next section.

The time is discrete in the model. There are two types of entities: establishments, which produce the consumption good, and consumers, who supply labor services to establishments. Two markets that operate in every period are the consumption good and labor markets.

The main focus here is to study long run effects on aggregate output of alternative tax treatments of the ESHI premiums. Thus, I state the consumer's and the establishment's problems in the environment with constant prices and define a stationary competitive equilibrium below. Consumption good is a numeraire with the price normalized to one.

1.2.1 Establishments

Production side of the economy is represented by heterogeneous production units or establishments.

Every period establishments experience idiosyncratic productivity shock, $s \in [\underline{s}, \bar{s}] \subset R_+$. Shocks are independent across establishments and follow the same first-order Markov process given by a function $F(s'|s)$ where $F(\cdot|s)$ denotes the distribution function for the next period's shocks s' for each value of the current period shock s.

Establishment of productivity s uses n units of labor to produce consumption good in accordance with a production function $f(s,n)$ which exhibits decreasing return to scale in n:

$$f(s,n) = sn^{\theta}, \qquad \theta \in (0,1).$$

Each establishment has to pay fixed cost of operation c_f measured in the units of output every period in which it produces. This cost is averted if establishment exits the market. Establishments operating in the current period survive in the next period with exogenous probability γ. Besides, all establishments surviving exogenous exit may voluntarily leave the market before the next period's information is revealed. Endogenous and exogenous exit combined determine overall exit rate in the economy.

In every period, there is an unlimited mass of potential ex-ante identical entrants which can enter the market after paying c_e units of consumption good as an entry cost. Once c_e is paid, entrants independently draw their initial productivity from the invariant distribution with the c.d.f. $G(s)$, $s \in [\underline{s}, \bar{s}]$.

Labor market is competitive while establishments are distinguished by the structure of the compensation package they offer in exchange for labor services. Besides paying wages, establishments may choose to sponsor HI for their employees purchased in a group health insurance marketplace at a cost H^G. I assume that insurance markets are not present in the model. Instead, establishments shop for HI in the marketplace at a fixed price, or HI premiums, H^G measured in the units of the numeraire good. Establishments cannot discriminate and select to cover only a portion of their workers. Thus, the establishments are constrained to offer HI to all employed workers in case of provision. Setting up a coverage for the establishment is associated with a random fixed cost $c_h \in [\underline{c}_h, \bar{c}_h] \subset R_+$ to be paid in every period when the coverage is offered. This cost represents a fixed cost of administering insurance plans, bargaining with insurance companies, collecting premiums, billing, advertising, any type of broker fees, etc [6] . It is measured in the units of output and wasted from the standpoint of the economy. Cost c_h is independent across establishments and follows a Markov process with a transition function $P(c_h'|c_h)$. Entrant's initial value of c_h is drawn from the distribution function $Q(c_h)$. Wages of establishments of different provision statuses are denoted w_0 and w_1 where subscripts 0 and 1 are, correspondingly, for establishments not sponsoring and sponsoring HI. The compensation package of the establishment having decided not to sponsor HI consists of the wage w_0 which represents per employee gross wage and the marginal cost of labor for this establishment. The compensation package of the establishment having decided to sponsor HI, (w_1, H^G), consists of the wage w_1 and HI H^G. Thus, the marginal cost of labor for the

[6]For the employers operating their own health plan as opposed to purchasing a plan from an insurance company, so called self-insured plan, c_h cost may also include premiums for stop-loss insurance which reimburses the employer for claims that exceed a predetermined level; any costs billed by the third-party administrator or a carrier that processes insurance claims and other aspects of the employee's HI.

establishment sponsoring HI is determined by the total cost of the package (w_1, H^G). Wages w_0 and w_1 are determined in equilibrium which is described later.

Events within a period unfold as follows. At the beginning of period t, $1-\gamma$ fraction of incumbents at each state leaves the market according to the exogenous exit rate. Next, surviving incumbents decide to exit or stay (endogenous exit). If establishment exits, it does not produce or pay fixed costs in subsequent periods. If establishment chooses to stay, it draws a productivity shock s from $F(s|s_{-1})$ and fixed cost of HI provision c_h from $P(c_h|c_{h_{-1}})$. Then, decisions about current employment n and HI provision χ are made, fixed costs c_f and, if insurance is offered, c_h are paid, production takes place and final good is sold to the consumers. As for the entrants, they are required to pay a one-time cost of entry c_e before values of shocks are revealed to them. Upon entry, the order of events in the first period is the same as for the establishments having already decided to stay, i.e., entrants cannot exit in the first period after they observe the shocks and produce at least one period. Also, productivity and fixed cost of setting up the insurance for entrants are drawn from G and Q, correspondingly. In subsequent periods, entrants join the pool of incumbents.

To formally state the establishment's decision problem, one needs to recognize that establishments in this economy are indexed by the productivity level s and fixed cost of setting up the insurance c_h. Note also that, whereas current values of s and c_h affect distributions of next periods shocks, employment and HI provision decisions do not change future paths of shocks. Thus, optimal employment and provision decisions of the establishment in the state (s, c_h) are derived from the static profit maximization problem which I specify below.

Instantaneous profit of the establishment in the state (s, c_h) hiring n units of labor are given by:

- if the establishment does not sponsor HI

$$\pi_0(s, c_h, n; w_0, w_1) = \; sn^\theta - w_0 n - c_f;$$

- if the establishment sponsors HI

$$\pi_1(s, c_h, n; w_0, w_1) = \; sn^\theta - (w_1 + H^G)n - c_f - c_h.$$

As before, lower subscripts 0 and 1 are to distinguish establishments not sponsoring and sponsoring HI.

Then, optimal employment and provision decisions of the establishment in the state (s, c_h) are derived as follows:

- employment of the establishment having decided to not sponsor or sponsor HI is deduced from, correspondingly,

$$\begin{aligned} \pi_0^*(s, c_h; w_0, w_1) &= \underset{n \geq 0}{\text{Max}} \; \{\pi_0(s, c_h, n; w_0, w_1)\}, \\ \pi_1^*(s, c_h; w_0, w_1) &= \underset{n \geq 0}{\text{Max}} \; \{\pi_1(s, c_h, n; w_0, w_1)\} \end{aligned} \tag{1.1}$$

  and given by

$$\begin{aligned} n_0(s, c_h; w_0, w_1) &= \left(\frac{s\theta}{w_0}\right)^{\frac{1}{1-\theta}}, \\ n_1(s, c_h; w_0, w_1) &= \left(\frac{s\theta}{w_1 + H^G}\right)^{\frac{1}{1-\theta}}; \end{aligned}$$

- HI provision decision is derived from

$$\pi^*(s, c_h; w_0, w_1) = \text{Max} \; \{\pi_0^*(s, c_h; w_0, w_1), \pi_1^*(s, c_h; w_0, w_1)\}$$

  and denoted by $\chi(s, c_h; w_0, w_1)$: $\chi(s, c_h; w_0, w_1) = 1$ when HI is sponsored, $\chi(s, c_h; w_0, w_1) = 0$ otherwise.

Incumbent who has already decided to stay in the current period and observed shock realizations confronts the following dynamic problem:

$$V(s, c_h; w_0, w_1) = \pi^*(s, c_h; w_0, w_1) + \beta\gamma \, \text{Max}\{0, E_{s'|s}[E_{c_h'|c_h} V(s', c_h'; w_0, w_1)]\} \tag{1.2}$$

where β is the discount factor, $E_{s'|s}(\cdot)$ and $E_{c_h'|c_h}(\cdot)$ denote conditional expectation operators of argument functions over the range of s' for given s and c_h' for given c_h. Expression $E_{s'|s}[E_{c_h'|c_h}V(s',c_h';w_0,w_1)]$ which is compared to the outside option normalized to zero in the second term of (1.2) is a future discounted value of the establishment in the state (s,c_h) if it decides to stay. Formally,

$$E_{s'|s}[E_{c_h'|c_h}V(s',c_h';w_0,w_1)] = \int_{\underline{s}}^{\bar{s}}\int_{\underline{c}_h}^{\bar{c}_h} V(s',c_h';w_0,w_1)dF(s'|s)dP(c_h'|c_h).$$

Optimal exiting decision of the establishment is denoted

$$x(s,c_h;w_0,w_1) = \begin{cases} 1, & \text{if} \quad E_{s'|s}[E_{c_h'|c_h}V(s',c_h';w_0,w_1)] \geq 0, \\ 0, & \text{else} \end{cases}$$

where $x(s,c_h;w_0,w_1) = 1$ corresponds to a decision to stay, and $x(s,c_h;w_0,w_1) = 0$ - to exit.

Potential entrant decides about entry before the state is known. Thus, the entrant's value net of entry costs is calculated as

$$V_e(w_0,w_1) = \int_{\underline{s}}^{\bar{s}}\int_{\underline{c}_h}^{\bar{c}_h} V(s,c_h;w_0,w_1)dG(s)dQ(c_h) - c_e. \qquad (1.3)$$

1.2.2 Consumers

The economy is populated by a continuum of identical infinitely lived consumers of measure 1 with preferences defined over the streams of the consumption good. On their own, consumers can access only individual HI marketplace offering coverage at $H^I \geq H^G$. Analogous to the establishment side, there is no market for individual HI and consumers shop for it in the marketplace at fixed price H^I measured in the units of the numeraire good. Premiums collected in the group and individual marketplaces are distributed back to the consumers as a part of the lump sum transfer from the government. Recall that HI can be purchased either privately or provided by the employer at no additional cost to the consumer. I assume that, independent of the source of provision, HI provides the *same* flow of health services *mandatory* for all individuals. Then, if consumers are identical, there is no loss of generality in assuming that consumers do not incur any utility from HI or health services. Say shortly, the assumption is that HI is not directly valued by the consumers and is required to have by everyone in this economy [7] .

In every period, all consumers are endowed with one unit of productive time, and supply of labor services is restricted to be either zero or one. Clearly, all individuals supply one unit of labor every period because they do not value leisure. Not only the labor market is competitive, the consumers can costlessly reallocate between the employers with different provision statuses and compensation structures. Then, in a given period the consumer chooses consumption, c_t, and the type of the employer to work for, $\psi_t \in \{0, 1\}$, where $\psi_t = 1$ corresponds to the employer offering HI and $\psi_t = 0$ otherwise. Consumer's problem then is:

$$\mathbf{U} = \underset{\{c_t, \psi_t\}_{t=0}^{\infty}}{\text{Max}} \sum_{t=0}^{\infty} \beta^t u(c_t)$$

subject to the budget constraint in all t

$$c_t = \begin{cases} w_0(1-\tau) - H^I + \Pi_t + T_t, & \text{if } \psi_t = 0, \\ w_1(1-\tau) - \tau^G H^G + \Pi_t + T_t, & \text{if } \psi_t = 1 \end{cases}$$

$$c_t \geq 0, \psi_t \in \{0, 1\}$$

[7] This assumption can be relaxed without affecting the findings of this paper. For example, one can assume that the consumers derive utility from having HI with consumer's tastes being distributed in accordance with some distribution. Then, a fraction of the population will choose not to have HI which will be characterized by a threshold of a taste parameter. For the remaining population that demands HI, the compensation packages of two types of establishments (sponsoring and not sponsoring HI) will need to satisfy the equilibrium condition (1.5) below to sustain both types of establishments in equilibrium. The rest of the equilibrium, including the subsidy to the establishments sponsoring HI (1.7), will remain the same. Given that the tax exemption would generate the same subsidy to the establishments sponsoring HI and, thus, the same distortion on the the production side, I keep the consumer side very simple.

where τ is the labor income tax levied on wages, τ^G is a tax levied on the premiums paid by employers (tax exemption of ESHI is the case when $\tau^G = 0$), Π_t is the consumer's share in the profit of all establishments, T_t it the lump sum transfer from the government to ensure a balanced budget constraint in every period. I assume that all consumers own equal share in the aggregate profits and face equal transfers from the government. Thus, Π_t and T_t also denote aggregate amounts.

Consumer's problem stated in this form reduces to maximizing per period utility by choosing employment at the establishment which ensures the highest disposable income remaining after HI purchase:

$$\tilde{\mathbf{U}} = \underset{\psi \in \{0,1\}}{\text{Max}} \ (1-\psi)[w_0(1-\tau) - H^I] + \psi[w_1(1-\tau) - \tau^G H^G]. \tag{1.4}$$

1.2.3 Equilibrium

In this section, I define equilibrium in which both establishments sponsoring and not sponsoring HI coexist.

Tax Exemption and Implicit Subsidy to Employers

Because consumers can costlessly move between the employers, equilibrium with two kinds of establishments stipulates that compensation packages at establishments sponsoring and not sponsoring HI provide the consumers with the same utility level. Equating disposable incomes of consumers employed at the establishments of different provision statuses in (1.4) delivers an equilibrium relationship between w_0 and w_1:

$$w_1 = w_0 - \frac{1}{1-\tau}\left(H^I - \tau^G H^G\right), \tag{1.5}$$

Using (1.5), the cost of labor for the establishments sponsoring HI is

$$w_1 + H^G = w_0 - \frac{1}{1-\tau}\left(H^I - H^G(1-\tau+\tau^G)\right). \tag{1.6}$$

Note that for any $\tau^G < \tau$ the second term in (1.6) is positive if $H^G \leq H^I$. This implies that employers sponsoring HI access implicit labor cost subsidy generated jointly by the differences in the premiums in individual and group HI marketplaces and tax exemption of the ESHI premiums. When $H^I = H^G = H$, difference in the tax treatment of the premiums is the sole source of the subsidy. In the latter case, if ESHI premiums are tax exempt ($\tau^G = 0$), the employers' per employee costs are $\frac{\tau}{1-\tau}H$ lower under the employer-sponsored provision than it would be otherwise. For exposition purposes, I denote the size of the subsidy h (note that h is independent of any equilibrium objects):

$$h = \frac{1}{1-\tau}\left(H^I - H^G(1-\tau+\tau^G)\right). \tag{1.7}$$

Intuitively, the subsidization effect is attained because in the case of the individual provision, consumers pay HI premium from their after-tax income. With employer-sponsored provision, premium is paid by the employer from the employee's before-tax income. In the latter case, individuals receive full value of benefits embedded in the HI and do not incur any tax on this part of their income. But then, the compensation package under employer provision costs less than wages necessary to provide consumers with the level of utility when HI coverage is absent.

In the remainder of the paper, I rely on the relationship in (1.6) which states that establishments sponsoring HI access implicit per employee labor cost subsidy h. Given (1.5) and (1.6), the set of prices to be determined in equilibrium reduces to one wage rate, w_0. Herewith, w_1 is determined as in (1.5) which is equivalent to $w_1 = w_0 - H^G - h$ [8] , and marginal cost of labor for the establishments sponsoring HI is given by (1.6) which is equivalent to $w_1 + H^G = w_0 - h$. In the remainder of

[8] If tax exemption of the ESHI premiums is eliminated, then $w_1 = w_0 - H^G$ holds, i.e., workers get a reduction in the wage by exactly the cost of HI born by the employer. Otherwise, employers extract extra value from providing HI lowering wages by more than it costs to purchase HI as in (1.6).

the paper, I utilize (1.5) and (1.6) to replace w_1 and $w_1 + H^G$ in the model, drop the subscript on w_0 and switch to condition equilibrium objects on a single wage w denoting w_0.

Definition of Equilibrium

Before I define equilibrium, there are several useful generalizations of the decisions to exit and sponsor HI that simplify characterization of equilibrium objects and aggregate variables.

First, for each s the decision to sponsor HI is characterized by a threshold $c_h^*(s)$ such that

$$\chi(s, c_h; w) = \begin{cases} 1, & \text{if} \quad c_h \leq c_h^*(s), \\ 0, & \text{else.} \end{cases}$$

The proof is obvious. Because for any positive subsidy $\pi_1^*(s, 0; w) > \pi_0^*(s, 0; w)$ $\forall s$, all establishments decide in favor of HI at $c_h = 0$. At the same time, $\pi_1^*(s, c_h; w)$ is strictly decreasing in c_h for a fixed s implying that there exit $c_h^*(s)$ such that $\pi_0^*(s, c_h^*(s); w) = \pi_1^*(s, c_h^*(s); w)$, and $\pi_1^*(s, c_h; w) > \pi_0^*(s, 0; w) = \pi_0^*(s, c_h; w)$ for all $c_h \leq c_h^*(s)$, i.e., establishments sponsor HI for all $c_h \leq c_h^*(s)$. Reverse for all $c_h > c_h^*(s)$.

Second, endogenous exit is summarized by the set of thresholds of productivity $s^*(c_h)$ $\forall c_h$ increasing in c_h. In words, the first time the establishment's state is (s, c_h), where $s < s^*(c_h)$, establishment optimally exits the market.

Now, I characterize some equilibrium objects, define equilibrium and formalize several aggregate variables for future discussion. Transition functions F and P, distribution functions G and Q, exogenous probability of survival γ, endogenous exit decisions of the establishments and mass of entry induce a distribution of establishments over the individual states in every period. Denote $\mu(S, C_h)$ to be the density function of the distribution of establishments over individual states at the end of the period $t-1$ (S and C_h are from the Borel sets of the support of distribution of s and c_h, correspondingly). Let M be the mass of establishments entering at the beginning of period t. Then, the aggregate state of the economy at the end of period t after exit of incumbents has occurred, new establishments have entered and all information has been revealed is given by:

$$\mu'([\underline{s}, s'), [\underline{c}_h, c'_h)) = \gamma \int_{\tilde{s}\in[\underline{s},s')} \int_{\tilde{c}_h\in[\underline{c}_h,c'_h)} \int_{c_h\in[\underline{c}_h,\bar{c}_h]} \int_{s\geq s^*(c_h)} dF(\tilde{s}|s)dP(\tilde{c}_h|c_h)d\mu(s,c_h) + \\ +M \int_{\tilde{s}\in[\underline{s},s')} \int_{\tilde{c}_h\in[\underline{c}_h,c'_h)} dQ(\tilde{c}_h)dG(\tilde{s}). \tag{1.8}$$

First term in (1.8) accounts for the transition of incumbents across the states acknowledging exogenous and endogenous exit. Second term takes into consideration the distribution of state variables across entrants.

At the end of period t, the mass of establishments sponsoring HI in S, $\mu'_1(S)$, is

$$\mu'_1([\underline{s}, s')) = \int_{\tilde{s}\in[\underline{s},s')} \int_{c_h\leq c^*_h(\tilde{s})} d\mu'(\tilde{s}, c_h).$$

Definition of Equilibrium. Given (1.5) and (1.6), a stationary recursive competitive equilibrium with two types of establishments and entry consists of the wage rate w^* denoting equilibrium value of w_0, employment decisions of the consumers $\psi^*_i(w^*)$ $\forall i \in [0, 1]$ where i denotes consumer's index, value functions for the incumbent establishment $V(s, c_h; w^*)$ and the entrant $V_e(w^*)$, establishment's optimal employment decisions $n(s, c_h; w^*)$, exiting decisions $x(s, c_h; w^*)$ and HI provision decisions $\chi(s, c_h; w^*)$, the sets of thresholds $c^*_h(s)$ $\forall s$ and $s^*(c_h)$ $\forall c_h$, mass of entry $M^* > 0$ and the distribution of establishments μ such that:

1. Given w^*, employment decisions $\psi^*_i(w^*)$ solve consumer's optimization problem (1.4) $\forall i \in [0, 1]$;

2. Given w^*, value function $V(s, c_h; w^*)$ is a solution to the establishment's problem (1.2); $n(s, c_h; w^*)$, $x(s, c_h; w^*)$ and $\chi(s, c_h; w^*)$ are the corresponding optimal decision rules;

3. Free entry condition is satisfied

$$V_e(w^*) = 0$$

where $V_e(w)$ is defined as in (1.3);

4. Labor market clears:

$$N_0^d(w^*) = N_0^s(w^*), \quad N_1^d(w^*) = N_1^s(w^*)$$

where $N_0^s(w)/N_1^s(w)$ and $N_0^d(w)/N_1^d(w)$ denote aggregate labor supply and demand (upper script is to denote supply or demand, lower script - the type of the establishment) calculated as:

$$N_0^d(w) = \int_{\underline{s}}^{\bar{s}} \int_{c_h^*(s)}^{\bar{c}_h} n(s, c_h; w) d\mu(s, c_h), \quad N_1^d(w) = \int_{\underline{s}}^{\bar{s}} \int_{\underline{c}_h}^{c_h^*(s)} n(s, c_h; w) d\mu(s, c_h),$$

$$N_0^s(w) = \int (1 - \psi_i^*(w^*)) di, \quad N_1^s(w) = \int \psi_i^*(w^*) di.$$

Values of $c_h^*(s)$ $\forall s$ are consistent with optimal provision decisions $\chi(s, c_h; w^*)$;

5. Stationary distribution of establishments over individual states, μ, evolves in accordance with (1.8) given the constant mass of entry M^* and threshold productivities $s^*(c_h)$ $\forall c_h$. Values of $s^*(c_h)$ are consistent with optimal exiting decisions of establishments $x(s, c_h; w^*)$;

6. The government budget constraint is satisfied:

$$T = \left(\tau w^* + H^I\right) N_0^s(w^*) + \left(\tau w_1(w^*) + \tau^{HI} H^G + H^G\right) N_1^s(w^*).$$

where $w_1(w^*)$ denotes equilibrium level of w_1 determined as $w_1(w^*) = w^* - H^G - h$.

Standard algorithm can be used to compute the equilibrium in this economy. Because wage is the only equilibrium object that affects the establishment's value, the equilibrium wage w^* is pinned down by the free entry condition. Given w^*, equilibrium distribution of establishments can be found for any arbitrary mass of entry. Let $\hat{\mu}(s, c_h)$ be the stationary distribution of establishments generated by a unit mass of entry. As in Hopenhayn and Rogerson (1993), $\mu(s, c_h) = M\hat{\mu}(s, c_h)$ holds here. Therefore, given optimal employment decisions $n(s, c_h; w^*)$ and $\hat{\mu}(s, c_h)$, equilibrium mass of entry M^* is pinned down by the labor market clearing condition. Note that, in the equilibrium, the consumers are indifferent between the employers. Thus, the alternative formulation of the labor market clearing condition would equate overall labor demand $N^d(w^*) = N_0^d(w^*) + N_1^d(w^*)$ and supply $N^s(w^*) = N_0^s(w^*) + N_1^s(w^*) = 1$:

$$\int_{\underline{s}}^{\bar{s}} \int_{\underline{c}_h}^{\bar{c}_h} n(s, c_h; w^*) d\mu(s, c_h) = 1.$$

Several aggregate variables of interest can be found as:

- gross aggregate output Y (does not account for any fixed costs):

$$Y = \int_{\underline{s}}^{\bar{s}} \int_{\underline{c}_h}^{\bar{c}_h} f(n(s, c_h; w), s) d\mu(s, c_h),$$

- aggregate output $\tilde{Y}$ (accounts for all kinds of fixed cost: c_f, c_h and c_e):

$$\tilde{Y} = Y - \int_{\underline{s}}^{\bar{s}} \int_{\underline{c}_h}^{\bar{c}_h} c_f d\mu(s, c_h) - \int_{\underline{s}}^{\bar{s}} \int_{\underline{c}_h}^{c_h^*(s)} c_h d\mu(s, c_h) - Mc_e,$$

- total subsidy to the establishments sponsoring HI S:

$$S = h \int_{\underline{s}}^{\bar{s}} \int_{\underline{c}_h}^{c_h^*(s)} n(s, c_h; w) d\mu(s, c_h),$$

- one period aggregate profit of all operating establishments Π:

$$\Pi = \int_{\underline{s}}^{\bar{s}} \int_{\underline{c}_h}^{\bar{c}_h} \pi^*(s, c_h; w) d\mu(s, c_h) - M^* c_e = \tilde{Y} - wN^d(w) + S,$$

- lump-sum transfer to consumers T:

$$T = \tau w N^s(w) + H^I N^s(w) - S,$$

- aggregate consumption C:

$$C = w(1-\tau)N^s(w) - H^I N^s(w) + \Pi + T.$$

Since in the equilibrium $N^d(w) = N^s(w) = 1$, it also follows that $C = \tilde{Y}$.

1.3 Quantitative Analysis

In this section, I quantify the distortionary effects of the subsidy generated by the tax exemption of ESHI premiums.

The exposition is split into several parts. First, I discuss the calibration procedure for the benchmark of the quantitative model and compare the calibrated benchmark to the data. Second, I perform numerical experiments of setting the subsidy to zero to evaluate output gains from eliminating the tax exemption under a fixed and endogenously determined distribution of establishments. Last, I evaluate and discuss the response of the economy to the regulation resembling most important features of the employer mandate of the Affordable Care Act in the presence of the tax exemption of ESHI premiums, i.e., in the presence of the subsidy.

1.3.1 Calibration

Calibration procedure is carried out so that the model mimics some features of the U.S. data on the establishment dynamics and ESHI. A large part of the calibration procedure follows the strategy in Hopenhayn and Rogerson (1993) but differs in postulating that the establishments sponsoring HI are granted access to the labor cost subsidy h.

To calibrate the model, I have to choose 16 parameters, as summarized in Table 2.1 below. The length of the model period corresponds to 1 year in the data so as standard in the literature. I assign the discount factor β is set to 0.96 which corresponds to the annual interest rate of 4%. The parameter of the decreasing return to scale θ in the establishment level production function is assigned 0.85 which has been used in the related papers, for example, Restuccia and Rogerson (2008) and Brugemann and Manovskii (2009). Also standard, I normalize $w = w_0 = 1$ in the benchmark and calibrate the cost of entry c_e so that the free entry condition holds under this normalization.

Next, I assign a value to the labor cost subsidy h in (1.7) imputing it from the data without solving the model. To approximate its' magnitude, I assume $H^I = H^G = H$ and $\tau^G = 0$, and measure the subsidy $h = \frac{\tau}{1-\tau}H$. Such estimate essentially provides a lower bound for h which I use for the benchmark [9] . In the next section, the numerical experiment of removing the subsidy is repeated for different values of h while the model is recalibrated for every given level of the subsidy.

To construct h from the data, I need to know employer's payments towards HI deducted from the determination of individual's taxable income, H, and tax rates at which these payments are exempt, τ. Since I adopt normalization $w = w_0 = 1$, nor-

[9]More stringent empirical estimate of h would distinguish between H^I and H^G and take into account the generosity of plans purchased individually and by the employers.

malizing the measure of H by an average per employee payroll of establishments not sponsoring HI in the data will harmonize H with $w = 1$ in the model. I assign H a value of an average per employee HI premium payments of establishments sponsoring HI equal to \$3,130.8 or 14.93% of the average per employee payroll of establishments not sponsoring HI in 1997 Robert Wood Johnson Foundation Employer Health Insurance Survey, Long and Marquis (1997) [10] (in the reminder of the paper, I refer to this data source as the Survey). To determine the subsidy, I apply the effective marginal tax rate $\tau = 0.4$ consistent with Prescott (2004). The labor cost subsidy h then is calculated to be \$2,087.2 or 0.0995 after normalization.

The remaining parameters are calibrated jointly by solving numerically for the equilibrium of the model for different sets of parameters and choosing one set at which the equilibrium statistics are closer to the targets. Table 2.1 lists the targets next to the parameters adjusting which the match is achieved. Table 2.2 provides the values of the targets in the data and corresponding benchmark statistics from the calibration exercise to illustrate the match of the model to the data. The proceeding discussion in this section elaborates on the choice of parameters and details of constructing the targets from the data.

Transition functions $F(s'|s)$ and $P(c_h'|c_h)$ are calibrated taking approximations of the AR(1) processes of $\ln(s_t)$ and $\ln(c_{h_t})$ on the discrete grids for s and c_h with 100 grid points each:

$$\ln(s_t) = a_s + \rho_s \ln(s_{t-1}) + \epsilon_t,$$

$$\ln(c_{h_t}) = a_{c_h} + \rho_{c_h} \ln(c_{h_{t-1}}) + \xi_t,$$

[10]The Survey contains information on the characteristics of employers and their offers of HI coverage (the plan types, premiums, benefits, cost-sharing, etc.). The data work carried out using the Survey is based on the sample of 21,545 private sector establishments weighted by the sampling weights provided in the Survey.

where

$$\epsilon_t \sim N(0, \sigma_s^2),\ a_s \geq 0,\ 0 \leq \rho_s < 1,$$

$$\xi_t \sim N(0, \sigma_{c_h}^2),\ 0 \leq \rho_{c_h} < 1.$$

The functions F and P are computed using Tauchen's discretization method (Tauchen (1986)).

Picking parameters for these processes requires some deliberation. First, observe that the establishment's dynamics in the model is driven by stochastic processes of both productivity s and fixed cost of setting up HI c_h. The latter determines HI provision decisions of establishments and, therefore, their access to the subsidy which affects the employment. Because changes in the establishment's employment over time are driven by changes in both s and c_h, a straightforward link between parameters of the $\ln(s_t)$ process and the law of motion for $\ln(n_t)$ cannot be established to calibrate ρ_s and σ_s as in Hopenhayn and Rogerson (1993). Thus, I take the following strategy.

Even though the establishments granted access to the subsidy adjust their employment, productivity remains a predominant driving force of changes in the establishment's employment over time. Hence, I pick ρ_s so that the estimated regression coefficient $\hat{\rho}$ of the AR(1) process of log employment

$$\ln(n_t) = a + \rho \ln(n_{t-1}) + \eta_t, \tag{1.9}$$

in the equilibrium of the model matches comparable estimate of the same process in the data. Hopenhayn and Rogerson (1993) and Lee and Mukoyama (2008) report estimates of the AR(1) process of plant-level log employment derived from the samples of manufacturing establishments. For calibration, I obtain alternative estimates of (1.9) utilizing information on establishments irrespective of their industry. Along with the question about the total number of active employees which identifies n_t, the Survey asks two questions allowing to approximate the number of employees in the previous year, n_{t-1}. It includes questions about a number of permanent employees added, Δn_t^+, and removed, Δn_t^-, from the payroll in the last 12 months. Using this information, I construct a variable $n_{t-1} = n_t - \Delta n_t^+ + \Delta n_t^-$ to find the number of employees in $t-1$ and estimate regression (1.9) restricting dataset to the establishments aged one year and older. Estimation procedure yields the regression coefficient of the process $\hat{\rho} = 0.9574$ [11] .

When choosing σ_s, I target to match the fraction of establishments in the smallest size category from 1 to 5 employees. Recall that changing over time c_h and access to the subsidy push establishments to decide differently about employment even in the absence of changes in productivity. This variation contributes to the variance in the establishment's growth rates over time and lowers σ_s necessary to match it in the data. The life-time growth of establishments slowers down if σ_s is low leading to the size distribution with a large share of small establishments. Meanwhile, model's predictions regarding the distortion might be significantly affected by the ability of the model to replicate the distributions of establishments and HI. Thus, I target a statistics related to the establishment's size distribution rather than often used variance of the growth rates of n.

The value of a_s, the constant of the process of log productivity, determines the mean of the process and is set to match the average employment level in the data.

Second, my strategy to pick the parameter values for the process of $\ln(c_{h_t})$ is to target the patterns of HI provision of establishments in the U.S. I choose a persistence,

[11] In the Survey, the number of active employees, n_t, refers to all types of employees including temporary and seasonal workers. At the same time, questions about Δn_t^+ and Δn_t^- are related to the permanent employees only. It is possible to estimate the regression (1.9) using the number of permanent employees because establishments report fractions of temporary and seasonal workers which can be used to eliminate them from n_t. The estimate of ρ from such regression are very similar to the one reported above.

ρ_{c_h}, so that the slope coefficients from the same regression of current period the establishment's HI provision dummy on the dummy for HI two years earlier coincide in the equilibrium of the model and the data:

$$d_t = a_d + \rho_d d_{t-2} + \zeta_t$$

where d_t and d_{t-2} are the dummy variables equal 1 if HI is sponsored and 0 otherwise in periods t and $t-2$. Such seemingly unusual choice is dictated by the availability of the information on the dynamics of HI provision of establishments in the Survey used to estimate this regression. As with the employment, the Survey asks two questions about the establishment's dynamics of HI provision status over time. Establishments not sponsoring HI at the time of the Survey are asked whether they offered HI within two years prior to the Survey date. In contrast, establishments sponsoring HI are asked about number of years HI has been offered. The clear drawback of the formulation of the first question is that it does not refer to a particular point in time. Similarly, the second question lacks a clear identification whether HI has been continuously offered during the period indicated by the establishment. Hence, any statistics on the dynamics of HI provision derived from these questions may not take into account all variation over time. For calibration purposes, I treat these questions as if the establishments were asked either about provision exactly two years ago or about the time they continuously offered HI. Then, limiting the dataset to the establishments of 2 years and older, I construct the dummy variables d_t and d_{t-2} and estimate the slope of the regression $\hat{\rho}_d = 0.8086$.

The remaining parameters of the $\ln(c_{h_t})$ process, a_{c_h} and σ_{c_h}, are tied to the distribution of HI provision across establishments. Specifically, I choose them to match the fraction of establishments providing HI in the first size category of establishments from 1 to 5 employees and overall fraction of establishments sponsoring HI.

For the distribution of the shocks at entry, I assume that both s and c_h have log normal distributions: $s \sim \ln \mathcal{N}(\mu_s^{ent}, \sigma_s^{ent})$ and $c_h \sim \ln \mathcal{N}(\mu_{c_h}^{ent}, \sigma_{c_h}^{ent})$. Parameters μ_s^{ent} and $\mu_{c_h}^{ent}$ are chosen to correspondingly match the share of entrants in the first size category from 1 to 5 employees and the fraction of those providing HI. The values of σ_s^{ent} and $\sigma_{c_h}^{ent}$ are set to match the model predictions about the average size of entrants and the overall rate of HI provision among them with the data. The data for calculating the targets for entrants were limited to the establishments aged 1 year and younger (I refer to these establishments as young in the remainder of the calibration section).

Values of c_f and γ are chosen so that entrant's and overall exit rates in the model match correspondingly the 1 year exit rate of young establishments and overall exit rate observed in the data.

As mentioned, parameters and their calibrated values, and targets in the data are summarized in Table 2.1. Table 2.2 displays the calibration targets and model statistics next to them. Tables 1.3 and 1.4 report distributional statistics of the U.S. economy and corresponding benchmark values.

Parameter	Target	Value
Subsidy, h	Equated to $\frac{H\tau}{1-\tau}$ at $\tau = 0.4$ and	0.0995
	H = av. HI spendings normalized by	
	av. payroll of est. not sponsoring HI [12]	
Discount factor, β	Interest rate	0.96
Decreasing return to scale, θ	Restuccia and Rogerson (2008)	0.85
Constant of the $\ln(s_t)$ process, a_s	Av. establishment size [13]	0.0148
AR(1) coefficient of the $\ln(s_t)$ process, ρ_s	AR(1) coefficient of the $\ln(n_t)$ process	0.9632
SD of the $\ln(s_t)$ process, σ_s	Establishment share, < 5 employees [14]	0.0613
Constant of the $\ln(c_{h_t})$ process, a_{c_h}	Fraction of establishments sponsoring HI, < 5	-0.3865
AR(1) coefficient of the $\ln(c_{h_t})$ process, ρ_{c_h}	Slope coefficient from the regression	0.9532
	of HI provision dummy on the dummy	
	for HI two years earlier	
SD of the $\ln(c_{h_t})$ process, σ_{c_h}	Fraction of est. sponsoring HI [15]	1.1
Mean of G distribution of s for entrants, μ_s^{ent} [16]	Share of young establishments, < 5	0.2547
SD of G distribution of s for entrants, σ_s^{ent}	Av. size of young establishments	0.2296
Mean of Q distribution of c_h for entrants, $\mu_{c_h}^{ent}$	Fraction of young est. sponsoring HI, < 5	4.2469
SD of Q distribution of c_h for entrants, $\sigma_{c_h}^{ent}$	Fraction of young est. sponsoring HI	7.1484
Survival probability, γ	1 year exit rate	0.9265
Fixed cost of operation, c_f	1 year exit rate of young establishments	0.4883
Entry cost, c_e	Entrant's value at $w = 1$	13.088

Table 1.1: Benchmark Parameters

[12] *Source*: value of H and targets for ρ_s, a_{c_h}, ρ_{c_h}, μ_s^{ent}, $\mu_{c_h}^{ent}$ and $\sigma_{c_h}^{ent}$ are obtained from the Survey, Long and Marquis (1997).

[13] *Source*: targets for a_s, σ_s^{ent}, γ and c_f are calculated from Business Employment Dynamics (BED) data for year 2000: `http://www.bls.gov/bdm/us_age_naics_00_table5.txt`; `http://www.bls.gov/bdm/us_age_naics_00_table6.txt`.

[14] *Source*: Choi and Spletzer (2012), private sector establishment data for year 2000 from Table 1.

[15] *Source*: imputed using the size distribution of establishments from Choi and Spletzer (2012) and the distribution of HI from the Survey, Long and Marquis (1997).

[16] Parameter values in this and next three rows are on the logarithmic scale.

Variable	Data	Model
Average establishment size	17.14	17.14
Average size of young establishments	7.94	7.94
1 year exit rate, overall	9.46%	9.59%
1 year exit rate of young establishments	21.6%	21.61%
AR(1) coefficient of $\ln(n_t)$ process	0.9574	0.9581
Slope coefficient from the regression of HI provision	0.8086	0.8032
dummy on the dummy for HI two years earlier		
Establishment share, < 5 employees	0.498	0.4982
Share of young establishments, < 5 employees	0.7026	0.7024
Fraction of establishments sponsoring HI, < 5 employees	0.338	0.3348
Fraction of young establishments sponsoring HI, < 5 employees	0.1515	0.1510
Fraction of establishments sponsoring HI	0.5159	0.5182
Fraction of young establishments sponsoring HI	0.2233	0.2233

Table 1.2: Calibration Targets and Model Values

	Establishment Size					
	< 5	5 to 9	10 to 19	20 to 49	50 to 99	≥ 100
Share of establishments, *data* [17]	*0.498*	0.209	0.137	0.095	0.033	0.027
Share of establishments, *model*	*0.4982*	0.1749	0.1320	0.1242	0.0406	0.0302
Fraction of est. sponsoring HI, *data* [18]	*0.338*	0.567	0.689	0.820	0.916	0.982
Fraction of est. sponsoring HI, *model*	*0.3348*	0.5470	0.7005	0.8004	0.8805	0.9344

Table 1.3: Distributional Statistics, Overall

	Establishment Size				
	< 5	5 to 9	10 to 19	20 to 49	≥ 50
Share of young establishments, *data*[18]	*0.7026*	0.165	0.0864	0.0341	0.0119
Share of young establishments, *model*	*0.7024*	0.1328	0.0793	0.0603	0.0252
Fraction of young est. sponsoring HI, *data*[18]	*0.1515*	0.3332	0.4333	0.4095	0.8741
Fraction of young est. sponsoring HI, *model*	*0.1510*	0.2819	0.4056	0.5071	0.6622

Table 1.4: Distributional Statistics, Entrants

Highlighted in Tables 1.3 and 1.4 are the shares of establishments in the first size category and the fraction of those sponsoring HI. These statistics are directly targeted and matching them does not come as a surprise. Less expected is that stochastic processes for s and c_h in tandem with the calibrated values of other parameters deliver a reasonably good fit of the size distribution of establishments and the distribution of HI which are not targeted. Overall, the model shows a good match to the data.

1.3.2 Results

In this section, I quantify the distortionary effects of the preferential tax treatment of the ESHI premiums. Specifically, I evaluate the output gains from setting the subsidy for the establishments sponsoring HI at $h = 0$. This hypothetical experiment is intended to represent a regulation eliminating the tax exemption of the ESHI premiums, or imposing $\tau^G = \tau$ on H^G in model terms, which wipes out tax benefits for the establishments sponsoring HI [19] . I find that the aggregate output increases by 1.73% (see the second column of Table 1.5).

[17] *Source*: Choi and Spletzer (2012), private sector establishment data for year 2000 from Table 1.

[18] *Source*: calculated using the Survey data, Long and Marquis (1997).

[19] More generally, the regulation imposes the same tax on H^I and H^G.

Tax exemption of the ESHI premiums has two key implications in the economy of this paper. One of them is a heterogeneity in the marginal costs of labor faced by already existing production units arising because only a fraction of establishments sponsoring HI is granted the access to the labor cost subsidy h (intensive margin). This wedge between the labor costs faced by the establishments of different HI provision statuses leads to the misallocation of production factors across operating establishments and lowers the output produced by them. Another implication suggests that distortions can be induced on the entry and exit of establishments (extensive margin). Because the large establishments receiving the subsidy tend to be old, the subsidy effectively redistributes resources from the young to the old establishments. Thus, removing the subsidy is likely to affect the present value of the life-time profits and impact the entry. Additionally, the stay/exit decisions of incumbents might be affected by a combination of the subsidy and general equilibrium effects on wages. By changing the number of establishments in operation, these distortions to the entry and exit can amplify or weaken the output losses associated with the misallocation of labor across already existing establishments.

I center the discussion of the quantitative results around these two implications and decompose the effects of the subsidy removal along the intensive and extensive margins. To understand the importance of the misallocation of labor across existing establishments, I set $h = 0$ and find an off-the-equilibrium allocation assuming that entry and exit decisions are fixed at their distorted levels in the benchmark. In this allocation, the wage rate clears the labor market while the free entry condition is not imposed. This exercise removes heterogeneity in the marginal costs of labor across establishments and efficiently allocates labor across existing establishments which allows me to evaluate the output gains from the reversal of the misallocation among the mass of existing establishments on the intensive margin. Numerically, I find that the output net of all fixed costs increases by 1.11% compared to the benchmark (as reported in the first row of third column in Table 1.5).

To characterize the effects of the subsidy on the entry and exit and evaluate the importance of the extensive margin, I compute the full equilibrium model with $h = 0$. In contrast to the previous analysis, all margins are in action here and, thus, comparing the two economies' responses regarding the aggregate output can illustrate the importance of endogenizing the distribution of establishments. As mentioned, the results for the full model indicate that the output accounting for all fixed costs to be paid by the establishments increases by 1.73% upon the removal of the subsidy (the second column in Table 1.5). The effect roots in the increased number of operating establishments achieved through the changes in the exit and entry. Here, the increase in number of entrants arises because the profits are redistributed from old to young establishments upon the removal of the subsidy which increases the value of entry above its' value in the benchmark and encourages entry. The incentives to exit end up being enhanced on average mainly because removing the subsidy increases the cost of labor for the establishments sponsoring HI which drives up the exit among them. Overall, despite that the elimination of the subsidy is accompanied by the increase in the exit, the increased entry dominates and evolves into a larger mass of establishments. Thus, a large part of the output gain is obtained on the extensive margin due to the changes in the entry and exit (1.73% vs. 1.11% when the distribution of establishments is fixed).

Table 1.5 provides more details on these results:

Variable	Equilibrium effects	Fixed number of establishments
Output, $\tilde{Y}$	**1.0173**	1.0111
$\tilde{Y} - c_h^{min}$	1.0173	1.0111
$\tilde{Y} - c_h^{total}$	1.0123	1.0061
$\tilde{Y} - c_h^{av}$	1.0112	1.0050
Entry, M	1.1121	1
Exit	1.0104	1
Mass of establishments	1.1006	1

Table 1.5: Effects of Eliminating the Subsidy (Relative to Benchmark)

Along with the aggregate output $\tilde{Y}$ accounting for the fixed costs of setting up the insurance as prescribed by the equilibria with fixed and endogenous number of establishments, Table 1.5 also reports three additional measures in lines 2-4 which diverge from the definition of $\tilde{Y}$ in a way they regard c_h. Observe that the output $\tilde{Y}$ demonstrates the extreme response of the economy when, having no other incentives to sponsor HI at $h = 0$, all establishments drop provision and no fixed costs c_h is wasted. In this case, economizing on the fixed costs of HI provision may already be a source of a significant increase in the output. Nevertheless, in a more realistic scenario, fixed costs c_h are likely to be shifted to another payer (at full or partially) who will take upon the role of sponsoring HI for the population covered under ESHI. Statistics reported in rows 2-4 explore some scenarios where the government institution covers the entire population in the government-sponsored universal health care program and bears the costs of setting up HI previously borne by private companies (these statistics are reported relative to the benchmark output $\tilde{Y}$) [20] . In the absence of the direct evidence on the government's efficiency in this kind of programs, I look at three following cases. First, statistics $\tilde{Y} - c_h^{min}$ illustrates the case when the government is very efficient at providing insurance and covers everyone at c_h^{min} corresponding to the minimal establishment level per covered employee fixed cost of HI provision. Essentially, this scenario does not affect the output gains compared to no c_h paid in $\tilde{Y}$ since per employee fixed c_h costs for the establishment facilitating HI provision in the most efficient way are negligible. Second measure, $\tilde{Y} - c_h^{total}$, evaluates the scenario in which the government is less efficient in implementing HI provision for the population and incurs the same cost as private companies did to cover their employees before. In this case, the output gains reduce to milder 1.23% and 0.61%. Third measure, $\tilde{Y} - c_h^{av}$, evaluates the least optimistic scenario when the government is as efficient in getting access to HI as the establishments are in the benchmark model and covers everyone at the benchmark average per covered employee fixed costs, c_h^{av}. Here, the output gains drop to 1.12% and 0.5% but, notably, are still positive. To sum up, these results demonstrate that even if the government cannot achieve the reduction in c_h compared to the private sector, either because c_h is unavoidable in nature or because it covers more individuals than are covered under ESHI, the output still increases when the tax exemption is eliminated.

1.3.3 ACA Employer Mandate

The most recent health care reform, the ACA, institutes many HI programs. Nevertheless, it does not alter the tax treatment of the premiums and, at the same

[20] Formally, c_h^{min} is the minimal establishment level ratio of fixed costs of setting up HI c_h to the number of employees in the benchmark; c_h^{total} is the sum of fixed costs of setting up HI borne by all establishments sponsoring HI in the benchmark; c_h^{av} is the ratio of the total fixed costs of HI provision borne by all establishments sponsoring HI to the total number of employees at these establishments.

time, introduces the employer mandate which might influence the effects produced by the tax exemption. In effect, the mandate levies taxes on large firms not sponsoring HI. By its very nature, such component is a distortionary policy because it affects asymmetrically businesses of different sizes. However, various effects of the ACA tax may mitigate or amplify the output losses caused by the subsidy which makes it interesting to consider the ACA in the scope of this paper. On the one hand, a tax levied on the large employers reallocates profits from the large to the small establishments, in the direction opposite to that induced by the subsidy. Thus, the ACA could encourage entry and, at least partially, offset the losses in the output due to decreased entry. On the other hand, such reallocation happens mostly at the expense of large establishments being a part of large firms not sponsoring HI. Thereby, the ACA tax increases the gap between the marginal costs of labor faced by the large establishments and entails more misallocation. This could undo the improvements in the output through the number of establishments on the extensive margin. The net effect of these two forces is a quantitative question which I explore by accommodating the model of this paper to include a simplified version of the employer mandate.

ACA Description and Model Adjustments

In practice, a component of the ACA known as the employer mandate introduces employer shared responsibility payment for large employers. This is a per year per full-time employee (FTE) fee of $2,000 for employers with 50 or more FTEs if insurance is not offered by the employer and at least one of employees receives a premium tax credit in an Exchange [21] (the first 30 FTE are exempt). Additionally, if at least one FTE receives a premium tax credit because coverage is either unaffordable or does

[21] Exchange refers to the Health Insurance Exchange Marketplace defined as individual insurance market where insurance plans cannot price or deny coverage based on preexisting conditions.

not cover 60% of total health care expenditures for typical population, the employer must pay the lesser of \$3,000 for each of those employees receiving a credit or \$2,000 for each FTE minus 30 [22] .

The model of this paper encompasses two forms of observed heterogeneity of establishments essential for the employer mandate: size and HI provision. However, the detailed specification of the mandate is not implementable in this environment because in real world the penalty is conditioned on the size of business in terms of FTE and premium tax credits received by employees in an Exchange. I take a simpler stand and implement ACA-like regulation with the following assumptions. The ACA tax $\hat{\tau}$ is applied at the *establishment level* if HI is not sponsored and the size of the establishment passes a threshold $n^{tr} = 50$ employees with first $n^{ex} = 30$ workers exempt. Regulation modified in this way imposes a wedge between the costs of labor for large establishments, similar to the ACA, but is implemented at the establishment rather than firm level [23] and waives all conditions for applying the penalty beyond size measured in the number of employees and HI provision.

In the model, such regulation requires to only adjust a static profit maximization problem of an establishment having decided to not sponsor HI given by (1.1). The problem now assumes a choice between staying below the threshold n_{tr} and avoiding tax, and passing n_{tr} and paying the tax. As so, the instantaneous profit of the

[22]More details on the employer shared responsibility payments is available at `http://obamacarefacts.com/obamacare-employer-mandate.php`.

[23]Because some small establishments with < 50 employees may be a part of a large firm taxed by the ACA, restricting the employer mandate to establishments with ≥ 50 employees is an obvious limitation (note that establishments with ≥ 50 employees do not create any tension with the simplified policy; for such establishments the regulation applies independent of whether they operate as a part of a multi-establishment firm or as a single-unit establishment). However, the Survey reveals that an extent of this happening might be small in the data. Specifically, only 16.44% of establishments with < 50 employees count as establishments in multi-establishment firms with ≥ 50 employees. The fraction of those to which the ACA tax might be applied is even smaller since the majority of large firms sponsor HI. Hence, implementation of the ACA at the firm level in the model might not be a channel that substantially changes the predictions of the simplified ACA regulation considered here.

establishment having decided not to sponsor HI can be written as:

$$\pi_0^*(s, c_h; w) = \text{Max } \{\pi_0^{*a}(s, c_h; w), \pi_0^{*b}(s, c_h; w)\},$$

where π_0^{*a}/π_0^{*b} is the profit of the establishment if it operates above/below the threshold n_{tr}. These profits are derived as:

$$\pi_0^{*a}(s, c_h; w) = \underset{n \geq n^{tr}}{\text{Max}} \; sn^\theta - (w + \hat{\tau})n + n^{ex}\hat{\tau} - c_f,$$

and

$$\pi_0^{*b}(s, c_h; w) = \underset{0 \leq n \leq n^{tr}}{\text{Max}} \; sn^\theta - wn - c_f.$$

ACA Effects

With calibrated parameters at hand, I evaluate the effect of the ACA tax on the aggregate output adding it in the distorted benchmark with the subsidy in place. The results imply a moderate decrease in the output by 0.13% (Table 1.6, second column).

In a fashion, similar to the removal of the subsidy, this effect can be decomposed along intensive and extensive margins. First, to explore the effects of the employer mandate on the intensive margin, I hold the distribution of establishments fixed as in the benchmark and adjust the wage rate to clear the labor market. With the introduced ACA tax, the misallocation of resources across existing establishments is aggravated because the tax increases the gap between marginal costs faced by the large establishments sponsoring and not sponsoring HI. The aggregate output drops by 0.14% compared to the benchmark level which is illustrated in the third column of Table 1.6. Second, to infer the effects on the extensive margin, I allow the entry and the exit to endogenously adjust in response to the ACA tax. Once the distribution of establishments is endogenously determined, the entry increases and the exit decreases. Both effects can be attributed to the redistribution of profits from the large to the small establishments caused by the tax. As a result of such redistribution, the present discounted value of life-time profits for young and small establishments increases relative to that for an average incumbent which leads to a higher entry and a lower exit. Together, these effects combine in a larger number of operating establishments compared to the benchmark. Nevertheless, it cannot manifest to the increase in the output and only slightly mitigates output losses caused by the increased misallocation due to the ACA tax. In this case, the output decreases by 0.13% below the benchmark (see the second column of Table 1.6). Table 1.6 illustrates the details regarding these results [24] (all statistics are relative to the benchmark except last four lines):

[24] The benchmark wage for this experiment was renormalized to the average annual wage in 2012 (http://www.bls.gov/oes/2012/may/oes_nat.htm) to accommodate a per employee ACA tax of $2,000.

Variable	Equilibrium effects	Fixed number of establishments
Output, $\tilde{Y}$	**0.9987**	0.9986
$\tilde{Y} - c_h^{total}$	0.9990	0.9989
Entry, M	1.0004	1
Exit	0.9979	1
Mass of establishments	1.0026	1
Share of establishments:		
- unconstrained and not taxed	0.4732	0.4734
- constrained and not taxed	0.0062	0.0062
- taxed	0.0019	0.0019
- sponsoring HI	0.5187	0.5185

Table 1.6: ACA Effects

The magnitude of the reported effects complemented by the fact that the tax of \$2,000 constitutes 4.37% of the equilibrium wage in the full model, it applies to only 0.19% of establishments and constrains 0.62% of them posits a fair question whether a tax with a limited application as here can give a rise to a large effect on the output, or it is a variety of confounding effects when the subsidy is in place that limits its revelation. Though not reported in details, I assess a distortionary power of the ACA tax introducing it in the economy without the subsidy and applying it in the closest proximity to the equilibrium corresponding to the third column of Table 1.6. In particular, I apply a labor cost tax of 4.37% to only those establishments that were taxed in the equilibrium of the full model. For establishments constrained at n^{tr}, I allow the endogenous choice between staying constrained and being taxed. All other establishments are not taxed. In this exercise, the output goes down by 0.03% which demonstrates a limited ability of the ACA tax to generate large negative effects on the output.

1.4 Discussions

1.4.1 Model Assumptions

In this section, I review the assumptions of the model and discuss the alternatives.

Incorporating Capital. When the output is produced with labor and capital inputs, the reallocation of labor resources following introduction of the labor cost subsidy for the establishments sponsoring HI is complemented by the reallocation of capital which can exacerbate the negative effect on the aggregate output.

It turns out that a model with labor and capital where the capital is rented by the establishments rather than owned and there are no capital adjustment costs is equivalent to the model with a single labor input endowed with a higher labor share. To see why, suppose that the production function is given by

$$\tilde{y}_t = \tilde{s}_t^{1-\lambda}(n_t^{\alpha}k_t^{1-\alpha})^{\lambda}$$

where k_t is capital traded at a price r and $\lambda < 1$.

From the static profit maximization problem of the establishment in the absence of any distortions, one can solve for k_t and, using the optimal solution $k_t(n_t)$, derive the establishment's revenues net of the costs of capital:

$$y_t = \tilde{s}_t^{1-\lambda}(n_t^{\alpha}k_t(n_t)^{1-\alpha})^{\lambda} - rk_t(n_t) = s_t n_t^{\theta}$$

where $\theta = \frac{\lambda\alpha}{1-\lambda(1-\alpha)}$ and s_t is a function of $\tilde{s}_t$, λ, α and r. Then, it is plausible to say that the production function $y_t = s_t n_t^{\theta}$ represents the relationship between output and labor after optimally accounting for the capital and predicts the same allocation of labor as the model with the capital. Observe, however, that θ is larger than the labor share $\lambda\alpha$ in the model with the capital.

While two setups, with and without capital, results in the same labor allocations in the absence of distortions, the fact that y_t accounts for the cost of capital might be a reason why these models deliver different effects in response to the same policy. In the case of the *labor cost subsidy*, the proportional change in the establishment level output $\tilde{y}_t$ in the model with two inputs is equal to the change in the output y_t predicted by the model with a single labor input as long as $\theta = \frac{\lambda\alpha}{1-\lambda(1-\alpha)}$. This also holds for the aggregate output of the economy with establishments varying in their access to the labor cost subsidy as in this paper. Therefore, the quantitative impact of the distortion on the gross output I find in the context of the calibrated model is identical to the one predicted by the model with labor and capital at the background in which λ and α can be shaped into $\theta = 0.85$ through $\theta = \frac{\lambda\alpha}{1-\lambda(1-\alpha)}$. One example would be a production function with parameters $\lambda = 0.9$ and $\alpha = 0.64$.

Static vs. Dynamic Model. It is worthwhile emphasizing that the establishment's dynamics in the model of this paper is a source of amplification of the distortionary effects of the subsidy through the entry margin. Generally, a static version of the current model with zero probability of establishment's survival to the next period can easily replicate the cross-sectional distributions of size and HI provision obtained in the benchmark. Note that a static version like that would feature the benchmark mass of operating establishments comprised of new entrants. Hence, it would produce the same effects on the output as the calibrated dynamic model under a fixed distribution of establishments because the elimination of the subsidy would entail the reallocation of resources across the same mass of existing establishments up to the point where the marginal products of labor are equalized. At the same time, such model would miss the establishment's size growth and the increase in probability of HI provision as the establishments are aging. Therefore, it would only account for the fact that the subsidy received by mostly large employers redistributes profits from small to large establishments which still might impact the entry when the distribution is endogenously determined, but it would overlook the fact that the subsidy redistributes profits from young to old establishments. Through this channel in the dynamic model, the profits are postponed into the future which lowers the ex ante value of entry and reduces the incentives of establishments to enter. Consequently, the output gains upon the endogenously determined distribution should be larger in the dynamic version of the model.

To illustrate the quantitative importance of the establishment's dynamics in amplifying the effects of the subsidy, I consider a static version of the model calibrated to exactly match the cross-sectional distributions of size and HI provision from the benchmark. In this version, establishments enter drawing the shocks from the benchmark steady state distribution of individual states $\mu(s, c_h)$, produce for one period, exit and are replaced with new entrants (i.e. the probability of survival is zero) [25] . With that, the allocations and the distribution of establishments in the static model coincide with those in the dynamic model in the presence of the subsidy. Next, I compare the output gains from the elimination of the subsidy in the static and dynamic versions decomposing them into the gains under a fixed and endogenous number of establishments. Table 1.7 reports the results (relative to the benchmark).

[25] Note that while a joint distribution of shocks at entry is readily available from the benchmark steady state distribution of individual states $\mu(s, c_h)$, the entry cost should be recalibrated so that the free entry condition is satisfied in the static model.

Variable	Equilibrium effects		Fixed number of establishments	
	Static model	Dynamic model	Static model	Dynamic model
Gross output, Y	1.0119	1.0213	1.0055	1.0055
Entry, M	1.0429	1.1121	1	1
Exit	1	1.0104	1	1
Mass of establishments	1.0429	1.1006	1	1

Table 1.7: Effects of Eliminating the Subsidy, Static vs. Dynamic Model

Notice that in the static model the distribution of establishments is comprised of new entrants and, thus, the output after accounting for the cost of entry differs from the benchmark. Then, even though the output gains in absolute terms are the same under the fixed distribution in both models, the proportional changes in the output are not. To highlight that the output gains are identical in the columns 4-5, I report gross output (aggregate production) rather than the output accounting for all fixed costs. Columns 2-3 in Table 1.7 demonstrate that dynamics introduces a sizable amplification effects on entry and output under the endogenous distribution of establishments. The mass of entrants increases by 4.29% in the static vs. 11.21% in the dynamic model which boosts the gains in the gross output from 1.19% to 2.13%.

The above discussion suggests that the results of the elimination of the subsidy under the endogenous distribution might be sensitive to how much size and HI provision growth is embedded in the establishment's life cycle in the quantitative model. In this respect, the benchmark economy reproduces well the age profile of the establishments in the data with reference to size and HI provision. Table A.1 in Appendix A reports the average size of establishments and a fraction of those sponsoring HI for different age categories. Notably, the size growth and the increasing probability of HI provision over the establishment's life-time are supported by two model assumptions: persistence of s and c_h shocks. Relaxing these assumptions in a way that lowers the size or HI provision growth over time would result in lower output gains through milder effects on the number of entrants. For example, under the i.i.d. c_h the probability of HI provision as the establishment ages increases only to the extent that a higher productivity s raises profits and allows to pay a bigger fixed cost c_h. In this case, a growth of provision over time would be lower compared to the benchmark where the probability of HI provision increases even in the absence of changes in productivity s.

1.4.2 Output Gains for Different Subsidy Levels

A crucial determinant of the distortion generated by the tax exemption of the ESHI premiums in this paper is the size of the subsidy h. My baseline calibrated value of h is 0.0995 or 9.95% of the wage paid by the establishments not sponsoring HI. However, the empirical subtleties of calculating H^I, H^G, τ and τ^G might affect the size of the subsidy size given in (1.7). Therefore, it is of interest to compute the output gains from eliminating the subsidy in the economies characterized by different values of h in order to account for any biases in the calibrated value. Some examples may be helpful here to illustrate why the size of the subsidy h might differ from the calibrated value.

For instance, the calibrated value of h takes into account only tax exempt employer's contribution towards HI premiums. At the same time, many employees have an opportunity to cover the remaining part of the premium with before-tax income if the employer offers benefits through a section 125 of the Internal Revenue Service

Code (cafeteria plan) which increases the subsidy [26] . Likewise, many employers offer employees the option of contributing to flexible spending accounts which permit employees to make pre-tax contributions to pay for eligible medical expenses [27] . Since this option is available to the employees through their employers, it contributes to the subsidization effect accruing to the employers due to the tax exemption of the employee's compensation. In a similar way, individually purchased HI plans may be more expensive than similar plans purchased by the employers having access to the group insurance market. Thus, the employers may also be entitled to an additional benefit due to the differences in the prices of HI plans captured in (1.7) through distinguishing between H^I and H^G [28] . On the contrary to these examples, one can argue that because the establishments not sponsoring HI tend to be low productivity, they are likely to fill in vacancies with low skilled and low paid workers. Thus, normalizing the value of H by w_0, as in the calibration exercise, might exaggerate the value of the subsidy.

Given these possibilities, I explore the implications of the subsidy varying from 0 to 0.15 or, say differently, from 0 to 15% of the wage paid by the establishments not sponsoring HI. To do this, I recalibrate the model for each subsidy level retaining other targets from the calibration section. Note that the parameter values calibrated simultaneously by solving the model are specific to the value of h. Thus, each alternative h requires solving the model and identifying a new set of parameters at which

[26] This channel might be substantial. For example, in 2012 41% of small firms (3 to 199 workers) and 91% of larger firms participated in the cafeteria plan. At the same time, the contributions of employees accounted for 16% and 26% of the premiums for single and family coverage correspondingly. This implies that for the majority of firms the tax exempt compensation is more than 10% larger than what accounted for in the calibration.

[27] For example, in 2012 17% of small firms and 76% of large firms offered flexible spending accounts to the employees.

[28] Notice that factors like take up ratio of the ESHI by the employees and differences in the cost of single and family coverage should be taken into account. I do not emphasize them here because the value of H in the calibration exercise refers to the average per employee payments towards HI premiums and spreads out total payments towards HI premiums across all employees.

the model matches the targets. After I obtain the parameter values for each h, I redo the experiment of dropping the subsidy under the fixed and endogenously determined distribution of establishments as in Section 1.3.2. Figure 1.1 illustrates the results. For comparison, I also plot the output predicted by a static version of the model with endogenously determined entry at each h.

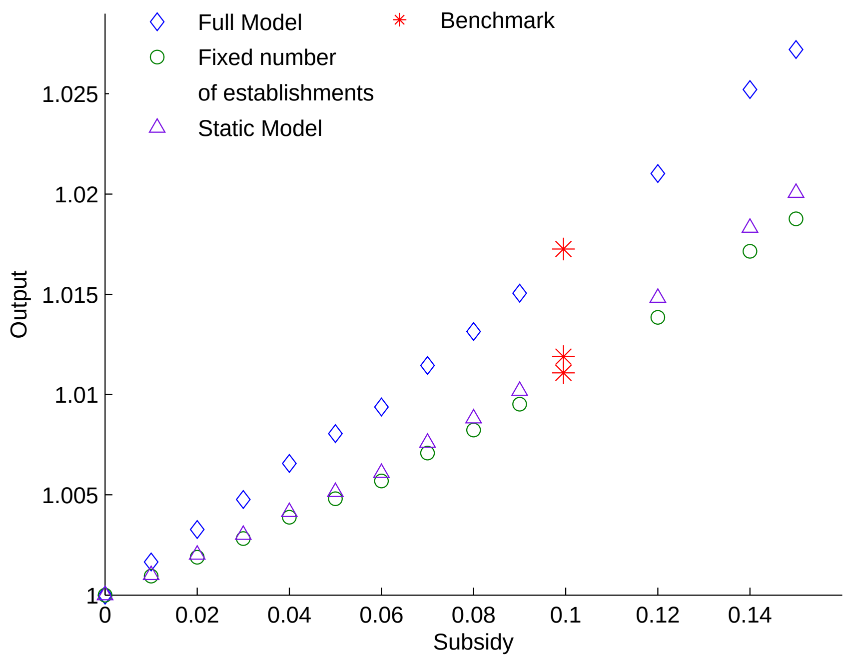

Figure 1.1: Output Gains for Different Subsidy Levels

One would expect that a larger subsidy h implies larger extent of the misallocation across existing establishments and of the profit redistribution from young to old establishments. As a result, the output gains from the removal of the subsidy among the mass of existing establishments and under the endogenous distribution are expected to increase in h. Figure 1.1 illustrates exactly this point. Moreover, the output in the full model increases at a higher rate than that under the fixed distribution illustrating a growing importance of the adjustments on the extensive margin as the subsidy increases. The highest gains in Figure 1.1 correspond to the subsidy

$h = 0.15$ and amount 1.88% and 2.72% under a fixed and endogenous distribution of establishments. Note that the increase in the output in a static model is not nearly as strong as in the full dynamic model and is almost identical to the output in the dynamic model when the distribution of establishments is fixed. In line with a previous discussion, this demonstrates that a static model misses the redistribution of profits from young to old establishments induced by the subsidy. Hence, the importance of the extensive margin in the static environment is very limited which is confirmed by the quantitative results illustrated in Figure 1.1.

1.4.3 Sources of Consumption Gains for the Consumer

The main prediction from elimination of the tax exemption of ESHI premiums in Section 1.3.2 is an increase in the aggregate output implying an increase in the individual and total consumption. I proceed to discuss the sources of consumers' benefits and mechanisms initiating them.

It is worth noticing that even though consumers in the benchmark receive the same value, there is a heterogeneity among them with respect to the structure of the compensation package defined by the HI provision status of the employer. Thus, the sources of benefits for the consumers affiliated with the employers of different HI provision statuses may differ. Rather than examining each consumer's type, I discuss consumption from the average consumer's viewpoint. Averaging elements of the consumer's budget constraint over consumers of different types, the average consumption $\bar{C}$ can be written as [29]

$$\bar{C} = \bar{w}(1 - \tau) - HN_0^s(w) + \Pi + T$$

[29] The expression for the average consumer's budget constraint should not be confused with the expression for the total consumption C given earlier. Even though total consumption is equal to the average consumption, the expression for C takes into account the relationship between w_0 and w_1 in (1.5) and is written in terms of the wage rate w denoting w_0. The expression for $\bar{C}$ does not use (1.5) and retains two types of wage compensation w_0 and w_1.

where $\bar{w} = w_0 N_0^s(w) + w_1 N_1^s(w)$ is an average wage compensation, $N_0^s(w)/N_1^s(w)$ is the labor supplied to all establishments not sponsoring/sponsoring HI. Notice, that the expression for $\bar{C}$ does not distinguish between H^I and H^G since in the calibrated model $H^I = H^G = H$. Also, the expression takes into account that tax collections on the premiums paid by the employers $\tau^G H N_1^s(w) = 0$ since $\tau^G = 0$ in the benchmark and $N_1^s(w) = 0$ in the equilibrium upon subsidy elimination.

Table 1.8 is a supplement to the effects of subsidy elimination in Table 1.5 illustrating some components of the average consumption:

Variable	Equilibrium effects	Fixed number of establishments
Worker's compensation	1.1768	1.1586
Profits	0.9498	1.1027

Table 1.8: Components of Average Consumption (Relative to Benchmark)

These results suggest that the benefits of eliminating the tax exemption of ESHI are encapsulated mainly in the increased compensation. To gain a perspective of the mechanism, observe that the drop of HI provision by the establishments is associated with the shift of HI premiums H from the establishments to the consumers and elimination of the implicit labor cost subsidy h previously transferred from the consumers to the establishments in the form of reduced wages. Being compensated with the equivalent increase in cash wages, the former shift of the premium payments increases consumer's compensation but has no impact on the value of the consumer because wages minus compelled payments of HI premiums stay constant. At the same time, the latter eliminated incidence of the subsidy h effectively makes the consumers better off while increasing the labor costs for the establishments. Regardless of the increased labor costs, the aggregate profits of establishments in the fixed distribution equilibrium in the second column also increase primarily because the allocation of resources improves which increases the output and establishments' revenues [30] . In this allocation, the profits settle at the level where the value of the entrant is larger than the entry costs which makes the entry profitable. Therefore, when the entry is allowed to adjust with endogenously determined distribution of establishments in the last column, the number of entrants increases creating additional demand for labor and pushing the consumer's compensation further up. With this, the profits decrease both because the compensation increases and payments towards the fixed costs of operation and entry costs pick up when the number of entrants increases.

1.5 Concluding Remarks

The analysis of this paper is motivated by the current tax treatment of the ESHI premiums in the United States. It is shown to generate an implicit labor cost subsidy to the employers sponsoring HI when there is a large number of worker competing for the jobs who can costlessly reallocate between the employers. The main argument of this paper is that, because provision among employers is not uniform, such subsidy can reduce the aggregate output through reallocating the labor resources across the establishments and affecting the number of operating establishments via entry and exit. Additionally, it is demonstrated quantitatively that the distortionary effect of the subsidy gets amplified when the related regulation, the ACA employer mandate, is introduced.

[30]Obviously, establishments drop HI provision which economizes on the fixed costs of setting up HI c_h and increases profits. Nevertheless, one could consider an alternative scenario where c_h payments would not matter for the aggregate profits while they would still increase reflecting improvements in the allocation of labor. For instance, one can assume that c_h payment are fully shifted from the establishments to the government who pushes them back onto the establishments in the form of a lump-sum tax. Then, drop in c_h costs paid by the establishments would effectively have no effect on profits, as well as on consumers' welfare.

The essential results of this paper can be summarized as follows. First, the quantitative analysis of the model calibrated to the U.S. data suggests that the elimination of the tax exemption of the ESHI premiums might increase the output by 1.11% if the distribution of establishments does not change. Second, it may significantly increase the number of establishments in operation when one accounts for endogeneity of entry and exit. This may give an additional boost to the output produced upon elimination of the tax exemption. In the model, the output increases by 1.73% when the distribution of establishments is determined endogenously. It is also demonstrated that even if the government, who might take upon the responsibility of providing HI, is less efficient and has to incur a higher cost of setting up HI compared to the private companies, the output gains from eliminating the tax exemption of ESHI premiums are still positive. Finally, the estimates of the output gains from the removal of the subsidy created by the tax exemption in the model calibrated with different levels of the subsidy indicate that the output can go up as much as 2.72% above the benchmark for the subsidy level equivalent to 15% of the average per employee payroll of establishments not sponsoring HI.

Chapter 2

MACROECONOMIC IMPLICATIONS OF HIGHER COST OF HEALTH SERVICES FOR THE UNEMPLOYED

2.1 Introduction

The majority of privately insured workers in the U.S. obtain their health insurance (HI) through the employment. Along with the benefit of reducing exposure to financial risk, HI provides access to lower priced health services because insurance companies contract with medical providers to provide services to their enrollees at discounted prices. At the same time, for those workers who lose their jobs, employer benefits including HI end immediately or shortly after leaving the employer. These individuals suffer a substantial financial loss related to health care as they lose HI and, therefore, have to pay inflated prices for medical services.

This link between the cost of health services and the employment status of the worker has potentially important implications for the functioning of the economy. Because workers lose HI after leaving the employer, higher health care cost for the uninsured increases the cost associated with unemployment which may be an impediment for worker's decision to switch jobs. As a result, the productivity of the economy as a whole and aggregate output may suffer if individuals who would like to move are constrained by the high cost of health services during the unemployment spell. Meanwhile, at the aggregate level, many employed workers might be affected by the cost of unemployment inflated by the cost of health services when HI is absent. In the data, the flows of workers out of jobs are accompanied by a substantial number of separations into unemployment. For example, in 2000-2008, more than 40% of workers separated from their jobs annually, and about 35% of these separations were into unemployment [1] . Because the extent of job reallocations which include unemployment spell is so sizable in the data, aggregate implications of the higher cost of health services faced by the unemployed workers might be significant. The objective of this paper is to quantify some of these implications. More specifically, I ask what are the effects of equalizing the cost of health services faced by the employed and the unemployed on labor market outcomes, aggregate labor productivity and aggregate output.

To address this question, I build a general equilibrium search model in the tradition of Lucas and Prescott (1974) island model where all workers are required to pay extra cost of health services contingent on the worker's employment status. In the model, the employed workers are assumed to face a lower cost of health services as compared with the unemployed. This assumption is motivated by two empirical observations. First observation relates to the evidence on the existing gap in the cost of health services between privately insured and the uninsured. In the paper, I examine aggregate hospital care statistics reported by the Healthcare Cost and Utilization Project (HCUP) and the American Hospital Association (AHA) which indicate that medical bills for the insured are about 50% discounted compared to the uninsured who are subject to a full charge [2] . Second supporting empirical observation points out to a bigger prevalence of lacking HI among the unemployed. For instance, 51% of the unemployed adults aged 18-64 years did not have HI compared with 18.2% of employed in the same age group in 2009-2010 (Driscoll and Bernstein (2012)).

[1] Annual levels for total separations and separations into unemployment are the sum of the 12 monthly levels. Average monthly separation rates for 2000-2008 were calculated using Bureau of Labor Statistics (BLS) data series on job turnover, `https://www.bls.gov/jlt/`. Average monthly employment-unemployment transition rates were calculated based on the BLS data retrieved from `https://www.bls.gov/cps/cps_flows.htm`.

[2] As mentioned before, the primary reason for this difference is the ability of insurance companies to negotiate lower rates for the insured.

In the course of analyzing the effects of the higher cost of health services for the unemployed, I first turn to a simple version of the model which allows for analytical representation of some results and provides important insights regarding the potential effects of the gap between the cost of health services in the employment and unemployment states in the full model. In this version of the model, the productivity of the island fluctuates between two values (high and low). Naturally, the mobility of workers is directed from the islands hit by the low productivity shock to the islands experiencing high productivity. The model demonstrates that increase in the cost of health services for the unemployed operates through two channels. First, it reduces mobility of workers leading to lower unemployment. Second, it worsens the allocation of workers across islands because higher cost of health services during unemployment induces the workers to stay in their locations after the island is hit by the low productivity shock. These two channels create an obvious trade-off: while the first channel tends to increase aggregate production, the second one reduces the level of output. Therefore, the interaction between these channels may cause aggregate output go either way. Quantitative exploration of the simple model shows that the second effect dominates and overall aggregate output decreases in the cost of health services faced by the unemployed as long as the time discount factor is sufficiently high.

While the same trade-off operates in a setup with a richer process for the productivity shocks, the quantitative importance of two channels might be affected in the presence of discounting. To address this issue, I move to a full model featuring multiple levels of the productivity shock and quantitatively investigate the magnitude of the effects discussed in the simple model. The model preserves the structure of Lucas and Prescott (1974) island model with varying costs of health services for the workers contingent on their employment status and, additionally, introduces unemployment insurance payments available to the unemployed. Because the model cannot be solved analytically, I undertake a calibration exercise to mimic selected observations on the U.S. economy. The important part of the calibration procedure is that I use the differences in the cost of health services for the privately insured and the uninsured inferred from the aggregate hospital statistics as the gap between the cost of health services faced by the employed and the unemployed in the benchmark. After the model is calibrated, I seek to understand the importance of equalizing the costs of health services across worker in the numerical experiment which reduces the cost of health services for the unemployed to the level faced by the employed. Quantitative results show that, similar to the simple model, change in the output is determined by a composition of changes in unemployment and allocation of workers across islands. Since lower cost of health services for the unemployed encourages worker to search more, the economy with equalized costs of health services is characterized by unemployment about 1.5 percentage points higher as compared with the benchmark. At the same time, due to the increased willingness of workers to search, the allocation of labor across islands or the composition of undertaken jobs improves as more workers are prone to leave their islands after experiencing low productivity shock. This leads to an increase in aggregate labor productivity by about 1.2%. Although the second effect increases aggregate output on its' own, the first one, increased unemployment, is quantitatively of a first order importance which leads to a decrease in aggregate output by 0.26%.

The structure of the remainder of the paper is as follows. Section 2.2 provides evidence on the gap in the cost of health services faced by the insured and the uninsured. Section 2.3 presents a simple model which illustrates the main trade-off between unemployment and allocative efficiency when employed and unemployed are subject to different costs of health services. Section 2.4 presents full equilibrium model. Calibration strategy and results of the numerical experiment equalizing costs of health services for the employed and the unemployed are discussed in Section 2.5. Section 2.6 concludes.

2.2 Evidence on the Gap Between the Cost of Health Services for the Privately Insured and the Uninsured

In this section, I examine the evidence on the gap in the rates paid for the medical services by different payers. The attention is concentrated on the gap faced by the uninsured and holders of private insurance to be used later in the calibration section of the paper.

It is a common practice that providers of medical services have a detailed and very comprehensive list of charges for medical services. These charges vary across providers and are used to determine a price of individual's medical condition (or a treatment) by adjusting for the intensity and a combination of received services. Then, the final charge of the treatment is billed to a payer (either to a patient or HI provider). Uninsured and other self-pay recipients [3] of medical services are presented with and expected to pay medical bills reflecting a full charge. Amounts paid by insurance companies of behalf of the patients are usually discounted and represent a result of the negotiation between the insurer and the medical services provider. In general, insurance companies may negotiate prices with every provider on every treatment, procedure or medical service on every plan. Aside from the fact that negotiated prices are not readily available to the public [4] , multiplicity of the negotiated prices creates a challenge for measuring differences in the cost of medical services for different

[3] Apart from the uninsured, individuals who qualify for the full charge are those whose HI provider does not have a contract with the medical service provider: international visitors, people whose health plans are lacking contract with the service provider, etc. Uninsured contribute a majority of self-pay patients. Therefore, I neglect potential biases in the aggregate statistics due to a detailed categorization of self-pay patients.

[4] To extract rents, the insurers and providers prefer that their competitors do not know the prices.

payers in the data. To my knowledge, data on the amounts billed to and paid by the privately insured and the uninsured for a comparable set of medical services are not available. Nevertheless, disintegrated pieces of information on hospital care can provide an insight onto the gap in the rates. The rest of this section discusses frequently used financial statistics on inpatient hospital care on the aggregate level [5] and utilizes them to approximate the differences in the cost of health services for the privately insured and the uninsured.

A common statistics used in financial reports of the hospital industry is a charge-to-cost ratio (CCR):

$$\text{Charge-to-Cost Ratio (CCR)} = \frac{\text{Charge}}{\text{Cost}}.$$

CCR measures charges billed to patients (Charge) relative to the Medicare allowable cost of medical services (Cost). It should be noted that the latter Medicare allowable cost is the cost of providing care determined by the Center of Medicare and Medicaid Services (CMS) which applies to *all* patients, not only those covered by Medicare or Medicaid [6] . Therefore, the ratio measures the magnitude of the markup that the hospitals charge to all patients collectively.

The average CCRs for the hospitals in the U.S. for the period 2001 through 2012 are shown below in Figure 2.1. The series is constructed based on the information provided by the Healthcare Cost and Utilization Project for the national inpatient sample in HCUP Cost-to-Charge Ratio Files (2012). The average is calculated using

[5] According to National Health Expenditure Accounts (2012), hospital care accounted for 32.3% of the national health expenditures or 38.2% of the personal health care expenditures in the United States in 2012. Other personal care expenditures include professional services (26.6% of national health expenditures), nursing homes and home care (8.1%), retail outlet sales of medical products (12.8%) and other health, residential, and personal care including program administration and net cost of private health insurance, and government public health activities (14.9%).

[6] The cost calculated by CMS includes only cost associated with patient's care while hospitals may incur additional expenditures attributed to cost which are not related to the patient care directly (for example, cost of amenities).

hospital-specific CCRs from the files which are based on all-payer, inpatient cost and charge information from the detailed reports of hospitals to the CMS. For hospitals with missing information about hospital-specific CCR, I use group average CCR also reported in HCUP Cost-to-Charge Ratio Files (2012). The group average CCR represents a weighted average for the hospitals in the group (defined by state, urban/rural, investor-owned/other, and number of beds) using the proportion of beds in the group as the weight for each hospital.

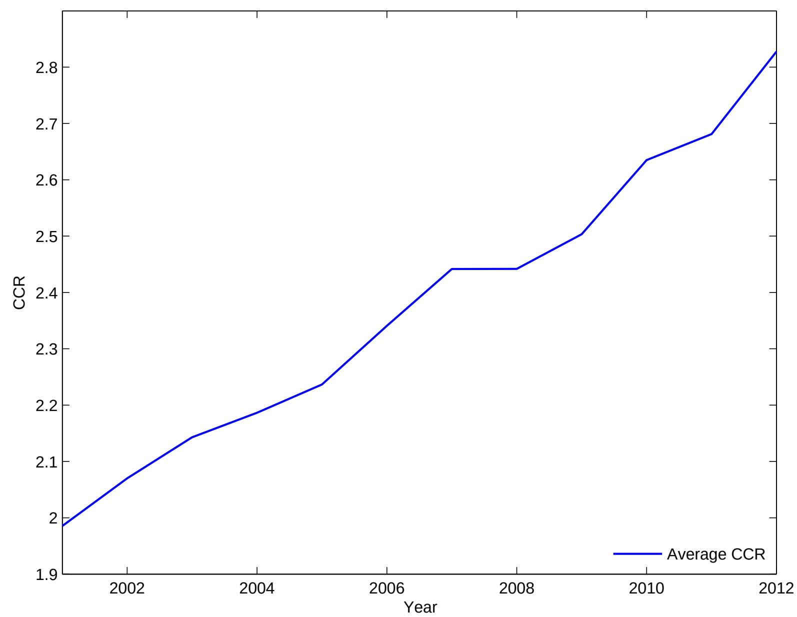

Figure 2.1: Average Charge-to-Cost Ratios for the U.S. Hospitals

The markup embedded in CCR can be viewed as an upper bound for the difference in the cost of health services faced by the insured and uninsured if the uninsured are billed and pay the charges in full according to CCR and insured face no markup. For example, Figure 2.1 illustrates that CCR for 2012 is approximately 2.8 which suggests that prices of medical services for the uninsured could be 2.8 times higher than for the insured. However, rather than postulating that the insured cover the cost with no markup, more careful analysis is needed which should take into account actual levels of reimbursing care by the insurance holders. This section proceeds inspecting additional statistics from the hospital industry to pin down health care payments relative to the cost for different categories of payers and, in particular, for the privately insured. Then, to infer the discounts for the privately insured in a final step, I compare estimated reimbursement level for the privately insured (payments relative to the cost) and average CCR in the data (throughout this section, the uninsured are assumed to pay full charges in accordance with CCR).

If hospitals were to collect a markup suggested by Figure 2.1, even after adjusting for the facts that 1) Medicare allowable costs might underestimate expenses perceived by hospitals as a cost and 2) markups for the outpatient care might be lower [7] , this would imply that the profit margin of the hospital industry (ratio of profits to revenues) must be high. For instance, if revenues were equal to the charges, the implied profit margin would be:

$$\text{Profit Margin} = \frac{\text{Revenue - Cost}}{\text{Revenue}} = \frac{\text{Charge - Cost}}{\text{Charge}} = 1 - \text{CCR}^{-1}.$$

Then, the average CCR for 2012 from Figure 2.1 would yield a profit margin of 64.6%. However, the data suggests it is much lower. Figure 2.2 exhibits aggregate total and operating hospital margins reported by the American Hospital Association (AHA) [8] :

[7] Outpatient care is a substantial part of hospital services. In the data, the proportion of outpatient services in the total revenues of the hospital industry has increased from 28% in 1994 to 46% in 2014. To my knowledge, statistics similar to CCR are not reported for the outpatient care. Therefore, I omit the discussion of outpatient care and assume that prices for outpatient care for different payers are determined in the same way as for the inpatient care.

[8] *Source*: Trendwatch Chartbook 2014. Trends Affecting Hospitals and Health System (2014).

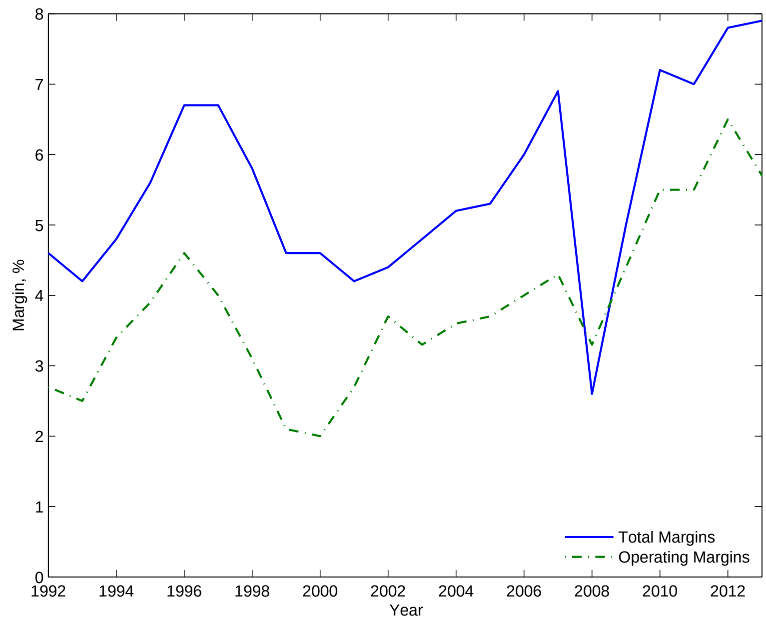

Figure 2.2: Aggregate Total and Operating Hospital Margins

The margins are calculated as the difference between total net revenue/operating revenue [9] and total expenses divided by total net revenue/operating revenue. According to Figure 2.2, aggregate total and operating margins for 2012 were 7.8% and 6.5% as opposed to 64.6% implied by the average CCR. Therefore, actual revenues are lower than the charges. Say differently, not all payers are subject to a full charge, and reimbursement levels for some of them can be significantly lower than CCR.

A statistics which determines reimbursement level for the payer is called a payment-to-cost ratio (PCR). Unlike CCR, PCR is a ratio of *payments received from a payer*, not billed, to the cost of providing care:

$$\text{Payment-to-Cost Ratio (PCR)} = \frac{\text{Payments}}{\text{Cost}}.$$

Generally, if PCRs for all categories of insured were known, then it would be straightforward to calculate the discounts comparing CCR and PCRs for each category. In

[9] Revenues are derived from activities related to the provision of health care for the patients.

the absence of the information about PCR for the privately insured, I retrieve it from profit margins using some assumptions outlined below.

The relationship between profit margin and PCRs for individual payer groups can be established if the profit margin is expressed in the following way:

$$\text{Profit Margin} = \frac{\sum_i (\text{Payments}_i - \text{Cost}_i) \cdot N_i}{\sum_i \text{Payments}_i \cdot N_i}$$

where Payments_i and Cost_i are, correspondingly, payments and cost for the unit of care for payer i, N_i is the number of units of care delivered to payer i. If the cost of a unit of care is the same for all payers $\text{Cost}_i = \text{Cost}$, which effectively means that the composition of received health care services does not differ across payers, the above expression for the profit margin can be written as:

$$\text{Profit Margin} = 1 - \frac{\text{Cost} \sum_i N_i}{\sum_i \text{Payments}_i \cdot N_i} = 1 - \left(\sum_i \frac{\text{Payments}_i}{\text{Cost}} \cdot \frac{N_i}{\sum_i N_i} \right)^{-1}.$$

The first term under the sum sign in the brackets is the PCR for group i. The second term represents the fraction of health care services received by group i. Then, it follows that:

$$\text{Profit Margin} = 1 - \left(\sum_i \text{PCR}_i \cdot f_i \right)^{-1}$$

where $f_i = \frac{N_i}{\sum_i N_i}$ is the proportion of health services provided to group i.

To find PCR for the privately insured relying on this expression, it suffices to know PCRs for all other categories of payers and the distribution of delivered care across payers. At the aggregate level, the following categories of payers are distinguished: Medicare and Medicaid holders, private payers including holders of private insurance and uninsured, uncompensated care, other government programs and non-patient costs. Next, I discuss some data and assumptions made regarding PCRs for these individual payers and the distribution of health care services.

As often claimed, a potential reason for low profitability margins of health care providers is the provision of care for Medicare and Medicaid beneficiaries which is reimbursed at lower rates as compared with other payers. In general, payment rates for Medicare and Medicaid with exception of managed care plans are set by the law rather than through a negotiation and may be even below the associated costs. Data for 2012 indicates that PCRs for Medicare and Medicaid [10] were correspondingly 85.9% and 88.9% (Trendwatch Chartbook 2014. Trends Affecting Hospitals and Health System (2014)). These ratios imply that hospitals received payment of 85.9 cents/ 88.9 cents for every dollar spent on Medicare/Medicaid patients in 2012 indicating underpayments for the care supplied under these programs [11] . Similar to Medicare and Medicaid, other government programs might be reimbursed at low rates. Thus, I assume that government programs and non-patient costs are compensated at the cost without markup (PCR is equal to 100%). As mentioned before, the uninsured are assumed to be subject to a full charge and pay in accordance with CCR (for calculations, I take PCR to be equal to the average CCR of 280% in 2012). Generally, hospitals may discount services for the uninsured on the case-by-case basis or provide charity care. However, most of the uninsured do not receive this privilege. A lot of times it happens because uninsured patients are not aware of such options or because negotiating prices on an individual basis with hospitals is considered a costly action making it more difficult to access care. Moreover, charity care may not necessarily be

[10] Payments here include managed care.

[11] At the same time, these programs are big which could substantially lower hospital's profit profits. According to Trendwatch Chartbook 2014. Trends Affecting Hospitals and Health System (2014), in 2012 Medicare and Medicaid contributed correspondingly 39.7% and 16.3% to the overall hospital care costs (include both inpatient and outpatient care). If one were to account for the proportion of costs and PCRs attributed to Medicare and Medicaid and assume that other payers pay in full in accordance with CCR, the implied profit margin for 2012 would be approximately approximately 42.3% which explains a large portion of the deviation of the observed profit margins in data from the one derived from CCR. However, the observed profit margin are far from being fully explained with the adjustment for underpayments by Medicare and Medicaid which suggests that payments from other payers, including privately insured, are lower than the charges.

aimed to only uninsured [12] and, at the aggregate level, it is included as a component of uncompensated care. Therefore, there is no convincing evidence that the uninsured face prices lower than the charges. Finally, I assume that uncompensated care is not reimbursed (PCR = 0).

I take the distribution of overall hospital care costs [13] across payers to represent the distribution of provided care. According to Trendwatch Chartbook 2014. Trends Affecting Hospitals and Health System (2014), in 2012 Medicare and Medicaid contributed correspondingly 39.7% and 16.3% to the overall hospital care costs along with 34% contributed by private payers, 6.1% by uncompensated care, 1.8% by other government programs and 2.2% by non-patient costs. To split the cost attributed to private payers between privately insured and uninsured, I use their share in the national inpatient costs. The distribution of national inpatient costs by primary payers bears a close similarity with the distribution of total costs including outpatient care. In 2012, 46% of inpatient costs were attributed to Medicare, 16% to Medicaid, 29% to private insurance, 5% to the uninsured, 4% to other costs. Note that private payers which include privately insured and uninsured account for exactly 34% of inpatient care as they do in the total costs including outpatient care. Thus, I assume that 5% of the overall hospital costs are spent on the care for the uninsured and 29% on the privately insured.

Expression for the profit margin combined with assumptions regarding rates of compensating care by different payers and the distribution of hospital care costs allows to compute that payments from privately insured patients (PCR) need to be 143.9% or 138.7% of the cost of care provided to them to obtain the profit margins

[12] For example, Uninsured and Overcharged: How Advocate Health Care Overcharges Chicago Hospital Patients (2003) mentions that only about half of charity care was supplied to the uninsured in Cook County in 2003.

[13] These costs include both inpatient and outpatient care.

observed in the data for 2012. Therefore, since the uninsured are assumed to pay a full charge or 280% of the cost, privately insured face health care prices 49.1% or 50.9% lower than the uninsured [14] . Section 2.5.1 uses these differences in the cost of medical services for the privately insured and uninsured to calibrate the gap in the cost of health services face by the employed and unemployed.

2.3 Simple Model

In this section, I write a simple model to offer insights onto the effects of higher prices of health services faced by the unemployed who lose access to HI upon leaving the employer on the labor market outcomes, aggregate labor productivity and aggregate output. The model modifies Lucas and Prescott (1974) island model with directed search by requiring all agents incur extra costs of medical services in every period. This cost is correlated with the employment status of the worker. More specifically, I explicitly assume that employed worker has access to HI through the employment and, thus, faces a lower cost of medical services. Once the worker leaves the island and becomes unemployed, she loses access to the private insurance, becomes uninsured until employed on an other island and faces a higher cost of health services. Discussion below suppresses time index for all variables since the analysis is focused on the comparison of stationary equilibria with reallocation of the workers between islands.

[14] There is some fragmented evidence in Uninsured and Overcharged: How Advocate Health Care Overcharges Chicago Hospital Patients (2003) confirming the discounts I obtain here. The average inpatient profit margin per uninsured discharge for Cook County residents was 55.1% and average provider insurance discount was 60% in 2001. This evidence is along the lines of profit margin of 64.6% on the uninsured and a discount of about 50% for the privately insured in the analysis conducted in this paper.

2.3.1 Environment

Consider an economy where ex-ante identical risk-neutral workers of measure 1 are hired by a continuum of labor markets, or islands, of measure 1. Here, the islands are defined as occupations. Each island has access to a decreasing return technology:

$$F(n, z) = zn^{\gamma}, \tag{2.1}$$

which is a function of island-specific productivity, z, and labor input, n, with a parameter of the decreasing return to scale γ, $\gamma < 1$. To obtain analytical results, in this section I simplify the exposition and assume that island's productivity fluctuates between two values z_1 and z_2, $z_1 < z_2$, in accordance with a transition function Q, s.t. $Q(z_i|z_i) = p_i$. Workers are assumed to own an equal share of aggregate profits Π and receive an equal lump-sum transfer from the government T.

There is no long-term employment relationship between an island and a worker. Thus, workers are rehired by the islands of worker's choosing in every period. If the worker stays on the island where she is currently located, she produces consumption good, receives island-specific wage and starts the next period in the same location. Otherwise, she needs to undergo one period of unemployment before joining her preferred island. During this period, the worker obtains no wage. Transitioning workers arrive to the destination island of their choosing at the end of the period, after production is complete. In other words, the unemployed workers do not produce in the current period and start the next period on the island of their choosing before realizations of the next period's productivity shocks are observed. This assumption implies that, in the current period, the island cannot employ more workers than the total number of workers located on the island at the beginning of the period, x.

In this economy, the state of the island, (x, z), is defined by the number of people who start period on the island before anyone leaves/arrives, x, and current period productivity level, z. Islands represent competitive labor markets implying that the workers take wages as given and island-specific wage is equal to the marginal product of the employed workers denoted as:

$$w(g(x,z),z) = F_1(g(x,z),z) = zf(g(x,z)) = z\gamma(g(x,z))^{\gamma-1},$$

where $g(x,z)$ is the employment level after workers' decisions to stay/leave have been made, $F_1(\cdot,\cdot)$ is a first order derivative of the production function F with respect to the first argument, and $f(x) = \gamma x^{\gamma-1}$.

Assume that, in every period, workers bear additional cost of health services perfectly correlated with their employment status. The main purpose for introducing this cost is to create an environment in which the cost of being unemployed are higher due to the loss of HI accompanied by a higher cost of health services. Denote this cost H_e if the worker is employed and $H_u > H_e$ if the worker is unemployed. Also, assume that H_u and H_e are collected by the government and distributed back to all workers through an equal lump-sum transfer T.

2.3.2 Stationary Equilibrium

In the stationary equilibrium with two levels of productivity shock, island's labor force (number of workers starting the period on the island) and employment lie in the set $\{x_1, x_2\}$, $x_2 > x_1$. In other words, labor force and employment switch back and forth between x_1 and x_2. The employment adjusts instantaneously to x_1 following negative productivity shock on the island with the labor force x_2 since the workers can leave the island before production takes place. In this case, the next period's labor force of an island is equal to x_1. However, the employment remains equal to x_1 on the island characterized by a labor force x_1 following positive productivity shock because the unemployed workers arrive on the island only at the end of the period. In this case, next period's labor force of an island is equal to x_2. Islands whose productivity does not change from previous to the current period maintain their employment and next period labor force equal to the current period labor force. Intuitively, if it was optimal to move to/leave these islands, it should have occurred in the previous periods. Therefore, the workers have no incentives to reallocate from/on these islands.

In a stationary equilibrium with reallocations, an island is in one of the following states:

1. (x_1, z_1): no workers leave and no arrive, the employment is $g(x_1, z_1) = x_1$ and the next period labor force is x_1;

2. (x_2, z_1): some workers leave and no arrive, the employment is $g(x_2, z_1) = x_1$ and the next period labor force is x_1;

3. (x_1, z_2): no workers leave and some arrive, the employment is $g(x_1, z_2) = x_1$ and the next period labor force is x_2;

4. (x_2, z_2): no workers leave and no arrive, the employment is $g(x_2, z_2) = x_2$ and the next period labor force is x_2.

In the stationary equilibrium, the mass of islands in each state is given by the stationary distribution $\lambda(x, z)$ and mobility of workers is directed from the islands (x_2, z_1) to the islands (x_1, z_2).

The values of the workers who have decided to stay on their islands are given by the following system of Bellman equations where employment levels and next period labor force for each island $(x, z)$are taken into account:

$$\begin{aligned} V(x_1, z_1) &= z_1 f(x_1) - H_e + \beta \left[p_1 V(x_1, z_1) + (1 - p_1) V(x_1, z_2) \right], \\ V(x_2, z_1) &= z_1 f(x_1) - H_e + \beta \left[p_1 V(x_1, z_1) + (1 - p_1) V(x_1, z_2) \right], \\ V(x_1, z_2) &= z_2 f(x_1) - H_e + \beta \left[p_2 V(x_2, z_2) + (1 - p_2) V(x_2, z_1) \right], \\ V(x_2, z_2) &= z_2 f(x_2) - H_e + \beta \left[p_2 V(x_2, z_2) + (1 - p_2) V(x_2, z_1) \right]. \end{aligned} \tag{2.2}$$

Here, the discount factor is denoted β. Additionally, I suppress worker's share in aggregate profits and government transfer because these are equal among the workers and, thus, have no effect on mobility decisions of risk-neutral workers.

As mentioned before, in the equilibrium with reallocations the workers move from the islands (x_2, z_1) to the islands (x_1, z_2). Therefore, the value of staying on the island (x_2, z_1) is equal to the value of being unemployed (the workers remaining on the island have to be indifferent between moving and staying). Also, the future discounted value of being unemployed is equal to the future discounted value of the worker on the island (x_1, z_2), the destination island of the transitioning worker (transitioning workers have to be indifferent between staying unemployed and being employed). Therefore, taking into account that the unemployed have to incur the cost of health services H_u, it follows that

$$V(x_2, z_1) = \underbrace{-H_u + \beta \left[p_2 V(x_2, z_2) + (1 - p_2) V(x_2, z_1) \right]}_{\text{Value of being unemployed}}, \tag{2.3}$$

where the right-hand side of the expression represent the value of being unemployed.

Combining (2.2) and (2.3) yields the following equilibrium relationship between x_1 and x_2 (a detailed equilibrium characterization is provided in Appendix B):

$$f(x_1)(z_1 + \beta(1 - p_1)z_2) - \beta p_2 z_2 f(x_2) + (H_u - H_e)(1 + \beta - \beta(p_1 + p_2)) = 0.$$

Given the reallocation technology (search friction) faced by the workers, the output maximizing allocation of workers across islands can be computed as a solution for the following problem:

$$\max_{x_1,x_2} \quad \lambda(x_1,z_1)F(x_1,z_1) + \lambda(x_1,z_2)F(x_1,z_2) + \lambda(x_2,z_1)F(x_1,z_1) + \lambda(x_2,z_2)F(x_2,z_2)$$

$$\text{s.t.} \;\; [\lambda(x_1,z_1) + \lambda(x_2,z_1)]\, x_1 + [\lambda(x_1,z_2) + \lambda(x_2,z_2)]\, x_2 = 1.$$

The last expression is a feasibility constraint which states that the number of workers located on the islands at the end of the period has to be equal to the total number of workers in the economy. As in the description of the equilibrium, there is no movement of workers on the islands (x_1, z_1) and (x_2, z_2). Therefore, these islands have correspondingly x_1 and x_2 workers at the end of the period, equal to what they have started with. The workers are reallocated away from the islands (x_2, z_1) which implies that islands in this state have x_1 workers at the end of the period. Finally, transitioning workers are allocated on the islands (x_1, z_2) which end the period with x_2 workers.

Given this problem of maximizing aggregate output, it is possible to show that (Appendix C contains a formal proof):

Proposition 1 *If β is sufficiently close to 1, then equilibrium allocation maximizes aggregate output when $H_u = H_e$.*

Proposition 1 implies that an ongoing reallocation in the competitive equilibrium fails to maximize aggregate output when workers take forward-looking mobility decisions if $H_u > H_e$. Intuitively, since workers face a higher cost of health services in the unemployment state, they attach a higher value to the employment on the islands with low productivity z_1 because the labor mobility occurring when the island experiences high productivity z_2 is costly. As a result, the allocation of labor across islands worsens which leads to a lower aggregate output.

To better understand the source of inefficiency, it is worth noting that an increased gap between H_u and H_e decreases unemployment. Because the workers attach a higher value to the islands which do not expose them to a costly unemployment, i.e. to the low productivity islands in this setup, the employment level x_1 on the islands (x_1, z_1), (x_2, z_1), (x_1, z_2) increases, and the employment x_2 on the islands (x_2, z_2) goes down. Therefore, the mobility of workers, as well as unemployment, decreases. This property is demonstrated numerically in Figure 2.3 where I vary value of H_u for a fixed set of other parameters. Generally, an arbitrary set of parameters generating an equilibrium with reallocations could be used to show a decrease in unemployment when H_u increases as compared with H_e. Thus, I omit a detailed discussion of the choice of the parameter values and only mention that parameters are selected to resemble U.S. economy and match the targets discussed in Section 2.5.1 (details on the parameterization of the two-period model are provided in Appendix D).

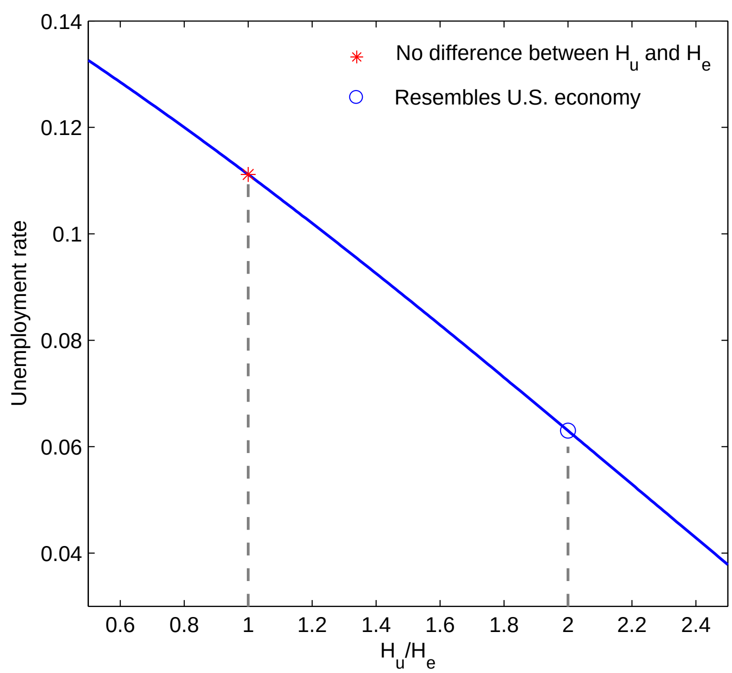

Figure 2.3: Equilibrium Unemployment for Different Values of H_u in the Two-Period Model

While decrease in unemployment on its' own could generate an increase in the output, labor productivity (output per worker) decreases when H_u increases. The primary reason for a decrease in labor productivity is that the allocation of labor across islands deteriorates. To see why, notice that the first-best allocation occurs when the marginal products (costs) of labor are equalized across islands. Then, even when $H_u = H_e$, the search friction associated with a period of unemployment during the transition results in the gap in the marginal costs of labor between departure and destination islands. As a result, the allocation of labor across islands moves away from its' first-best and aggregate labor productivity decreases. A higher cost of health services for the unemployed further increases the cost associated with unemployment and attaches workers to the islands that they would leave otherwise. As a results, the gap in the marginal costs of labor between departure and destination island is amplified. Therefore, the labor allocation across islands deteriorates and labor productivity goes down even further. To demonstrate this property, Figures 2.4 and 2.5 display variance in the marginal costs of labor across islands and aggregate labor productivity for different levels of H_u (values are reported relative to those under $H_u = H_e$).

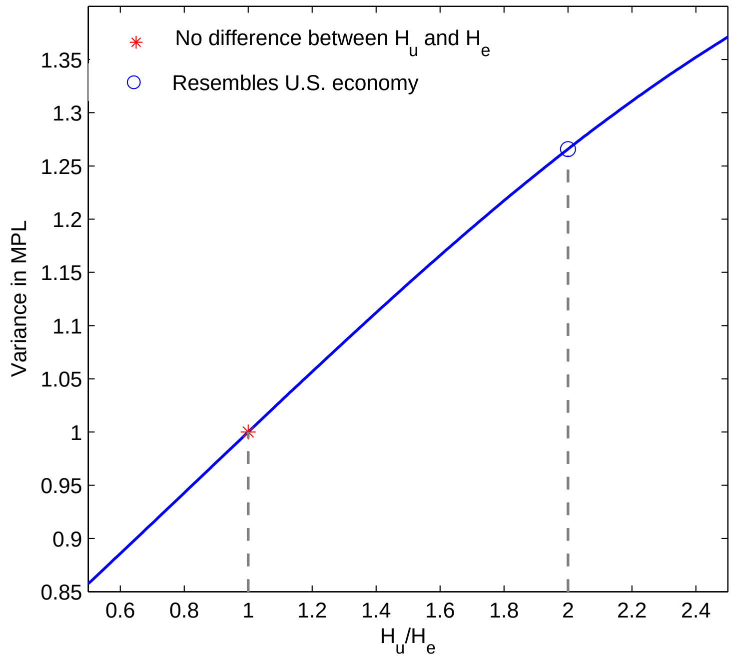

Figure 2.4: Variance in the Marginal Costs of Labor Across Islands for Different Values of H_u in the Two-Period Model

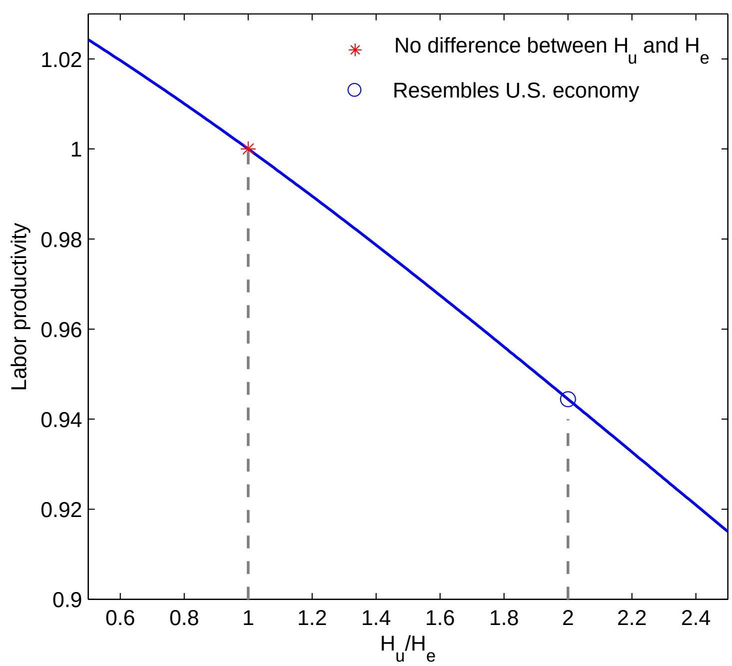

Figure 2.5: Labor Productivity for Different Values of H_u in the Two-Period Model

Figures 2.4 shows that variance in the marginal cost of labor across islands increases in H_u signaling about worse allocation of workers as compared with the first-best. These changes in the allocation of workers are reflected in the lower labor productivity in Figures 2.5 which decreases as H_u increases.

The two forces, decreased unemployment and decreased labor productivity, operate in opposite directions and create a natural source of ambiguity regarding the effects on aggregate output. Nevertheless, when $H_u > H_e$, a decreased labor productivity dominates and the aggregate output decreases in H_u. This result is demonstrated in Figure 2.6. Additionally, analytical results in Appendix E confirm graphical analysis performed in this section [15] .

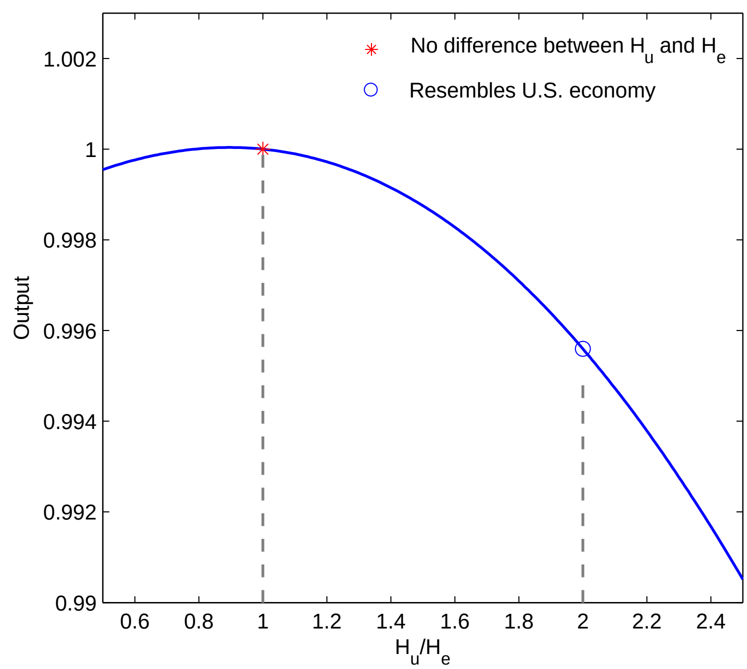

Figure 2.6: Aggregate Output for Different Values of H_u in the Two-Period Model

[15]Comparative statics results for the aggregate output are obtained for the case when β is close to 1.

It is worth noting that the value of the β used to obtain results in Figures 2.3 - 2.6 is equal to 0.9898, which is distinctly different but close to 1. Nevertheless, Figure 2.6 demonstrates that, even in the presence of discounting effects, the output maximizing level of H_u is close to H_e.

Needless to say, that two-state process for the productivity of the island in this section is a simplification allowing to highlight the trade-off between effects on unemployment and labor productivity produced by the gap between H_u and H_e. In a more realistic setup, a richer structure for the productivity process of the island could be desirable. At the same time, analytical results suggesting that output maximizing allocation occurs when H_u is close to H_e and, thus, that equalizing H_u and H_e leads to an increase in aggregate output, are conditional on a simple case of two-state productivity process in the absence of discounting. In a framework with a discounting, which is further enriched by a more complex productivity shock process, the impact of equalizing the H_u and H_e is harder to predict. Therefore, to quantify the effects of equalizing H_u and H_e in the data, in the next section I frame a full model which features more complex structure of the productivity shock process, as well as unemployment insurance payments to the unemployed. Inclusion of the latter unemployment insurance is justified not only because it is an integral component of the U.S. economic system, but also because it affects the differences in the net non-wage payments received/paid by the employed and unemployed which might influence worker's mobility decisions. Then, I parameterize the full model so that predictions of the model are consistent with observations for the U.S. economy and evaluate the effects of the policy eliminating the gap in the cost of health services for the employed and unemployed on unemployment, aggregate labor productivity and aggregate output.

2.4 The Main Model

This section considers a full equilibrium model which, along with varying costs of medical services faced by individuals in different employment states, incorporate unemployment insurance benefits.

2.4.1 Economy

The model resembles the main features of the economic environment described in Section 2.3. Namely, the production takes place on the islands using labor as an input in a decreasing return to scale production technology. At the beginning of a period after shocks are realized workers decide whether to stay on the island or leave it. If a worker decides to stay, then she engages in the production, receives labor income and pays costs of medical services H_e. If worker decides to leave, she becomes unemployed in the current period and arrives to the island of her choice at the end of the period after production stage is complete. The unemployed are required to pay the cost of health services H_u. Different from Section 2.3, I assume that every unemployed receives unemployment benefits b.

As before, production function of an island if of the form (2.1) and island-specific productivity shocks follow Markov process given by the transition function $Q(z'|z)$. The state of an island is described by a pair (x, z) where x is a population (labor force) on the island at the beginning of the period before workers leave, and z is a current period productivity of the island. Wage on the island (x, z) is denoted $w(g(x, z), z)$ and is determined as $w(g(x, z), z) = F_1(g(x, z), z)$ where $g(x, z)$ is the equilibrium employment level, $F_1(\cdot, \cdot)$ is a first order derivative of the production function F with respect to the first argument. Production function (2.1) implies that wages are given by:

$$w(g(x,z),z) = z\gamma(g(x,z))^{\gamma-1}. \tag{2.4}$$

The value of the worker finding herself on the island (x, z) at the beginning of the period, denoted by $V(x, z)$, represents her decision whether to stay or leave. If the worker leaves, she becomes unemployed. The value of leaving the island is equal to the present value of unemployment benefits net of the cost of health services incurred while being unemployed, H_u, plus future discounted value of employment on the island of her choice, denoted by θ and determined in equilibrium. If the worker decides to stay, she earns labor income, pays cost of health services H_e and begins next period in the same location.

The problem faced by the worker on the island (x, z) is described by the following Bellman equation:

$$V(x,z) = \max\left\{b - H_u + \theta, w(g(x,z),z) - H_e + \beta E\left[V(x',z')|x,z\right]\right\}.$$

where z' stands for the next period's realization of the productivity shock, and x' denotes the next period's labor force available on the island (number of workers starting the next period on the island).

For a given value of θ, consider the following cases which might occur.

Case 1. No workers leave and no arrive. Then, $x' = x$, $g(x, z) = x$ and yjr value of the worker is

$$V(x,z) = w(x,z) - H_e + \beta E\left[V(x,z')|x,z\right].$$

Lack of new arrivals means that continuation value of a worker on the island is lower than continuation value of the unemployed θ (otherwise, some unemployed would arrive on the island). In other words,

$$\beta E\left[V(x,z')|x,z\right] \leq \theta.$$

Case 2. All workers stay and some arrive. Then, $x' > x$, $g(x,z) = x$ and the value of the worker starting period in this location is

$$V(x,z) = w(x,z) - H_e + \beta E\left[V(x',z')|x,z\right].$$

Workers arrive until the continuation value of a worker on the island is equalized to the continuation value of unemployed:

$$\beta E\left[V(x',z')|x,z\right] = \theta.$$

which determines next period's population on the island, x', and permits a simpler expression for the value of a worker:

$$V(x,z) = w(x,z) - H_e + \theta.$$

Case 3. Some workers leave and no arrive. Then, $x' = g(x,z) < x$ and the value of the worker starting on the island is given by

$$V(x,z) = b - H_u + \theta.$$

In this case, workers leave until the present discounted value of the worker staying on the island is equalized to the value of being unemployed, i.e.:

$$w(x',z) - H_e + \beta E\left[V(x',z')|x,z\right] = b - H_u + \theta.$$

This equation determines the next period's population on the island, x'.

Combining all these cases, one arrives at the following Bellman equation which describes the problem faced by the worker on the island (x,z):

$$V(x,z) = \max\left\{b - H_u + \theta, w(x,z) - H_e + \min\left\{\theta, \beta E\left[V(x,z')|x,z\right]\right\}\right\}. \tag{2.5}$$

2.4.2 Equilibrium

For a given value of θ, let the distribution of islands at a point of time be given by $\lambda_t(x, z; \theta)$. Then, there exists next period's distribution of islands, $\lambda_{t+1}(x, z; \theta)$, consistent with optimal mobility decisions of individuals and transition function for productivity shocks. If optimal decision rules involve reallocations of agents across islands, then there exists a unique stationary distribution of islands $\lambda(x, z; \theta)$. The economy-wide labor market clearing condition equates total labor force of 1 and population (labor force) of all islands which pins down the value of θ in equilibrium:

$$\int\int x\lambda(x, z; \theta)dxdz = 1. \tag{2.6}$$

Definition. An equilibrium is a set of functions $V(\cdot)$ representing the value function of a worker, $g(\cdot)$ representing optimal island's employment, $\lambda(\cdot)$ describing the distribution of islands across states, $w(\cdot)$ representing wages and the continuation value of unemployment θ such that:

- Given the price function $w(\cdot)$ and the value of being unemployed θ, functions $V(\cdot)$ and $g(\cdot)$ maximize individual's utility (solve equation (2.5)).

- Wages $w(\cdot)$ on every island are determined competitively, i.e. the workers on every island are paid their marginal product of labor in accordance with (2.4).

- Invariant distribution $\lambda(\cdot)$ is consistent with island's optimal employment levels and transition functions Q.

- Economy-wide labor market clearing condition (2.6) is satisfied.

- The government budget is balanced:

$$T + (1 - U)H_e + UH_u - Ub = 0,$$

where U is the mass of unemployed workers (workers transitioning between islands), T is government transfer.

2.5 Policy Experiment

The primary interest of this section is to study implications of varying cost of health services for the unemployment rate, aggregate labor productivity and aggregate output. It is accomplished by comparing steady state outcomes of the benchmark economy featuring different costs of health services for the worker of different employment statuses with the outcomes of the economy where policies towards eliminating gaps in the costs of medical services are implemented.

There is a number of ways in which these gaps may be removed. The most obvious one is to provide HI to all uninsured. The negotiated prices would then apply universally and the argument of different rates would lose its' relevance. This is a direction in which Affordable Care Act is able to address higher rates charged to the uninsured by providing easier access to HI. Possible alternatives include a requirement for all payers to pay a single rate listed on the chargemaster files of medical care providers or subsidizing differences in the rates for the uninsured.

The first of the suggested policies implies lowering down prices toward their lower end while the second option is different in bumping up the prices for all payers towards their upper end. Third option preserves the gap in the rates while equalizing cost of health services through subsidies. In the environment considered here, it is not essential what kind of policy to study. Equal cost of health services, independent on the policy approach, have no effect on worker's mobility decisions and allocative efficiency if financed through non-distortionary taxes whenever necessary [16] . There-

[16] In the last policy scenario, the way in which the subsidy is financed may play an important role for the quantitative results. Nevertheless, since the focus of the paper is to understand the effects of equalizing the cost of health services across workers rather than to study its' interaction with a

fore, in the numerical experiment I choose to fix the cost of health services for the employed at the benchmark level and vary the cost of health services for the unemployed. It is assumed that lower cost for the unemployed are made possible due to subsidies financed through the equal lump-sum taxes on all workers.

2.5.1 Calibration

This section describes the choice of parameter values for the benchmark economy. The values are chosen to match selected observations on the U.S. economy including unemployment and different costs of health services faced by the workers in different employment statuses in line with previous findings.

Besides the choice of model period, the following parameters are to be determined in the calibration: discount factor β; the extent of the decreasing return to scale γ; cost of health services for the employed and unemployed, H_e and H_u; transition function of island's productivity $Q(z'|z)$ and the unemployment insurance benefits b.

I choose model period to be a quarter. This implies that job reallocations will involve at least one quarter of unemployment. Though it is possible that average duration of unemployment in the model economy will be substantially higher, this choice is intended to capture a relatively short duration of unemployment in the U.S. In the data, unemployment duration between 1983 to 2008, a period of a stable variation in the duration of unemployment, averaged to 16.1 weeks [17] or 1.34 model periods.

---

particular taxation scheme, it is reasonable to consider a subsidy financed through non-distortionary taxes.

[17] The data is retrieved from FRED, iterative data tool provided by the Federal Reserve Bank of St. Louis at `https://fred.stlouisfed.org/series/LNU03008275#0`.

I set two parameters following standard practice without solving the model. The discount factor is set to be $\beta = 0.9898$ which corresponds to the annual interest rate of 4%. Parameter of the decreasing return to scale γ is set to be equal to 0.85 [18] .

The remaining parameters are calibrated jointly. This means that the model is solved for different sets of parameters. The set for which the model predictions match empirical targets from the data discussed below is selected.

Turning to the cost of health services, H_e and H_u, recall that Section 2.2 finds that the discounted rates for the privately insured are about 50% of that for the uninsured. For the calibration, I assume that the cost of health services for the employed workers who have access to ESHI are exactly 50% of the cost faced by the unemployed, $H_e = 0.5H_u$, an average between discounts found in Section 2.2. The level of H_e is chosen such that the ratio of the total health expenditures to GDP in the model economy matches the share of health consumption expenditures [19] to GDP in the data. The national health expenditures are constantly growing in the U.S. Therefore, rather than matching data on health expenditures for a particular year, I target the average share of health consumption expenditures in GDP for a period from 2001 to 2013 equal to 15.2% [20] .

The transition matrix Q is taken to approximate the following AR(1) process:

$$\ln(z_t) = \rho \ln(z_{t-1}) + \epsilon_t,$$

[18] This choice is in line with papers, for example, Restuccia and Rogerson (2008), Brugemann and Manovskii (2009), which use the same extent of the decreasing return to scale in the establishment level production function.

[19] Matching the share of health consumption expenditures rather than national health expenditures leaves aside the investment component of the latter. It is arguable if investments are financed through the payments of individuals. Thus, I choose to not take into account this part of health expenditures. To shed light on the size of the investment component, it should be noted that investments account for about 6% of total health expenditures or 1% of GDP.

[20] Data on health care expenditures come from National Health Expenditure Accounts (2012). GDP annual series are retrieved from Bureau of Economic Analysis interactive data tool, `https://www.bea.gov/index.htm`.

where

$$\epsilon_t \sim N(0, \sigma_\epsilon^2),\ a \geq 0,\ 0 \leq \rho < 1.$$

Discrete approximation of this Markov process is made using Tauchen's method (Tauchen (1986)) on a grid of evenly-spaced 20 grid points over the interval $[-3.5\sqrt{\sigma_\epsilon^2/(1-\rho^2)},\ 3.5\sqrt{\sigma_\epsilon^2/(1-\rho^2)}]$, where $\sqrt{\sigma_\epsilon^2/(1-\rho^2)}$ is a standard deviation of the process of $\ln(z_t)$.

The persistence of the process for the productivity shocks ρ is selected to match the average unemployment rate of 6.3% for the period 2000-2012 [21] . Choosing variance σ^2 of log productivity process, I seek to capture the extent of frictional wage dispersion in the data. One of the commonly used summary statistics of wage dispersion is a mean-min ratio, which is the average wage divided by the lowest observed wage. Hornstein *et al.* (2007) use various methodologies and data sets to measure mean-min ratios in the data and arrive to the range of estimates between 1.5 and 2. An appealing value chosen by these authors for their analysis is 1.7, which I also target here with the choice of σ^2.

Finally, I parameterize unemployment benefits b to match the average replacement ratio in the data for the period 2001-2012 which is equal to 40.9%. Note that this index measures the average benefits amount relative to the average wage rather than average individual replacement ratio equal to 46.7% for the same time period. Therefore, the model counterpart of the replacement ratio is the ratio of the unemployment benefits b to the average wage.

Table 2.1 summarizes calibrated parameters, their definitions, targeted moment in the data and the implied parameter values:

[21] Data sources: unemployment series `https://data.bls.gov/pdq/SurveyOutputServlet`.

Parameter	Target	Value
Discount factor, β	Interest rate	0.9898
Extent of the decreasing return to scale, γ	Restuccia and Rogerson (2008)	0.85
Cost of health services for the employed, H_e	Health consumption expenditures relative to GDP	0.1513
Cost of health services for the unemployed, H_u	Calculated from $H_e = 0.5H_u$	0.3025
AR(1) coefficient of the process of $\ln(z_t)$, ρ	Unemployment rate	0.7736
Variance of the process of $\ln(z_t)$, σ_ϵ^2	Mean-min wage ratio	0.0257
Unemployment benefits, b	Average replacement ratio	0.3688

Table 2.1: Model Parameters and Targets

Table 2.2 summarizes the targeted moments in the data and displays their counterparts in the calibrated model:

Target	Data	Model
Health consumption expenditures relative to GDP	0.152	0.152
Unemployment rate	0.063	0.062
Mean-min wage ratio	1.7	1.7
Average replacement ratio	0.409	0.409

Table 2.2: Calibration Targets, Data vs. Model Predictions

The model matches the selected targets well. Notably, it replicates the dispersion of wages as measured by the mean-min ratio when the latter is directly targeted. It is known to be a hard task for a wide spectrum of models of equilibrium unemployment to generate an amount of frictional wage dispersion observed in the data. For example, Hornstein *et al.* (2007) point out that standard search and matching models generate about 20-fold lower mean-min ratio, when the latter is not targeted. In this regards, calibration procedure can claim only a partial success because mean-min ration was directly targeted while other unmatched moments deviate from the data. For example, the calibrated model predicts that duration of unemployment is 1.0124 model periods while the data suggests it is 1.34 (exceeds model counterpart by about 4 weeks). To match observations on duration of unemployment, the model would need to be augmented, perhaps, deviating from the fully directed search. However, whether the augmented model could jointly match the duration of unemployment and variation in wages is an open question.

2.5.2 Numerical Experiment

This section evaluates effects of the regulation eliminating differences in the cost of health services across workers on unemployment rate, aggregate labor productivity and aggregate output. In model terms, the numerical experiment refers to setting $H_u = H_e$ enacted in the above model for the calibrated values of parameters. The results come from contrasting the performance of the benchmark economy featuring $H_e = 0.5H_u$, as calibrated in the previous section, and the new stationary equilibrium with $H_u = H_e$. Table 2.3 shows the results. Along with the main aggregates of interest, the table also reports measures of wage dispersion (variance and mean-min ratio), subsidy-to-output ratio under the policy regime $H_u = H_e$, health expenditures, aggregate output net of health expenditures and duration of unemployment. All values except unemployment, subsidy-to-output ratio and duration of unemployment are normalized by their corresponding benchmark values.

Variable	Benchmark	$H_u = H_e$
Unemployment rate	0.062	0.0765
Aggregate output	1	0.9974
Labor productivity	1	1.0118
Variance in wages	1	0.9120
Mean-min wage ratio	1	0.9703
Health expenditures	1	0.9442
subsidy-to-output ratio		0.0109
Aggregate output net of health expenditures	1	1.0062
Duration of unemployment	1.0124	1.0246

Table 2.3: Long-Run Effects of Eliminating Differences in the Cost of Health Services

Table 2.3 shows that equalizing cost of health services across workers leads to a decrease in aggregate output by 0.26%. The lower output comes from the interaction of two effects operating in opposite directions. One the other hand, the policy increases unemployment rate from 6.2% to 7.45% which adversely affects aggregate production. On the one hand, it improves labor productivity measured as output per employed worker by 1.18% which increases aggregate output. However, a higher productivity created by the policy cannot compensate a decrease in the output due to lower employment, and aggregate output goes down.

Increase in unemployment rate caused by the policy is straightforward to understand. A decrease in the cost of health services for the unemployed when H_u is set at H_e level makes transition between jobs less painful which incentivizes people to switch jobs and, thus, increases unemployment. In line with this reasoning, unemployment goes up from 6.2% to 7.65%. Additionally, as cost of health services for the unemployed decreases, the average unemployment duration increases from 1.0124 to

1.0246. This response agrees with the fact that, upon arrival on the island and observing a new productivity, the workers tend to accept a job less easily since the option of moving to another island is more attractive when the cost of being unemployed are lower.

Increase in labor productivity bears a lot upon how the workers are allocated across islands. To better understand the effect of the policy on the allocation of workers across islands, below I discuss effects on worker's allocation and productivity pertaining to the search friction. Next, I turn to explain how the policy equalizing the cost of health services changes the allocation of workers and leads to the increase in labor productivity.

In this environment, mobility of workers between islands is associated with worker flows moving from the islands suffering low productivity shocks to the islands experiencing high productivity shocks. As some workers leave the islands with low productivity, marginal products of labor adjust and wages for the workers remaining on the island increase. As a results of this adjustment, some workers optimally decide to stay on the island they would otherwise leave if wages were fixed since wages become high enough to overweight benefits of transitioning to another island. In the equilibrium scenario of a frictionless economy, such model mechanics would lead to an economy-wide wage rate consistent with economy-wide labor market clearing condition. In the world with a search friction, loosing labor income and going through the costly unemployment prevents some departures from the islands experiencing low productivity and decreases arrivals to the islands experiencing high productivity. As a result, low productivity islands hire more worker and high productivity ones hire less workers compared to the frictionless world, and variation in marginal products of labor across islands arises. Obviously, emerging dispersion in marginal products of labor lowers labor productivity as compared with the first-best allocation where the same workers are allocated across islands so that the marginal products are equalized.

This discussion suggests that any policy increasing the cost of being unemployed amplifies the adverse effect of the search friction on the allocation of workers across islands and lowers labor productivity. Higher cost of being unemployed prevents more worker from moving to another island, and leads to more workers allocated to low productivity islands and less workers allocated to high productivity islands. Consequently, dispersion in marginal products of labor/wages across islands increases signaling about worse allocation of workers across islands and lower productivity.

Conversely, a policy lowering the cost of being unemployed, as the one considered in this paper, drives a higher willingness of workers to reallocate to new jobs after their island is hit by the low productivity shock. In other words, it mitigates the sources driving worker's unwillingness to move from the low productivity shock islands to high productivity ones. As a result, the pattern of worker's allocation across islands is altered: less worker are willing to stay on the low productivity islands and more workers settle on the high productivity ones. Say differently, the allocation of workers across islands moves towards the allocation in the frictionless world. Figure 2.7 below illustrates the distribution of employment across islands characterized by different productivity levels in the benchmark economy, under the policy setting $H_u = H_e$ and in the first-best. As seen from the figure, the allocation of workers under the policy $H_u = H_e$ is shifted to the right compared with the benchmark, in the direction closer to that in the frictionless economy.

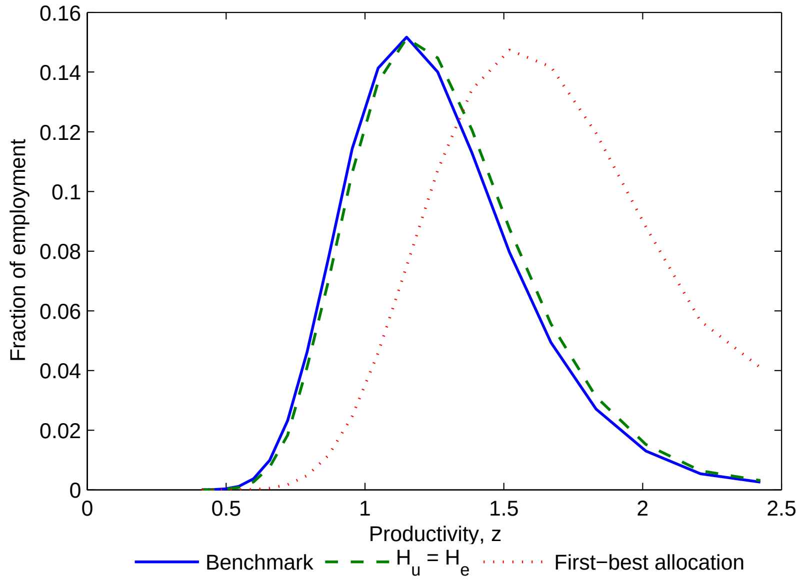

Figure 2.7: Distribution of Employment Across Islands with Different Productivity Levels

As more workers leave low productivity islands and stay on the high productivity ones, the dispersion in the marginal products of labor across islands decreases which signals about a better allocation of workers across islands which leads to an increase in labor productivity. These effects can be seen in Table 2.3 demonstrating that, with the introduction of the policy, variance in wages (marginal costs of labor) and mean-min ration decrease by correspondingly 8.8% and 2.97% and labor productivity increases by 1.18%.

To summarize the effects of the policy, the numerical results show that, when the cost of health services are equalized across workers, losses due to higher rotation of the workers between jobs (higher unemployment) induced by a lower cost of unemployment overtake gains in labor productivity due to a better allocation of workers across jobs (less workers on low productivity islands and more workers on the high productivity ones). As a result, aggregate output decreases by 0.26%.

2.6 Concluding Remarks

The link between access to HI and employment is an essential attribute of the U.S. health care system which is often viewed as a serious distortion influencing aggregate economic outcomes. In this paper, I studied effects of one distortion induced by this link - correlation between employment status of the worker and the cost of health services. On the one hand, the cost of health services are higher if the patient does not have HI. On the other hand, majority of workers receive HI through the employment and lose access to it if they leave the employer. As a result, the unemployed face higher cost of health services which affects worker's incentive to reallocate between jobs and impacts the performance of the economy. The focus of this paper is on a quantitative exploration of the effects of higher cost of health services for the unemployed on unemployment, aggregate labor productivity and aggregate output.

On methodological ground, I build the analysis upon classical theoretical framework of Lucas and Prescott (1974) extended to include cost of health services for the workers which are correlated with worker's employment status. Numerical experiment of equalizing these costs across workers in the calibrated model shows that the aggregate output decreases by 0.26%. Nevertheless, output decrease is hard to conjecture a priori as there operates an important trade-off. First, lowering the cost of health services for the unemployed enables workers to search more which increases unemployment and decreases aggregate output. Second, it permits workers to pursue search and undertake more productive jobs which improves the allocation of resources and raises aggregate output. In the model environment of this paper, improvement in allocative efficiency is reflected in higher labor productivity and is a result of a better allocation of workers across existing jobs rather than due to new and more productive jobs being created [22] . In turn, the allocation of labor existing jobs improves due to less workers being hired by low productivity islands and more workers being hired by high productivity ones which decreases variance in the marginal cost of labor across islands. As a result, the increased unemployment is accompanied by a counterbalancing force of improved labor productivity which, however, in the quantitative experiment of this paper cannot overweight reduction in output induced by higher unemployment.

[22] A possibility that unemployment insurance may raise aggregate output by inducing creation of riskier but more productive jobs is discussed in Acemoglu and Shimer (1999) and Acemoglu and Shimer (2000).

REFERENCES

Acemoglu, D. and R. Shimer, "Efficient Unemployment Insurance", Journal of Political Economy **107**, 5, 893–928 (1999).

Acemoglu, D. and R. Shimer, "Productivity Gains from Unemployment Insurance", European Economic Review **44**, 7, 1195–1224 (2000).

Bhattacharya, D., N. Guner and G. Ventura, "Distortions, Endogenous Managerial Skills and Productivity Differences", Review of Economic Dynamics **16**, 1, 11–25 (2013).

Brugemann, B. and I. Manovskii, "Fragility: Quantitative Analysis of the US Health Insurance System", Working Paper (2009).

Buera, F., B. Moll and Y. Shin, "Well-Intended Policies", Review of Economic Dynamics. **16**, 1, 216–230 (2013).

Choi, E. J. and J. R. Spletzer, "The Declining Average Size of Establishments: Evidence and Explanations", Monthly Labor Review pp. 50–65 (2012).

Driscoll, A. K. and A. B. Bernstein, "Health and Access to Care Among Employed and Unemployed Adults: United States, 2009-2010", NCHS Data Brief , 83 (2012).

Employer Health Benefits Survey, Annual Survey, The Henry J. Kaiser Family Foundation, `http://files.kff.org/attachment/report-2015-employer-health-benefits-survey` (2015).

Gabler, A. and M. Poschke, "Experimentation by Firms, Distortions, and Aggregate Productivity", Review of Economic Dynamics **16**, 1, 26–38 (2013).

Guner, N., G. Ventura and Y. Xu, "Macroeconomic Implications of Size-Dependent Policies", Review of Economic Dynamics **11**, 4, 721–744 (2008).

HCUP Cost-to-Charge Ratio Files, Healthcare Cost and Utilization Project (HCUP). Agency for Healthcare Research and Quality, `http://www.hcup-us.ahrq.gov/db/state/costtocharge.jsp` (2001-2012).

Hopenhayn, H. A., "Entry, Exit, and Firm Dynamics in Long Run Equilibrium", Econometrica **60**, 5, 1127–1150 (1992).

Hopenhayn, H. A., "On the Measure of Distortions", Working Paper (2012).

Hopenhayn, H. A. and R. Rogerson, "Job Turnover and Policy Evaluation: A General Equilibrium Analysis", The Journal of Political Economy **101**, 5, 915–938 (1993).

Hornstein, A., P. Krusell and G. L. Violante, "Frictional Wage Dispersion in Search Models: A Quantitative Assessment", Working Paper 13674, National Bureau of Economic Research, `http://www.nber.org/papers/w13674` (2007).

Huang, K. X. D. and G. W. Huffman, “Unemployment and Welfare Implications of the Current U.S. Tax Treatment of Employer-Provided Medical Insurance”, Macroeconomic Dynamics **18**, 07, 1547–1580 (2014).

Jeske, K. and S. Kitao, “U.S. Tax Policy and Health Insurance Demand: Can a Regressive Policy Improve Welfare?”, Journal of Monetary Economics **56**, 2, 210–221 (2009).

Lee, Y. and T. Mukoyama, “Entry, Exit and Plant-Level Dynamics”, Working Paper, Feferal Reserve Bank of Cleveland (2008).

Long, S. and M. Marquis, “Robert Wood Johnson Foundation Employer Health Insurance Survey [Community Tracking Study and State Initiatives in Health Care Reform Program], 1997 [Computer File]”, ICPSR02935-v2, Washington, DC: RAND Corporation [producer], 2000. Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor], 2005-07-06, doi:10.3886/ICPSR02935.v2 (1997).

Lucas, R. J. and E. C. Prescott, “Equilibrium Search and Unemployment”, Journal of Economic Theory **7**, 2, 188–209 (1974).

National Health Expenditure Accounts, Center for Medicare and Medicaid Services, `http://www.cms.gov/Research-Statistics-Data-and-Systems/Statistics-Trends-and-Reports/NationalHealthExpendData/index.html` (2012).

Options for Reducing The Deficit: 2014 to 2023, Congressional Budget Office, Congress of The United States, `https://www.cbo.gov/sites/default/files/cbofiles/attachments/44715-OptionsForReducingDeficit-3.pdf` (2013).

Prescott, E. C., “Why Do Americans Work So Much More Than Europeans?”, Federal Reserve Bank of Minneapolis Quarterly Review pp. 2–13 (2004).

Restuccia, D. and R. Rogerson, “Policy Distortions and Aggregate Productivity with Heterogeneous Establishments”, Review of Economic Dynamics **11**, 4, 707–720 (2008).

Roys, N. and F. Gourio, “Size-Dependent Regulations, Firm Size Distribution, and Reallocation”, 2013 Meeting Papers 199, Society for Economic Dynamics, `https://ideas.repec.org/p/red/sed013/199.html` (2013).

Tauchen, G., “Finite State Markov-Chain Approximations to Univariate and Vector Autoregressions”, Economics Letters **20**, 2, 177–181 (1986).

Trendwatch Chartbook 2014. Trends Affecting Hospitals and Health System, American Hospital Association, `http://www.aha.org/research/reports/tw/chartbook/index.shtml` (2014).

Uninsured and Overcharged: How Advocate Health Care Overcharges Chicago Hospital Patients, SEIU Hospital Accountability Project, http://www.thenightministry.org/070_facts_figures/030_research_links/040_healthcare/Discriminatory_Pricing__DP_Advocate.pdf (2003).

Veracierto, M., "Employment Flows, Capital Mobility, and Policy Analysis", Working Paper, Federal Reserve Bank of Chicago (2000).

Vereshchagina, G., "Between-Firm Redistribution of Profit in Competitive Industries: Why Labor Market Policies May Not Work", Working Paper (2005).

APPENDIX A

ADDITIONAL BENCHMARK STATISTICS FOR CHAPTER 1

	Age Category			
	≤ 1 years	2 years	3 to 5 years	≥ 6 years
Average size, *data* [1]	7.94	10.70	12.61	22.16
Average size, *model*	7.94	9.86	11.39	21.03
Fraction of establishments sponsoring HI, *data* [2]	0.2233	0.2873	0.3131	0.5411
Fraction of establishments sponsoring HI, *model*	0.2233	0.2677	0.3178	0.6495

Table A.1: Average Size of Establishment and HI Provision, by Establishment Age

[1] *Source*: calculated from Business Employment Dynamics (BED) data, year 2000.

[2] *Source*: calculated using the Survey data, Long and Marquis (1997).

APPENDIX B

CHARACTERIZATION OF COMPETITIVE EQUILIBRIUM IN THE TWO-PERIOD MODEL

To find the solution of the model, I solve the system of Bellman equations for the workers having decided to stay on the island (2.2) and the unemployed (2.3) stated in Section 2.3.2. Then, I obtain a stationary distribution of islands across individual states, $\lambda(x, z)$. Finally, I characterize the equilibrium in the economy evoking economy-wide labor market clearing condition.

Expressions (2.2) and (2.3) state that:

$$\begin{aligned}
V(x_1, z_1) &= z_1 f(x_1) - H_e + \beta \left[p_1 V(x_1, z_1) + (1 - p_1)V(x_1, z_2)\right], \\
V(x_2, z_1) &= z_1 f(x_1) - H_e + \beta \left[p_1 V(x_1, z_1) + (1 - p_1)V(x_1, z_2)\right], \\
V(x_1, z_2) &= z_2 f(x_1) - H_e + \beta \left[p_2 V(x_2, z_2) + (1 - p_2)V(x_2, z_1)\right], \\
V(x_2, z_2) &= z_2 f(x_2) - H_e + \beta \left[p_2 V(x_2, z_2) + (1 - p_2)V(x_2, z_1)\right]
\end{aligned}$$

and

$$V(x_2, z_1) = -H_u + \beta \left[p_2 V(x_2, z_2) + (1 - p_2)V(x_2, z_1)\right].$$

First two equations are the same which reduces the system to

$$V(x_2, z_1) = z_1 f(x_1) - H_e + \beta \left[p_1 V(x_1, z_1) + (1 - p_1)V(x_1, z_2)\right], \tag{B.1}$$
$$V(x_1, z_2) = z_2 f(x_1) - H_e + \beta \left[p_2 V(x_2, z_2) + (1 - p_2)V(x_2, z_1)\right], \tag{B.2}$$
$$V(x_2, z_2) = z_2 f(x_2) - H_e + \beta \left[p_2 V(x_2, z_2) + (1 - p_2)V(x_2, z_1)\right], \tag{B.3}$$
$$V(x_2, z_1) = -H_u + \beta \left[p_2 V(x_2, z_2) + (1 - p_2)V(x_2, z_1)\right]. \tag{B.4}$$

From (B.3),

$$V(x_2, z_2) = \frac{1}{1 - \beta p_2} \left(z_2 f(x_2) - H_e + \beta(1 - p_2)V(x_2, z_1)\right).$$

Combining the above expression with (B.4) yields:

$$V(x_2, z_1) = \frac{1}{1 - \beta} \left[-(1 - \beta p_2)H_u - \beta p_2 H_e + \beta p_2 z_2 f(x_2)\right]. \tag{B.5}$$

Now, combine (B.2) and (B.4) to get

$$V(x_1, z_2) = z_2 f(x_1) - H_e + H_u + V(x_2, z_1)$$

and substitute it in (B.1):

$$V(x_2, z_1) = \frac{1}{1 - \beta} \left[\beta(1 - p_1)H_u - (1 - \beta(1 - p_1))H_e + f(x_1)(z_1 + \beta(1 - p_1)z_2)\right]. \tag{B.6}$$

Equating (B.5) and (B.6) yields an equilibrium relationship between employment levels x_1 and x_2 as given in equation (2.3.2) in Section 2.3.2:

$$f(x_1)(z_1 + \beta(1 - p_1)z_2) - \beta p_2 z_2 f(x_2) + (H_u - H_e)(1 + \beta - \beta(p_1 + p_2)) = 0. \tag{B.7}$$

Now, I proceed to derive a stationary distribution of islands across individual states, $\lambda(x, z)$.

First, it should be noted that the equilibrium is characterized by the steady state distribution of islands across productivities z_1 and z_2. The mass of islands experiencing productivity shock z_i is denoted $\lambda(z_i)$, $i \in \{1, 2\}$, and found by solving the following system of equations:

$$\begin{bmatrix} \lambda(z_1) & \lambda(z_2) \end{bmatrix} \begin{bmatrix} p_1 & 1-p_1 \\ 1-p_2 & p_2 \end{bmatrix} = \begin{bmatrix} \lambda(z_1) & \lambda(z_2) \end{bmatrix}.$$

The solution to the system yields:

$$\lambda_1 = \frac{1-p_2}{2-p_1-p_2}, \quad \lambda_2 = \frac{1-p_1}{2-p_1-p_2}.$$

Recall, that islands in the state (x_1, z_2) are the islands that have just experienced a positive productivity shock which happens with probability $1 - p_1$. Therefore, the mass of islands in the state (x_1, z_2) is:

$$\lambda(x_1, z_2) = \frac{(1-p_1)(1-p_2)}{2-p_1-p_2}$$

which also implies that

$$\lambda(x_1, z_1) = \frac{p_1(1-p_2)}{2-p_1-p_2}.$$

Analogously, the mass of islands $\lambda(x_2, z_1)$ and $\lambda(x_2, z_2)$ are given by:

$$\lambda(x_2, z_1) = \frac{(1-p_1)(1-p_2)}{2-p_1-p_2}, \quad \lambda(x_2, z_2) = \frac{p_2(1-p_1)}{2-p_1-p_2}.$$

The labor market clearing condition declares that the number of workers located on the islands at the end of the period have to be equal to the total number of workers in the economy. There is no movement of workers on the islands (x_1, z_1) and (x_2, z_2). Therefore, they have correspondingly x_1 and x_2 at the end of the period equal to what they have started with. The workers are reallocated from the islands (x_2, z_1) implying that the islands started the period with x_2 workers have x_1 at the end of the period. Finally, the moving workers allocate on the island (x_1, z_2) at the end of the period. Therefore, these islands have x_2 workers at the end of the period. Thus, the labor market clearing condition is:

$$[\lambda(x_1, z_1) + \lambda(x_2, z_1)]\, x_1 + [\lambda(x_1, z_2) + \lambda(x_2, z_2)]\, x_2 = 1,$$

or

$$\lambda(z_1)x_1 + \lambda(z_2)x_2 = 1. \tag{B.8}$$

Solving the system of equations (B.7) and (B.8) yields equilibrium employment levels of x_1 and x_2.

APPENDIX C

PROOF OF PROPOSITION 1 IN CHAPTER 2

The problem of maximizing aggregate output in Section 2.3.2 is written as:

$$\max_{x_1,x_2} \quad \lambda(x_1,z_1)F(x_1,z_1)+\lambda(x_1,z_2)F(x_1,z_2)+\lambda(x_2,z_1)F(x_1,z_1)+\lambda(x_2,z_2)F(x_2,z_2)$$
$$\text{s.t.} \quad [\lambda(x_1,z_1)+\lambda(x_2,z_1)]\,x_1+[\lambda(x_1,z_2)+\lambda(x_2,z_2)]\,x_2=1.$$

As the distribution of islands across individual states in the steady state of the competitive equilibrium does not depend on the employment levels x_1 and x_2, $\lambda(x,z)$ is this problem is characterized by the expressions derived in Appendix B. Then, the feasibility constraint can be written as (B.8) from Appendix B which also implies that $\frac{\partial x_2}{\partial x_1}=-\frac{\lambda(z_1)}{\lambda(z_2)}$.

Then, the F.O.C. to the above problem of maximizing aggregate output becomes:

$$(z_1\lambda(x_1,z_1)+z_2\lambda(x_1,z_2)+z_1\lambda(x_2,z_1))\,f(x_1)-\frac{\lambda(z_1)}{\lambda(z_2)}\lambda(x_2,z_2)f(x_2)=0.$$

Substituting expressions for $\lambda(x,z)$ and $\lambda(z)$ derived in Appendix B gives an output-maximizing level of x_1 as a solution to:

$$f(x_1)(z_1+(1-p_1)z_2)-p_2z_2f(x_2)=0. \tag{C.1}$$

where x_2 is obtained from (B.8) in Appendix B.

Recall, that the relationship between equilibrium levels of x_1 and x_2 is given by:

$$f(x_1)(z_1+\beta(1-p_1)z_2)-\beta p_2z_2f(x_2)+(H_u-H_e)(1+\beta-\beta(p_1+p_2))=0. \tag{C.2}$$

Comparing (C.1) and (C.2), one can conclude that, if the discount factor is close to 1, then (C.2) implies (C.1) is $H_u=H_e$. Say differently, if $\beta\to 1$, then competitive equilibrium allocation maximizes aggregate output when $H_u\to H_e$.

APPENDIX D

PARAMETERIZATION OF THE TWO-PERIOD MODEL

This section describes the choice of parameter values for the two-period model used for graphical analysis in Section 2.3.

Majority of the parameters are chosen to match observations for the U.S. economy serving as a basis for calibration of the full model in Section 2.5.1. More specifically, I choose model period to be a quarter and set the discount factor to be $\beta = 0.9898$. Parameter of the decreasing return to scale γ is taken to be equal 0.85. It is assumed that $H_u = 0.5H_e$ and H_e is chosen to match the ratio of health consumption expenditures to GDP.

The transition matrix Q is constructed in the following way. The value of $p_1 = Prob(z_1|z_1)$ is chosen so that the unemployment in the model is equal to the average unemployment rate of 6.3% in the data. The value of $p_2 = Prob(z_2|z_2)$ is chosen to mimic the average duration of unemployment of 16.1 weeks (1.34 model periods).

Finally, the model is constructed for two productivity levels, z_1 and z_2. In the full model, the grid for z is approximated on the interval $\pm 3.5 \cdot$ standard deviation around the mean of the process for $\ln(z)$. So, once the parameters of the process are picked, the grid for the productivity is known and there is no need to match levels of z to any target. In the two-period model, z_1 and z_2 are free parameters which have to be disciplined to match some targets. In this regards, I normalize the value of z_1 to be 1. The value of z_2 is selected to match wage mean-min ratio of 1.7 (this target is used in Section 2.5.1 to calibrate variance of the process for $\ln(z)$ in the full model).

Table D.1 summarizes chosen parameter values:

Parameter	Value
Discount factor, β	0.9898
Extent of the decreasing return to scale, γ	0.85
Cost of medical services for the employed, H_e	0.2440
Cost of medical services for the unemployed, H_u	0.4880
$Prob(z_1\|z_1)$, p_1	0.7875
$Prob(z_2\|z_2)$, p_2	0.7475

Table D.1: Parameter Values in the Two-Period Model

It should be noted, that all targeted values are matched precisely in the two-period model.

APPENDIX E

COMPARATIVE STATICS

In this Appendix, I provide a comparative statics results for some equilibrium objects to support the discussion in Section 2.3.2. First, I show that, unemployment decreases in H_u. Second, I verify that the allocation of labor deteriorates and moves away from the fist-best allocation. This part is accomplished by demonstrating that the gap in the marginal products of labor increases in H_u compared to no gap in the first-best. Finally, I show that when $\beta \to 1$ and $H_u > H_e$, aggregate output decreases in H_u. The latter also establishes that an increase in employment is not enough to offset decrease in labor productivity as H_u increases and, hence, aggregate output decreases.

I start demonstrating that unemployment level decreases in H_u. The number of people departing from each individual island (x_2, z_1) is equal to $x_2 - x_1$. At the same time, the mass of islands from which the departures happen is given by $\lambda(x_2, z_1)$. Then, the number of the unemployed in the economy is:

$$U = \lambda(x_2, z_1)(x_2 - x_1).$$

Thus,

$$\frac{\partial U}{\partial H_u} = \lambda(x_2, z_1)\left(\frac{\partial x_2}{\partial H_u} - \frac{\partial x_1}{\partial H_u}\right). \tag{E.1}$$

Express x_2 from the labor market clearing condition (B.8):

$$x_2 = 1 - \frac{\lambda(z_1)}{\lambda(z_2)}x_1,$$

which implies that:

$$\frac{\partial x_2}{\partial H_u} = -\underbrace{\frac{\lambda(z_1)}{\lambda(z_2)}}_{>0}\frac{\partial x_1}{\partial H_u}.$$

Therefore, the sign of $\frac{\partial x_2}{\partial H_u}$ is just the opposite of the sign of $\frac{\partial x_1}{\partial H_u}$.

To determine the sign on $\frac{\partial x_1}{\partial H_u}$, substitute expression for x_2 in (B.7) and denote left-hand side of (B.7) as $G(x_1)$:

$$G(x_1) = f(x_1)(z_1 + \beta(1-p_1)z_2) - \beta p_2 z_2 f(1 - \frac{\lambda(z_1)}{\lambda(z_2)}x_1) + (H_u - H_e)(1 + \beta - \beta(p_1 + p_2)).$$

Then, according to Implicit Function Theorem,

$$\frac{\partial x_1}{\partial H_u} = -\left(\frac{\partial G}{\partial x_1}\right)^{-1}\frac{\partial G}{\partial H_u} = -\frac{1 + \beta - \beta(p_1 + p_2)}{f'(x_1)(z_1 + \beta(1-p_1)z_2) + \beta p_2 z_2 f'(1 - \frac{\lambda(z_1)}{\lambda(z_2)}x_1)}.$$

Because production function $F(x, z)$ is a decreasing return to scale, it follows that $f'(x) < 0$ and, hence, implies that the denominator of this fraction is negative. Notice that $1 + \beta - \beta(p_1 + p_2) \geq 1 + \beta - 2\beta = 1 - \beta > 0$ which implies that the nominator of the fraction is positive. Therefore,

$$\frac{\partial x_1}{\partial H_u} > 0 \quad \Rightarrow \quad \frac{\partial x_2}{\partial H_u} < 0.$$

In words, it implies that the employment on the islands (x_1, z_1), (x_1, z_2), (x_2, z_1) increases while that on the island (x_2, z_2) decreases. It also follows from expression (E.1) that $\frac{\partial U}{\partial H_u} < 0$, i.e. unemployment decreases in H_u.

Now, I demonstrate that an increase in H_u worsens the allocation of labor and increases the gap in the marginal products of labor across the islands. First, notice that the first-best allocation occurs at the point when the marginal products of labor of the islands are equalized:

$$z_1 f(x_1^*) = z_2 f(x_2^*) \quad \text{or} \quad f(x_1^*) = \frac{z_2}{z_1} f(x_2^*),$$

where x_1^* and x_2^* denote the first-best employment on the islands experiencing productivity shocks z_1 and z_2.

Express $f(x_1)$ from the equilibrium condition (B.7):

$$f(x_1) = \frac{\beta p_2 z_2}{z_1 + \beta(1-p_1)z_2} f(x_2) - (H_u - H_e)\frac{1+\beta-\beta(p_1+p_2)}{z_1+\beta(1-p_1)z_2}.$$

If $H_u = H_e$, then the condition above implies that:

$$f(x_1) = \frac{\beta p_2 z_2}{z_1 + (1-p_1)z_2} f(x_2) < \frac{z_2}{z_1 + (1-p_1)z_2} f(x_2) < \frac{z_2}{z_1} f(x_2). \tag{E.2}$$

Since labor market clearing condition (B.8) is satisfied in both first-best and equilibrium allocations and $f'(x)$ is decreasing in x, the equilibrium employment levels satisfy:

$$x_1 > x_1^* \quad \text{and} \quad x_2 < x_2^*$$

which essentially gives a deviation of marginal products of labor from each other as defined by E.2.

It was shown before that x_1 increases and x_2 decreases in H_u. Then, because $f'(x)$ is decreasing in x, it translates into an increase in the gap between left-hand and right-hand side of inequality (E.2), or an increase in the gap between marginal products of labor across islands. As a result, the allocation of labor worsens compared to the first-best.

Finally, I show that the output decreases in H_u when β is close to 1 and $H_u > H_e$. The present value of the economy's output Y is given by the objective function of output maximization problem in Appendix C. Differentiating the objective function with respect to H_u yields:

$$\frac{\partial Y}{\partial H_u} = \left(z_1\lambda(x_1,z_1) + z_2\lambda(x_1,z_2) + z_1\lambda(x_2,z_1)\right) f(x_1)\frac{\partial x_1}{\partial H_u} + \lambda(x_2,z_2) f(x_2)\frac{\partial x_2}{\partial H_u}.$$

Using expressions for $\lambda(x, z)$ and evoking labor market clearing condition to find $\frac{\partial x_2}{\partial H_u}$, one can show that:

$$\frac{\partial Y}{\partial H_u} = \left(f(x_1)(z_1 + (1-p_1)z_2) - p_2 z_2 f(x_2)\right)\frac{\partial x_1}{\partial H_u}.$$

As argued before, $\frac{\partial x_1}{\partial H_u} > 0$. At the same time, when $\beta \to 1$, it follows from the equilibrium condition (B.7) that

$$f(x_1)(z_1 + (1-p_1)z_2) - p_2 z_2 f(x_2) \to \underbrace{-(H_u - H_e)(2 - p_1 - p_2)}_{<0 \text{ if } H_u > H_e}.$$

Hence, $\frac{\partial Y}{\partial H_u} < 0$ when $\beta \to 1$ and $H_u > H_e$, i.e. aggregate output decreases in H_u.