



Publicly Accessible Penn Dissertations

1-1-2014

Essays on Health Insurance Market Design and Labor Market Interactions

Naoki Aizawa

University of Pennsylvania, naizawa@sas.upenn.edu

Follow this and additional works at: <http://repository.upenn.edu/edissertations>



Part of the [Economics Commons](#)

Recommended Citation

Aizawa, Naoki, "Essays on Health Insurance Market Design and Labor Market Interactions" (2014). *Publicly Accessible Penn Dissertations*. 1187.

<http://repository.upenn.edu/edissertations/1187>

This paper is posted at ScholarlyCommons. <http://repository.upenn.edu/edissertations/1187>

For more information, please contact libraryrepository@pobox.upenn.edu.

Essays on Health Insurance Market Design and Labor Market Interactions

Abstract

This dissertation aims to develop empirical frameworks to assess a variety of health insurance market policies and explore the optimal policy design taking into account their impacts on the labor market and public insurance program.

The first chapter (co-authored with Hanming Fang) presents and empirically implements an equilibrium labor market search model where risk averse workers facing medical expenditure shocks are matched with firms making health insurance coverage decisions. We use our estimated model to evaluate the equilibrium impact of the 2010 Affordable Care Act (ACA) and find that it would reduce the uninsured rate among the workers in our estimation sample from 20.12% to 7.27%.

The second chapter evaluates the current health insurance exchange (HIX) system implemented under the ACA and examines its optimal design, accounting for adverse selection and equilibrium labor market interactions. I develop and empirically implement a life cycle equilibrium labor market search model integrated with the pre-ACA health insurance market. Counterfactual experiments show that the ACA decreases not only the uninsured rate but also aggregate labor productivity. Next, I examine the optimal design of HIX by choosing the values of three major design components---tax penalties on the uninsured, premium subsidies and age-based rating regulations. I find that the optimal combination of these components makes it less beneficial for older workers relative to younger workers to purchase health insurance from HIX. Implementing the optimal structure leads to higher labor productivity and a slightly lower uninsured rate.

The third chapter (co-authored with You Suk Kim) studies the incentives for private insurers to use advertising to attract low-cost, healthy individuals and the impacts of advertising on selection, competition, and welfare in the context of the Medicare Advantage (MA). We develop and estimate an equilibrium model of the MA market, which incorporates strategic advertising by insurers. We find that advertising has positive effects on overall demand, but a much larger effect on the demand of the healthy individuals. Moreover, we find that advertising accounts for 15% of the selection of healthier individuals into MA. The impact of risk adjustment policies is also examined.

Degree Type

Dissertation

Degree Name

Doctor of Philosophy (PhD)

Graduate Group

Economics

First Advisor

Hanming Fang

Subject Categories

Economics

ESSAYS ON HEALTH INSURANCE MARKET DESIGN AND LABOR MARKET
INTERACTIONS

Naoki Aizawa

A DISSERTATION

in

Economics

Presented to the Faculties of the University of Pennsylvania in Partial
Fulfillment of the Requirements for the Degree of Doctor of Philosophy

2014

Supervisor of Dissertation

Hanming Fang, Professor of Economics

Graduate Group Chairperson

George Mailath, Walter H. Annenberg Professor in the Social Sciences

Dissertation Committee:

Andrew Shephard, Assistant Professor of Economics

Holger Sieg, The J.M. Cohen Term Professor of Economics

Kenneth I. Wolpin, Walter H. and Leonore C. Annenberg Professor in the Social Sciences

ESSAYS ON HEALTH INSURANCE MARKET DESIGN AND LABOR MARKET
INTERACTIONS

©COPYRIGHT

2014

Naoki Aizawa

Acknowledgements

First and foremost, I would like to express my sincerest thanks to my main advisor Hanming Fang for his constant encouragement, support and guidance. He was very helpful from the first day that I took his class. He has greatly helped my development as a researcher in many ways—as an instructor in his second year Ph.D. class; as a co-author in our joint project; as a supervisor for my teaching assistant job in his undergraduate class; and as the main advisor of my dissertation.

I am also grateful to my dissertation committee members, Professors Andrew Shephard, Holger Sieg, and Kenneth I. Wolpin. I have benefited tremendously from my interactions with them, and I owe a great debt to them. I also thank Professors Petra Todd, Robert Town, and Yuichi Yamamoto for their encouragement and supports at various stages of my dissertation. Moreover, I benefited from Professors Aislinn Bohren, Xu Cheng, Harold Cole, Francis Diebold, Cecilia Fielner, Camilo Garcia-Jimeno, Jonathan Kolstad, Dirk Krueger, George Mailath, Iourii Manovskii, Antonio Merlo, Guillermo Ordonez, Mallesh Pai, Andrew Postlewaite, Katja Seim, Xun Tang, and Rakesh Vohra for their comments and suggestions to my dissertation and their assistance during my job market process.

Special thanks is expressed to my classmate You Suk Kim. We shared the office during our first year and then our collaboration eventually led to a joint paper which is the third chapter of my dissertation. I would also like to thank other fellow students, especially Diego Amador, Sina Ates, Luigi Bocola, Michael Chirico, Jaesung Choi, Nils Gornemann, Nicolas Grau, Ilwoo Hwang, Mitsuru Katagiri, Rossa O’Keeffe-O’Donovan, Felipe Saffie,

Hikaru Saijo, Minchul Shin, Dongho Song, Jan Tilly, Chamna Yoon, and Naoki Wakamori for their time discussing about my research. Without their help, this dissertation could not be completed.

A dissertation fellowship from the Center for Retirement Research at the Boston College, Grant 5001537 from the Social Security Administration, the Russell Ackoff Fellowship from the Wharton Risk Center at the University of Pennsylvania, and the Maloof Family Dissertation Fellowship from the Department of Economics at the University of Pennsylvania are gratefully acknowledged.

Finally, I thank my parents, Masatoshi and Sumiko, my brother Takeshi, and my grandmother Misao for their understanding that I have pursued this career and for their constant encouragement and unconditional love during my days at Penn. I dedicate this dissertation to my grandfather Susumu.

ABSTRACT

ESSAYS ON HEALTH INSURANCE MARKET DESIGN AND LABOR MARKET INTERACTIONS

Naoki Aizawa

Hanming Fang

This dissertation aims to develop empirical frameworks to assess a variety of health insurance market policies and explore the optimal policy design taking into account their impacts on the labor market and public insurance program.

The first chapter (co-authored with Hanming Fang) presents and empirically implements an equilibrium labor market search model where risk averse workers facing medical expenditure shocks are matched with firms making health insurance coverage decisions. We use our estimated model to evaluate the equilibrium impact of the 2010 Affordable Care Act (ACA) and find that it would reduce the uninsured rate among the workers in our estimation sample from 20.12% to 7.27%.

The second chapter evaluates the current health insurance exchange (HIX) system implemented under the ACA and examines its optimal design, accounting for adverse selection and equilibrium labor market interactions. I develop and empirically implement a life cycle equilibrium labor market search model integrated with the pre-ACA health insurance market. Counterfactual experiments show that the ACA decreases not only the uninsured rate but also aggregate labor productivity. Next, I examine the optimal design of HIX by choosing the values of three major design components—tax penalties on the uninsured, premium subsidies and age-based rating regulations. I find that the optimal combination of these components makes it less beneficial for older workers relative to younger workers to purchase health insurance from HIX. Implementing the optimal structure leads to higher labor productivity and a slightly lower uninsured rate.

The third chapter (co-authored with You Suk Kim) studies the incentives for private

insurers to use advertising to attract low-cost, healthy individuals and the impacts of advertising on selection, competition, and welfare in the context of the Medicare Advantage (MA). We develop and estimate an equilibrium model of the MA market, which incorporates strategic advertising by insurers. We find that advertising has positive effects on overall demand, but a much larger effect on the demand of the healthy individuals. Moreover, we find that advertising accounts for 15% of the selection of healthier individuals into MA. The impact of risk adjustment policies is also examined.

Contents

Acknowledgements	iii
Abstract	v
Contents	vii
List of Tables	xii
List of Figures	xvi
1 Equilibrium Labor Market Search and Health Insurance Reform	1
1.1 Introduction	1
1.2 Related Literature	8
1.3 An Equilibrium of Model of Wage Determination and Health Insurance Provision	11
1.3.1 The Environment	11
1.3.2 Timing in a Period	16
1.3.3 Analysis of the Model	17
1.3.4 Steady State Equilibrium	28
1.4 Qualitative Assessment of the Model	29
1.4.1 Numerical Simulations	29
1.4.2 Comparative Statics	33
1.4.3 Identification of γ, d, C and σ_f	35

1.5	Data Sets	36
1.5.1	Survey of Income and Program Participation	37
1.5.2	Medical Expenditure Panel Survey (MEPS)	38
1.5.3	Robert Wood Johnson Foundation Employer Health Insurance Survey	38
1.5.4	Summary Statistics	39
1.6	Estimation Strategy	42
1.6.1	First Step	44
1.6.2	Second Step	47
1.6.3	Deriving the Likelihood Functions of Workers' Labor Market Dy- namics	48
1.6.4	Employer-Side Moments	50
1.7	Estimation Results	51
1.7.1	Parameter Estimates	51
1.7.2	Within-Sample Goodness of Fit	54
1.8	Counterfactual Experiments	58
1.8.1	Model for the Counterfactual Experiments	60
1.8.2	Parameterization of the Counterfactual Policies	65
1.8.3	Main Result	68
1.8.4	Assessing the Effects of Components of the ACA	77
1.8.5	Role of Tax Exemption for ESHI Premium	78
1.8.6	Other Counterfactual Experiments	81
1.9	Conclusion	86

2 Health Insurance Exchange Design in an Empirical Equilibrium Labor

	Market Model	89
2.1	Introduction	89
2.2	Model	99
2.2.1	Environment	99

2.2.2	Analysis of the Model	106
2.3	Data Set	119
2.3.1	Survey of Income and Program Participation	120
2.3.2	Medical Expenditure Panel Survey (MEPS)	121
2.3.3	Kaiser Family Employer Health Benefits Survey	122
2.3.4	Descriptive Statistics	122
2.4	Estimation	123
2.4.1	Identification	123
2.4.2	Estimation Strategy	125
2.5	Estimation Result	128
2.5.1	Model Fit	131
2.6	Counterfactual Experiments: Evaluating the ACA and the components of HIX	136
2.6.1	The Model for evaluating the ACA	136
2.6.2	The Main Results	140
2.7	Normative Analysis: Optimal Design of Health Insurance Exchanges . . .	150
2.7.1	The Government Problem	151
2.7.2	Main Results	155
2.7.3	Discussion: the role of Medicaid expansion	163
2.8	Conclusion	164

3 Advertising Competition and Risk Selection in Health Insurance Markets: Evidence from Medicare Advantage 166

3.1	Introduction	166
3.2	Background on Medicare Advantage	172
3.3	Data and Preliminary Analysis	174
3.3.1	Data	174
3.3.2	Preliminary Analysis	182

3.4	Model	186
3.4.1	Demand	188
3.4.2	Supply	192
3.5	Identification and Estimation	194
3.5.1	Demand	195
3.5.2	Supply	198
3.5.3	Estimation Algorithm	201
3.6	Estimates	202
3.6.1	Utility	202
3.6.2	Cost	206
3.7	Counterfactual Experiments	209
3.7.1	Ban of Advertising	209
3.7.2	Risk Adjustment	212
3.8	Conclusion	215
A Appendix to Chapter 1		217
A.1	Numerical Algorithm to Solve the Equilibrium of the Benchmark Model	217
A.2	Omitted Formula in Section 1.6.1	221
A.3	Derivation of Likelihood Function	222
A.4	Estimation Procedure	224
A.5	Equilibrium of the Counterfactual Economy	225
A.5.1	Steady State Equilibrium for the Post-Reform Economy	225
A.5.2	Numerical Algorithm for the Counterfactual Policy Experiments	226
A.5.3	Steady State Distribution of Employment in Counterfactual Experiments	229
B Appendix to Chapter 2		234
B.1	Numerical Algorithm	234

B.2	Omitted Derivations in the Counterfactual Experiments	236
B.2.1	Value Function	236
B.2.2	Steady State Worker Distribution	238
B.3	Parameterization of Policy Parameters	241
B.3.1	Penalties associated with individual mandate	241
B.3.2	Penalties associated with employer mandate	242
B.3.3	Medicaid expansion	243
B.3.4	Premium subsidies in HIX	244
B.3.5	Age-based pricing regulation	245
C	Appendix to Chapter 3	246
C.1	Constructing Health Status	246
	Bibliography	252

List of Tables

1.1	Predictions of the Baseline Model: Benchmark and Comparative Statistics.	30
1.2	Understanding Why High-Productivity Firms Are More Likely to Offer Health Insurance than Low Productivity Firms.	32
1.3	Summary Statistics: SIPP 1996.	40
1.4	Summary Statistics: Comparison between MEPS 1997-1999 and SIPP 1996.	41
1.5	Summary Statistics: RWJ-EHI 1997.	43
1.6	Summary Statistics of the Subsample of the MEPS 1997-1999 Used in the Estimation of Medical Expenditure Distributions in the First Step.	46
1.7	Parameter Estimate from Step 1.	52
1.8	Parameter Estimate from Step 2.	53
1.9	Fit for Medical Expenditure Distributions: Model vs. Data.	55
1.10	Worker-Side Moments in the Labor Market: Model vs. Data.	56
1.11	Employer-Side Moments: Model vs. Data.	57
1.12	Counterfactual Policy Experiments: Evaluation of the ACA and its Two Variations.	69
1.13	Counterfactual Policy Experiments: Evaluation of Various Components of the ACA (1).	75
1.14	Counterfactual Policy Experiments: Evaluation of Various Components of the ACA (2).	76
1.15	Counterfactual Policy Experiments: Evaluating the Effects of Eliminating the Tax Exemption for EHI Premium under the Benchmark and the ACA.	79

1.16	Counterfactual Policy Experiments: Evaluation of Alternative Policy Arrangements.	82
2.1	Summary Statistics: SIPP 2004.	123
2.2	Summary Statistics for MEPS 2004-2007.	123
2.3	Summary Statistics: Kaiser 2004-2007.	124
2.4	Parameter Estimate for Individual Preference	128
2.5	Parameter Estimate for Latent Medical Expenditure Shocks and Health Transition Process	129
2.6	Parameter Estimate for Individual Labor Market Activities	130
2.7	Parameter Estimate for the Distribution of Workers	130
2.8	Parameter Estimate for Firm-side Characteristics	131
2.9	Model fit: mean 4-month wage conditional on coverage, age and education.	131
2.10	Model fit: coverage rate among the employed conditional on age and education.	132
2.11	Model fit: employment rate conditional on age and education.	132
2.12	Model fit: health status conditional on age and education.	132
2.13	Model fit: mean health status conditional on insurance, employment and education.	133
2.14	Model fit: Annual health transition	134
2.15	Model fit: mean annual medical expenditure conditional on age and education and health.	135
2.16	model fit for firm-side moment	135
2.17	Counterfactual Policy Experiments: Evaluation of the ACA and its Variations.	141
2.18	Counterfactual Policy Experiments: the Impact of the ACA by Age Groups.	143

2.19 Counterfactual Policy Experiments: the Impact of the ACA by Health Status.	144
2.20 Counterfactual Policy Experiments: Partial Equilibrium.	149
2.21 Optimal Policy Parameters under Case 1	155
2.22 Optimal Policy Parameters under Case 2	156
2.23 Optimal Policy Parameters under Case 3: Allowance of Age-Dependent Subsidies and Individual Mandate	157
2.24 Counterfactual Policy Experiments: Optimal Design of HIX under case 3 and its Variations.	158
2.25 Comparison of the uninsured rate by different age groups under the ACA and under the optimal structure of HIX in Case 3.	161
2.26 Optimal Policy Parameters under Case 3 and Exogenous Offer Distribution	162
2.27 Counterfactual Policy Experiments: Comparison: Outcomes in general equilibrium under (1) optimal policies obtained with endogenous offer distribution of compensation packages and (2) optimal policies obtained with exogenous offer distribution of compensation packages.	163
3.1 Capitation Payment and Health Expenditure by Health Status in Los Angeles County	167
3.2 Capitation Payments and Demographic Characteristics	177
3.3 Plan Characteristics Included in Analysis	181
3.4 Summary Statistics at County Level	182
3.5 Incentives for Risk Selection	183
3.6 Relationship between Health Status and Over-payment by Location	185
3.7 Relationship between Advertising and Capitation Payments	187
3.8 Health Status and Insurer Choice by Medicare Beneficiaries	187
3.9 Estimates for Key Parameters in Utility	203
3.10 Elasticity of Demand with Respect to Advertising and Premiums	204

3.11	Estimates for Parameters in Mean Utility (δ_{jmt})	206
3.12	Estimates for Preference Heterogeneity	207
3.13	Estimates for Marginal Costs of Providing Insurance	208
3.14	Estimates for Marginal Costs of a Unit of Advertising	208
3.15	Ban on Advertising	210
3.16	Consumer's Surplus with a Ban on Advertising	211
3.17	Health Compositions in traditional Medicare vs MA (Ban on Advertising)	212
3.18	Risk Adjustment	214
3.19	Consumer's Surplus with Risk Adjustment	215
3.20	Health Risk Compositions in traditional Medicare vs MA (Risk Adjustment)	216
C.1	Logit Regression for Positive Medicare Claims Cost	250
C.2	Regression of Medicare Claims Costs on Health Characteristics	251

List of Figures

1.1	The Distributions of Workers' Accepted Wages by Health Insurance Status: Model vs. Data.	55
1.2	Size Distribution of Firms: Model vs. Data.	59
1.3	Size Distribution of Firms by Insurance Offering Status: Model vs. Data.	59
2.1	The fraction of employed workers in top 50% of firms (by productivity)	146
2.2	The fraction of employed individuals in top 50% of firms (by productivity): by individual type	147
2.3	Optimal subsidies rate over ages	157
2.4	Optimal subsidies rate over income	159
2.5	The fraction of employed workers in top 50% of firms (by productivity)	160
3.1	Relationship between Health Status and Over-payment by Location	185

Chapter 1

Equilibrium Labor Market Search and Health Insurance Reform

This chapter is co-authored with Hanming Fang.

1.1 Introduction

The Affordable Care Act (hereafter, ACA), signed into law by President Barack Obama in March 2010, represents the most significant reforms to the U.S. health insurance and health care market since the establishment of Medicare in 1965.¹ The health care reform in the U.S. was partly driven by two factors: first, a large fraction of the U.S. population does not have health insurance (close to 18% for 2009); second, the U.S. spends a much larger share of the national income on health care than the other OECD countries (health care accounts for about one sixth of the U.S. GDP in 2009).² There are many provisions in the ACA whose implementation will be phased in over several years, and some of the most significant changes will take effect from 2014. In particular, four of the most important

¹The Affordable Care Act refers to the Patient Protection and Affordable Care Act (PPACA) signed into law by President Obama on March 23, 2010, as well as the Amendment in the Health Care and Education Reconciliation Act of 2010.

²See OECD Health Data at www.oecd.org/health/healthdata for a comparison of the health care systems between the U.S. and the other OECD countries.

components of the ACA are as follows:³

- **(Individual Mandate)** All individuals must have health insurance that meets the law's minimum standards or face a penalty when filing taxes for the year, which will be 2.5 percent of income or \$695, whichever is higher.^{4, 5}

- **(Employer Mandate)** Employers with more than 50 full-time employees will be required to provide health insurance or pay a fine of \$2,000 per worker each year if they do not offer health insurance, where the fines would apply to the entire number of employees minus some allowances.

- **(Insurance Exchanges)** State-based health insurance exchanges will be established where the unemployed, the self-employed and workers who are not covered by employer-sponsored health insurance (ESHI) can purchase insurance. Importantly, the premiums for individuals who purchase their insurance from the insurance exchanges will be based on the average health expenditure risks of those in the exchange pool.⁶ Insurance companies that want to participate in an exchange need to meet a series of statutory requirements in order for their plans to be designated as "qualified health plans."

- **(Premium Subsidies)** All adults in households with income under 133% of Federal poverty line (FPL) will be eligible for receiving Medicaid coverage with no cost sharing.⁷

³Detailed formulas for the penalties associated with violating the individual and employer mandates, as well as for that for the premium subsidies, are provided in Section 1.8.2.

⁴These penalties would be implemented fully from 2016. In 2014, the penalty is 1 percent of income or \$95 and in 2015, it is 2 percent or \$325, whichever is higher. Cost-of-living adjustments will be made annually after 2016. If the least inexpensive policy available would cost more than 8 percent of one's monthly income, no penalties apply and hardship exemptions will be permitted for those who cannot afford the cost.

⁵This component of the ACA was one of the core issues in the U.S. Supreme Court case *National Federation of Independent Business v. Sebelius*, 567 U.S. 512 (2012) where twenty-six States, several individuals and the National Federation of Independent Business challenged the constitutionality of the individual mandate and the Medicaid expansion. The U.S. Supreme Court ruled on June 28, 2012 to uphold the constitutionality of the individual mandate on a 5-to-4 decision.

⁶States that opt not to establish their own exchanges will be pooled in a federal health insurance exchange.

⁷This represents a significant expansion of the current Medicaid system because many States currently cover adults with children only if their income is considerably lower, and do not cover childless adults at all. The U.S. Supreme Court ruled on June 28, 2012 that the law's provision that, if a State does not comply with the ACA's new coverage requirements, it may lose not only the federal funding for those requirements, but all of its federal Medicaid funds, is unconstitutional. This ruling allows states to opt out of ACA's Medicaid expansion, leaving each state's decision to participate in the hands of the nation's

For individuals and families whose income is between the 133 percent and 400 percent of the FPL, subsidies will be provided toward the purchase of health insurance from the exchanges.

The goal of this study is to understand how the health care reform will change the health insurance and labor markets. Would the ACA significantly reduce the uninsured rate? Would more employers be offering health insurance to their employees? How would the reform affect workers' wages, health and productivity? How would it affect employment and firm size distributions? What is the impact on total health expenditures and on government budget? We are also interested in several counterfactual policies. For example, how would the remainder of the ACA perform if its individual mandate component had been struck down by the Supreme Court? What would happen if the current tax exemption status of employer-provided insurance premium is eliminated? Are the premium subsidies necessary for the insurance exchanges to overcome the adverse selection problem? Can we identify alternative reforms that can improve welfare relative to the ACA?

An equilibrium model that integrates the labor and health insurance markets is necessary for us to understand the general equilibrium implications of the health insurance reform. First, the United States is unique among industrialized nations in that it lacks a national health insurance system and most of the working-age population obtain health insurance coverage through their employers. According to Kaiser Family Foundation and Health Research and Educational Trust (2009), more than 60 percent of the non-elderly population received their health insurance sponsored by their employers, and about 10 percent of workers' total compensation was in the form of ESHI premiums.⁸ Second, there have been many well-documented connections between firm sizes, wages, health insurance offerings and worker turnovers. For example, it is well known that firms that do not offer

⁸Among those with private coverage from any source, about 95% obtained employment-related health insurance (see Selden and Gray (2006)).

health insurance are more likely to be small firms, to offer low wages, and to experience higher rate of worker turnover. In the 1997 Robert Wood Johnson Foundation Employer Health Insurance Survey which we use later in our analysis, we find that the average size was about 8.8 for employers that did not offer health insurance, in contrast to an average size of 33.9 for employers that offered health insurance; the average annual wage was \$20,560 for workers at firms that did not offer health insurance, in contrast to average wage of \$29,077 at firms that did; also, annual separation rate of workers at firms not offering health insurance was 17.3%, while it was 15.8% at firms that did. Moreover, in our data sets, workers in firms that offer health insurance are more likely to self report better health than those in firms that do not offer health insurance.

In this study we present and empirically implement an equilibrium labor market search model integrated with health insurance market. Our model is based on Burdett and Mortensen (1998) and Bontemps, Robin, and Van den Berg (1999, 2000).⁹ But we depart from these standard models by incorporating health and health insurance; thus we endogenize the distributions of wages and health insurance provisions, employer size, employment and worker's health. In our model workers observe their own health status which evolves stochastically. Workers' health status affects both their medical expenditures and their labor productivity. Health insurance eliminates workers' out-of-pocket medical expenditure risks and affects the dynamics of their health status. In the benchmark model, we assume that workers can obtain health insurance only through employers. Both unemployed and employed workers randomly meet firms and decide whether to accept their job offer, compensation package of which consists of wage and ESHI. Firms, which are heterogenous in their productivity, post compensation packages to attract workers. The cost of providing health insurance, which will be used to determine ESHI premiums, is

⁹Their model theoretically explains both wage dispersion among *ex ante* homogeneous workers and the positive correlation between firm size and wage. Moscarini and Postel-Vinay (2012) demonstrate that the extended version of this model, which allows firm productivity heterogeneity and aggregate uncertainty, has very interesting but also empirically relevant properties about firm size and wage adjustment over the business cycles.

determined by both the health composition of its workforce and a fixed administrative cost. When deciding the compensation packages, the firms anticipate that their choice of compensation packages will affect the health composition of their worker as well as their sizes in the steady state.

We characterize the steady state equilibrium of the model in the spirit of Burdett and Mortensen (1998). We estimate the parameters of the baseline model using data from Survey of Income and Program Participation (SIPP, 1996 Panel), Medical Expenditure Panel Survey (MEPS, 1997-1999), and Robert Wood Johnson Foundation Employer Health Insurance Survey (RWJ-EHI, 1997). The first two data sets are panels on worker-side labor market status, health and health insurance, while the third one is a cross-sectional establishment level data set which contains information such as establishment size and health insurance coverage. Because the data on the supply-side (i.e., workers) and demand-side (i.e. firms) of labor markets come from different sources, we estimate the model using GMM for the case of combinations of data sets, as proposed by Imbens and Lancaster (1994) and Petrin (2002). We show that our baseline model delivers a rich set of predictions that can *qualitatively* and *quantitatively* account for a wide variety of the aforementioned phenomenon observed in the data including the correlations among firm sizes, wages, health insurance offering rates, turnover rates and workers' health compositions.

In our empirical analysis, we find that a critical driver to explain these correlations is the positive effect of health insurance on the dynamics of health status. While it is true that firms by offering health insurance can benefit from the tax exemption of the insurance premium, they also attract unhealthy workers who both increase their health insurance costs and decrease their labor productivity – this is the standard *adverse selection* problem. This creates a potential disincentive for firms to offer health insurance. In Section 1.4.1, we show that in the presence of the positive effect of health insurance on health, the degree of the adverse selection problem faced by high-productivity firms offering health insurance is *less severe* than that for low-productivity firms. The reason

is that, a high-productivity firm offering health insurance can poach workers from a much wider range of firms, including a larger fraction of workers who worked in firms that already offer insurance and thus are healthier; in contrast, a low-productivity firm offering health insurance can only poach workers from firms with even lower productivity, most of which do not offer health insurance and thus have less healthy workers.

Moreover, the adverse selection problem that firms offering health insurance suffer is attenuated over time by the *positive effect of health insurance on health*. Importantly, however, this effect from the improvement of health status of the workforce is captured more by high productivity firms due to what we term as “*retention effect*,” which simply refers to the fact that high-productivity firms tend to offer higher wages and retain workers longer (see Fang and Gavazza (2011) for an evidence for this mechanism). These effects jointly allow our model to generate a positive correlation between wage, health insurance, and firm size; and they moreover explain why health status of employees covered by ESHI is better than that of uninsured employees in the data.¹⁰

We use our estimated model to examine the impact of the previously-mentioned four key components of the ACA. We find that the implementation of the ACA would significantly reduce the uninsured rate among the workers in our estimation sample from 20.12% in the benchmark economy to 7.27%. This large reduction of the uninsured rate is mainly driven by the unemployed (3.2% of the population) and 9.65% of the employed workers with relatively low wages participating in the insurance exchange with their premium supported by the income-based subsidies. We find that the employer mandate increases the health insurance offering rate for firms with 50 or more workers from 91.13% in the benchmark to 99.93% under the ACA; however, the health insurance offering rate is somewhat reduced for firms with less than 50 workers. Moreover, the employer mandate leads to a slight increase in the fraction of firms with less than 50 workers, with a small but

¹⁰In fact, we will show in Table 1.15 that, due to these effects, the incentives for firms, even the more productive ones, to offer health insurance is largely unaffected in a counterfactual environment where the tax exemption of ESHI premiums is eliminated.

noticeable clustering of firms with size just below the employer mandate threshold of 50. Overall, there is only a small increase in the fraction of employed workers receiving ESHI, from 82.53% in the benchmark to 82.84% under the ACA. We also find that the ACA would raise the health expenditure by about 8%, and lead to an increase in the fraction of healthy workers in the population.

We also investigate the effect of the ACA if its individual mandate component were removed, a scenario that would have resulted had the Supreme Court ruled the individual mandate unconstitutional (see Footnote 5). We find that a significant reduction in the uninsured rate would also have been achieved: the uninsured rate in our simulation under “ACA without individual mandate” would be 12.18%, significantly lower than the 20.12% under the benchmark. The premium subsidy component of the ACA would have in itself drawn all the unemployed (healthy or unhealthy) and the low-wage employed (again both healthy and unhealthy) in the insurance exchange. In fact, if we were to remove the premium subsidies, instead of the individual mandate, from the ACA, we find that the insurance exchange will suffer from adverse selection problem so severe as to render it non-active at all. However, the presence of the exchange, though non-active, still changes the workers’ outside option and thus affects the firms’ decisions, which leads to a small reduction of the uninsured rate from 20.12% in the benchmark to 17.14-17.28% under “ACA without the subsidies.”

Interestingly, we find that, under a policy of “ACA without the employer mandate,” the uninsured rate would be 6.44%, lower than the uninsured rate under the full ACA. Without the employer mandate, the health composition of the workers in the health exchange pool is improved, which leads to a decrease in the premium in the exchange. This makes it less desirable for individuals to stay uninsured and subject to individual mandate penalty. We also find that the equilibrium under the policy of “ACA without employer mandate” achieves a higher average productivity, higher average wages and higher worker’s average utility, without increasing the government spending.

We also simulate the effects of eliminating the tax exemption for ESHI premium both under the benchmark and under the ACA. We find that, while the elimination of the tax exemption for ESHI premium would reduce the probability of all firms, especially the larger ones, offering health insurance to their workers, the overall effect on the uninsured rate is rather modest. We find that the uninsured rate would increase from 20.12% to 23.39% when the ESHI tax exemption is removed in the benchmark economy; and it will increase from 7.27% to 9.15% under the ACA. We also experimented with the effect of prohibiting firms from offering ESHI. We find that it would lead to a large reduction in the fraction of active firms in the labor market, which suggests that ESHI allows low-productive firms to be active in the market because they can potentially extract the workers' risk premium.

The remainder of this chapter is structured as follows. In Section 1.2, we review the related literature; in Section 1.3, we present the model of the labor market with endogenous determinations of wages and health insurance provisions; in Section 1.4, we present a qualitative assessment of the workings of the model; in Section 1.5, we describe the data sets used in our empirical analysis; in Section 1.6, we explain our estimation strategy; in Section 1.7, we present our estimation results and the goodness-of-fit; in Section 1.8, we describe the results from several counterfactual experiments; and finally in Section 1.9, we conclude and discuss directions for future research.

1.2 Related Literature

This study is related to three strands of the literature. First and foremost, it is related to a small literature that examines the relationship between health insurance and labor market. Dey and Flinn (2005) propose and estimate an equilibrium model of the labor market in which firms and workers bargain over both wages and health insurance offerings to examine the question of whether the employer-provided health insurance system leads to inefficiencies in workers' mobility decisions (which are often referred to as "job lock")

or “job push” effects). However, because a worker/vacancy match is the unit of analysis in Dey and Flinn (2005), their model is not designed to address the relationship between firm size and wage/health insurance provisions, which is important to understand the size-dependent employer mandate in the ACA. Moreover, in Dey and Flinn (2005), workers’ health status and health expenditures are not explicitly modeled, and firms’ heterogeneous costs of offering health insurance are also exogenous. In our framework, we explicitly incorporate workers’ health and health expenditures, and endogenize health insurance costs and premium. We believe these features are essential to assess the general equilibrium effects of the ACA on population health, health expenditures and health insurance premiums.

Bruegemann and Manovskii (2010) develop a search and matching model to study firms’ health insurance coverage decision. In their model, firm sizes are discrete to highlight the effect of fluctuations in the health composition of employees on the dynamics of firm’s coverage decision, and they argue that the insurance market for small firms suffers from adverse selection problem because those firms try to purchase health insurance when most of their employees are unhealthy. Our study provides a complementary channel which has received little attention in the literature: it is harder for small firms to overcome adverse selection problems because they cannot retain their workers long enough to capture the benefits from the advantageous dynamic effects of health insurance on health. This channel arises in our environment because we allow for on-the-job searches and explicitly model the dynamic effect of health insurance on health, both of which are absent in their model. Moreover, our model endogenously generates reasonable wage distributions, which are important to study the impact of income-based premium subsidies and individual mandates.

The channel that worker turnover discourages firm’s health insurance provision is related to Fang and Gavazza (2011). They argue that health is a form of general human capital, and labor turnover and labor-market frictions prevent an employer-employee pair

from capturing the entire surplus from investment in an employee’s health, generating under-investment in health during working years and increasing medical expenditures during retirement. We advance their insights by showing that in an equilibrium model of labor market, it also reduces the adverse selection problem for high-productivity firms relative to low-productivity firms, and helps explain why high-productivity firms have a stronger incentives to provide health insurance to their workers. Moreover, our primary focus is about health insurance coverage provision and labor market outcomes, while theirs is about the life-cycle medical expenditure.

Second, there are a growing number of empirical analyses examining the likely impact of the ACA by focusing the Massachusetts Health Reform, implemented in 2006, which has similar features with the ACA. Kolstad and Kowalski (2012b,a,c) study the effect on medical expenditure, selection in insurance markets, and labor markets. Courtemanche and Zapata (2012) found that Massachusetts reform improves the health status of individuals. They study these issues based on a “difference-in-difference” approach. These approaches are very informative to understand the overall and likely impact of reform. By structurally estimating an equilibrium model, we complement this literature by providing a quantitative assessment of the *mechanisms* generating such outcomes. Moreover, we provide the assessment of various other counterfactual policies such as the removal of tax exclusion of ESHI premiums. In a recent paper, Pashchenko and Porapakkarm (2013) evaluates the ACA using a calibrated life-cycle incomplete market general equilibrium model. They consider several individual decisions such as health insurance, consumption, saving, and labor supply, but they do not model firms’ decision of offering health insurance as well as firm size distribution. Therefore, their model is not designed to address the effects of ACA on firms’ insurance coverage and wage offer decisions and the equilibrium effects of size-dependent employer mandate.

Third, this study is related to a large literature estimating equilibrium labor market

search models.¹¹ Van den Berg and Ridder (1998) and Bontemps, Robin, and Van den Berg (1999, 2000) empirically implement Burdett and Mortensen (1998)’s model. Their empirical frameworks have been widely applied in subsequent studies investigating the impact of various labor market policies on labor market outcomes. Among this literature, our study is mostly related to Shephard (2012) and Meghir, Narita, and Robin (2012), which also allow for multi-dimensional job characteristics as in our study: wage and part-time/full-time in Shephard (2012), wage and formal/informal sector in Meghir, Narita, and Robin (2012), and wage and health insurance offering in our study. However, in Shephard (2012) a firm’s job characteristics is assumed to be exogenous, while in our study employers endogenously choose job characteristics. In Meghir, Narita, and Robin (2012) firms choose whether to enter the formal or informal sectors so in some sense their job characteristics are also endogenously determined; however, in Meghir, Narita, and Robin (2012), workers are homogeneous so firms’ decision about which sector to enter does not affect the composition of the types of workers they would attract. In contrast, in our model, workers are heterogenous in their health, thus employers endogenously choose job characteristics, namely wage and health insurance offering, by taking into account their influence on the initial composition of its workforce as well as the subsequent worker turnover.

1.3 An Equilibrium of Model of Wage Determination and Health Insurance Provision

1.3.1 The Environment

Consider a labor market with a continuum of firms with measure normalized to 1 and a continuum of workers with measure $M > 0$.¹² They are randomly matched in a frictional

¹¹See Eckstein and Wolpin (1990) for a seminal study that initiated the literature.

¹²Throughout the study, we use “workers” and “firms” interchangeably with “individuals” and “employers” respectively.

labor market. Time is discrete, and indexed by $t = 0, 1, \dots$, and we use $\beta \in (0, 1)$ to denote the discount factor for the workers.¹³

Workers have constant absolute risk aversion (CARA) preferences:¹⁴

$$u(c) = -\exp(-\gamma c), \quad (1.1)$$

where $\gamma > 0$ is the absolute risk aversion parameter.

Workers' Health. Workers differ in their health status, denoted by h , and they can either be **Healthy** (H) or **Unhealthy** (U). In our model, a worker's health status has two effects. First, it affects the distribution of health expenditures. Specifically, we model an individual's health expenditure distributions as follows. Let $x \in \{0, 1\}$ denote an individual's health insurance status, where $x = 1$ means that he has health insurance. We assume that the probability that an individual will experience a medical shock is given by:

$$\Pr(m > 0 | h, x) = \Phi(\alpha_0 + \beta_0 1\{h = U\} + \gamma_0 x), \quad (1.2)$$

and conditional on a medical shock, the realization of the medical expenditure is drawn from a log normal distribution:

$$m | (h, x) \sim \exp(\alpha_m + \beta_m 1\{h = U\} + \gamma_m x + \epsilon_{hx}), \quad (1.3)$$

where $\epsilon_{hx} \sim N(0, \sigma_{hx}^2)$ and is independently and identically distributed across time periods. Note that in (1.2) and (1.3) we allow both the individual's health and health insurance status to affect the medical expenditure distributions; moreover, in (1.3) we allow that the log normal medical expenditure distributions to be conditionally heteroskedas-

¹³In our empirical analysis, a "period" corresponds to four months.

¹⁴Alternatively we can assume constant relative risk aversion (CRRA) preferences as in Rust and Phelan (1997), but then would have to deal with the issue of possible negative consumption.

tic.¹⁵ In subsequent analysis, we will use \tilde{m}_h^x to denote the random medical expenditure for individuals with health status h and health insurance status x as described by (1.2) and (1.3), and use m_h^x to denote the expectation of \tilde{m}_h^x which is given by:¹⁶

$$m_h^x \equiv E\tilde{m}_h^x = \exp(\alpha_m + \beta_m 1\{h = U\} + \gamma_m x) \exp\left(\frac{\sigma_{hx}^2}{2}\right) \Phi(\alpha_0 + \beta_0 1\{h = U\} + \gamma_0 x). \quad (1.4)$$

Second, a worker's health status affects his productivity. Specifically, if an individual works for a firm with productivity p , he can produce p units of output if he is healthy, but he can produce only $d \times p$ units of output if he is unhealthy where $1 - d$ represents the productivity loss from being unhealthy.¹⁷

In each period, worker's health status changes stochastically according to a Markov Process. The period-to-period transition of an individual's health status depends on his health insurance status. We use $\pi_{h'h}^x \in (0, 1)$ to denote the probability that a worker's health status changes from $h \in \{H, U\}$ to $h' \in \{H, U\}$ conditional on insurance status $x \in \{0, 1\}$. The transition matrix is thus, for $x \in \{0, 1\}$,

$$\pi^x = \begin{pmatrix} \pi_{HH}^x & \pi_{UH}^x \\ \pi_{HU}^x & \pi_{UU}^x \end{pmatrix}, \quad (1.5)$$

where $\pi_{UH}^x = 1 - \pi_{HH}^x$ and $\pi_{HU}^x = 1 - \pi_{UU}^x$.

¹⁵In our empirical application, we truncate it at the top at twice the highest medical expenditure observed in the data. Our specification allows us to capture two of the most salient features of the medical expenditure distributions: they are heavily skewed to the right and there is a sizable fraction of individuals with zero medical expenditure. Similar specifications have been used in Einav, Finkelstein, Ryan, Schrimpf, and Cullen (2013).

¹⁶The conditional variances of the medical expenditures can also be analytically characterized:

$$\begin{aligned} & \text{Var}(\tilde{m}_h^x | h, x) \\ &= \exp[2(\alpha_m + \beta_m 1\{h = U\} + \gamma_m x)] \exp(\sigma_{hx}^2) [\exp(\sigma_{hx}^2) - 1] \Phi(\alpha_0 + \beta_0 1\{h = U\} + \gamma_0 x) \end{aligned}$$

We use these moments in our first-step estimation described in Section 1.6.1.

¹⁷One can alternatively assume that the productivity loss only occurs if an individual experiences a bad health shock. Because an unhealthy worker is more likely to experience a bad health shock, such a formulation is equivalent to the one we adopt in the study.

Firms. Firms are heterogeneous in their productivity. In the population of firms, the distribution of productivity is denoted by $\Gamma(\cdot)$ which we assume to admit an everywhere continuous and positive density function. In our empirical application, we specify Γ to be lognormal with mean μ_p and variance σ_p^2 , i.e., $p \sim \ln N(\mu_p, \sigma_p^2)$.

Firms, after observing their productivity, decide a package of wage and health insurance provision, denoted by (w, x) where $w \in R_+$ and $x \in \{0, 1\}$. If a firm offers health insurance to its workers, it has to incur a fixed administrative cost $C > 0$. We assume that any firm that offers health insurance to its workers is self-insured, and will charge an insurance premium from its workers each period to cover the necessary reimbursement of all the realized health expenditures in addition to the administrative cost C .¹⁸

Importantly, we assume, due to regulations in Health Insurance Portability and Accountability Act (HIPAA) which prohibits discrimination against employees and dependents based on their health status, that all the workers in a given firm will receive the *same* compensation package (wage and health insurance offering *regardless of* their health status).¹⁹

Health Insurance Market. In the baseline model, which is intended to represent the pre-ACA U.S. health insurance market, we assume that workers can obtain health insurance only if their employers offer them. This is a simplifying assumption meant to capture the fact that the individual private insurance market is rather small in the U.S. In our counterfactual experiment, we will consider the case of competitive private insurance market to mimic the health insurance exchanges that would be established under the ACA.

¹⁸In principle, firms should also be able to decide on the premium if they decide to offer health insurance. However, because we require that firms be self-insured, the insurance premium will be determined in equilibrium by the health composition of workers in steady state.

¹⁹HIPAA is an amendment of Employee Retirement Security Act (ERISA), which is a federal law that regulates issues related to employee benefits in order to qualify for tax advantages. A description of HIPAA can be found at the Department of Labor website: <http://www.dol.gov/dol/topic/health-plans/portability.htm>

Labor Market. Firms and workers are randomly matched in the labor market. In each period, an unemployed worker randomly meets a firm with probability $\lambda_u \in (0, 1)$. He then decides whether to accept the offer, or to remain unemployed and search for jobs in next period. We assume that all new-born workers are unemployed.

If an individual is employed, he meets randomly with another firm with probability $\lambda_e \in (0, 1)$. If a currently employed worker receives an offer from another firm, he needs to decide whether to accept the outside offer or to stay with the current firm. An employed worker can also decide to return to the unemployment pool.²⁰ Moreover, each match is destroyed exogenously with probability $\delta \in (0, 1)$, upon which the worker will return to unemployment. As we discuss in Section 1.3.2, we assume that individual may experience both the exogenous job destruction and the arrival of the new job offer within in the same period.²¹

To generate a steady state for the labor market, we assume that in each period any individual, regardless of health and employment status, will leave the labor market with probability $\rho \in (0, 1)$. An equal measure of newborns will enter the labor market unemployed and their initial health status will be healthy with probability $\mu_H \in (0, 1)$.

Income Taxes and Unemployment Benefit. Workers' wages are subject to a non-linear tax schedule, but the ESHI premium is tax exempt in the baseline model. For the *after-tax income* $T(y)$, we follow the specification in Kaplan (2012) which approximates the U.S. tax code by:²²

$$T(y) = \tau_0 + \tau_1 \frac{y^{(1+\tau_2)}}{1 + \tau_2} \tag{1.6}$$

²⁰Returning to unemployment may be a better option for a currently employed worker if his health status changed from when he accepted the current job offer, for example.

²¹This specification is used by Wolpin (1992) and more recently by Jolivet, Postel-Vinay, and Robin (2006). This allows us to account for transitions known as “job to unemployment, back to job” all occurring in a single period, as we observe in the data.

²²Robin and Roux (2002) also studied the impact of progressive income tax within the framework of Burdett and Mortensen (1998).

where $\tau_0 > 0$, $\tau_1 > 0$ and $\tau_2 < 0$.

1.3.2 Timing in a Period

At the beginning of each period, we should imagine that individuals, who are heterogeneous in their health status, are either unemployed or working for firms offering different combinations of wage and health insurance packages. We now describe the explicit timing assumptions in a period that we use in the derivation of the value functions in Section 1.3.3. We believe that our particular timing assumptions simplify our derivation but they are not crucial.

1. Any individual, whether employed or unemployed, and regardless of his health status, may leave the labor market with probability $\rho \in (0, 1)$;
2. If an employed worker stays in the labor market matched with a firm with productivity p , then:
 - (a) he produces output p if healthy and $d \times p$ if unhealthy;
 - (b) the firm pays wage and collects insurance premium if it offers health insurance;
 - (c) he receives a medical expenditure shock, the distribution of which depends on his beginning-of-the-period health status;
 - (d) he then observes the realization of the health status that will be applicable next period;
 - (e) he randomly meets with new employers with probability λ_e ;
 - (f) the current match is destroyed with probability $\delta \in (0, 1)$, in which case the worker must decide whether to accept the outside offer, if any, or to enter unemployment pool;
 - (g) if the current match is not destroyed, then he decides whether to accept the outside offer if any, to stay with the current firm, or to quit into unemployment.

3. Any unemployed worker experiences the following in a period:
 - (a) he receives the unemployment benefit b ;
 - (b) he receives a medical expenditure shock, the distribution of which depends on his beginning-of-the-period health status;
 - (c) he then observes the realization of the health status that will be applicable next period;
 - (d) he randomly meets with employers with probability λ_u , and decides whether to accept the offer if any, or to stay unemployed.

4. Time moves to the next period.

1.3.3 Analysis of the Model

In this section, we characterize the steady state equilibrium of the model. The analysis here is similar to but generalizes that in Burdett and Mortensen (1998). We first consider the decision problem faced by a worker, for a postulated distribution of wage and insurance packages by the firms, denoted by $F(w, x)$, and derive the steady state distribution of workers of different health status in unemployment and among firms with different offers of wage and health insurance packages (w, x) . We then solve the firms' optimization problem and provide the conditions for the postulated $F(w, x)$ to be consistent with equilibrium.

Value Functions

We first introduce the notation for several valuation functions. We use $v_h(y, x)$ to denote the expected *flow* utility of workers with health status h from income y and insurance status $x \in \{0, 1\}$; and it is give by:

$$v_h(y, x) = \begin{cases} u(T(y)) & \text{if } x = 1 \\ E_{\tilde{m}_h^0} u(T(y) - \tilde{m}_h^0) & \text{if } x = 0, \end{cases} \quad (1.7)$$

where $T(y)$ is after-tax income as specified in (1.6) and \tilde{m}_h^0 is the random medical expenditure for uninsured individual as specified by (1.2) and (1.3). Note that when an individual is insured, i.e., $x = 1$, his medical expenditures are fully covered by the insurance. As long as \tilde{m}_h^0 is not always 0, $v_h(y, 1) > v_h(y, 0)$; i.e., regardless of workers' health, if wages are fixed, then all workers desire health insurance.

Let U_h denote the value for an unemployed worker with health status h at the beginning of a period; and let $V_h(w, x)$ denote the value function for an employed worker with health status h working for a job characterized by wage-insurance package (w, x) at the beginning of a period. U_h and $V_h(\cdot, \cdot)$ are of course related recursively. U_h is given by:

$$\frac{U_h}{1 - \rho} = v_h(b, 0) + \beta E_{h'|h,0} \left[\begin{array}{c} \lambda_u \int \max\{V_{h'}(w, x), U_{h'}\} dF(w, x) \\ +(1 - \lambda_u)U_{h'} \end{array} \right], \quad (1.8)$$

where the expectation $E_{h'}$ is taken with respect to the distribution of h' conditional on the current health status h and insurance status $x = 0$ because unemployed workers are uninsured in the baseline model. (1.8) states that the value of being unemployed, normalized by the survival rate $1 - \rho$, consists of the flow payoff $v_h(b, 0)$, and the discounted expected continuation value where the expectation is taken with respect to the health status h' next period, whose transition is given by $\pi_{h'h}^0$ as described in (1.5). The unemployed worker may be matched with a firm with probability λ_u and the firm's offer (w, x) is drawn from the distribution $F(w, x)$. If an offer is received, the worker will choose whether to accept the offer by comparing the value of being employed at that firm $V_{h'}(w, x)$, and the value of remaining unemployed $U_{h'}$; if no offer is received, which occurs with probability $1 - \lambda_u$, the worker's continuation value is $U_{h'}$.

Similarly, $V_h(w, x)$ is given by

$$\begin{aligned} \frac{V_h(w, x)}{1 - \rho} &= v_h(w, x) \\ &+ \beta \lambda_e \left\{ \begin{aligned} &(1 - \delta) \mathbf{E}_{h'| (h, x)} \left[\int \max \{ V_{h'}(\tilde{w}, \tilde{x}), V_{h'}(w, x), U_{h'} \} dF(\tilde{w}, \tilde{x}) \right] \\ &+ \delta \mathbf{E}_{h'| (h, x)} \left[\int \max \{ U_{h'}, V_{h'}(\tilde{w}, \tilde{x}) \} dF(\tilde{w}, \tilde{x}) \right] \end{aligned} \right\} \\ &+ \beta(1 - \lambda_e) \left\{ (1 - \delta) \mathbf{E}_{h'| (h, x)} [\max \{ U_{h'}, V_{h'}(w, x) \}] + \delta \mathbf{E}_{h'| (h, x)} [U_{h'}] \right\}. \end{aligned} \quad (1.9)$$

Note that in both (1.8) and (1.9), we used our timing assumption that a worker's health status next period depends on his insurance status this period even if he is separated from his job at the end of this period (see Section 1.3.2).

Workers' Optimal Strategies

Standard arguments can be used to show that a worker's decision about whether to accept a job offer is characterized by "generalized reservation wage" policies. Note that in our model, both unemployed and employed workers make decisions about whether to accept or reject an offer, and their reservation wages will depend on their state variables, i.e., their employment status including the terms of their current offer (w, x) if they are employed, and their health status h .

Optimal Strategies for Unemployed Workers. First, consider an unemployed worker. As the right hand side of (1.7) is increasing in w , $V_h(w, x)$ is increasing in w . On the other hand, U_h is independent of w . Therefore, the reservation wage for an unemployed worker with health status h satisfies:

$$U_h = V_h(\underline{w}_h^x, x), \quad (1.10)$$

so that if an unemployed worker meets a firm with offer (w, x) , he will accept the offer if $w > \underline{w}_h^x$ and reject otherwise. Because a worker's expected flow utility $v_h(w, x)$ as described in (1.7) and the law of motion for health as described in (1.5) both depend on

his current health and health insurance status, the reservation wages of the unemployed also differ across these statuses.

Optimal Strategies for Currently-Employed Workers: Job-to-Job Transitions.

Now we consider the reservation wages for a currently-employed worker. Let (w, x) be the wage-insurance package offered by his current employer; and let (w', x') be the one offered by his potential employer. Then, the reservation wage for the employed worker with health status h to switch, denoted by $\underline{s}_h^{x'}(w, x)$, must satisfy

$$V_h(w, x) = V_h\left(\underline{s}_h^{x'}(w, x), x'\right). \quad (1.11)$$

A worker with health status h on a current job (w, x) will switch to a job (w', x') if and only if $w' > \underline{s}_h^{x'}(w, x)$. It is straightforward from (1.9) that

$$\underline{s}_h^{x'}(w, x) = w \text{ if } x = x'.$$

However, when $x \neq x'$, the exact value of $\underline{s}_h^{x'}(w, x)$ must be solved from (1.11); in particular, it will differ by worker's health and health insurance status. It can be easily shown that

$$\underline{s}_h^{x'}(w, x) > w \text{ if } x = 1 \text{ and } x' = 0; \quad \underline{s}_h^{x'}(w, x) < w \text{ if } x = 0 \text{ and } x' = 1.$$

Once we solve $\underline{s}_h^x(\cdot, \cdot)$, we can use its definition as in (1.11) to obtain, for any new offer (w', x') ,

$$V_h(w', x') = V_h(\underline{s}_h^x(w', x'), x),$$

thus a worker with a current offer (w, x) will accept the new offer (w', x') if and only if

$$w < \underline{s}_h^x(w', x'). \quad (1.12)$$

We will use this characterization in the expressions for steady state conditions in Section 1.3.3.

Optimal Strategies for Currently-Employed Workers: Quitting to Unemployment. Finally, a worker with health status h who is currently on a job (w, x) may choose to quit into unemployment. This may happen because of the changes in workers' health condition since he last accepted the current job offer and the possibility that the offer arrival probability for unemployed worker, λ_u , may be higher than that for an employed worker, λ_e . Clearly a worker with health status h and health insurance status x will quit into unemployment only if the current wage w is below a threshold. Let us denote the threshold wages for quitting into unemployment by \underline{q}_h^x . Clearly, \underline{q}_h^x must satisfy

$$V_h(\underline{q}_h^x, x) = U_h. \quad (1.13)$$

Comparing (1.13) with (1.10), it is clear that $\underline{q}_h^x = \underline{w}_h^x$. Thus we can conclude that employed workers will quit to unemployment only if his health status changed from when he first started on the current job. Moreover, if $\underline{w}_H^x < \underline{w}_U^x$, then a currently unhealthy worker who accepted a job (w, x) with wage $w \in (\underline{w}_H^x, \underline{w}_U^x)$ when his health status was H may now quit into unemployment; if $\underline{w}_H^x > \underline{w}_U^x$ instead, then a currently healthy worker who accepted a job (w, x) with wage $w \in (\underline{w}_U^x, \underline{w}_H^x)$ when his health status was U may now quit into unemployment.

Steady State Condition

We will focus on the steady state of the dynamic equilibrium of the labor market described above. We first describe the steady state equilibrium objects that we need to characterize and then provide the steady state conditions.

In the steady state, we need to describe how the workers of different health status h are allocated in their employment (w, x) . Let u_h denote the measure of unemployed workers

with health status $h \in \{U, H\}$; and let e_h^x denote the measure of employed workers with health insurance status $x \in \{0, 1\}$ and health status is $h \in \{U, H\}$. Of course, we have

$$\sum_{h \in \{U, H\}} (u_h + e_h^0 + e_h^1) = M. \quad (1.14)$$

Let $G_h^x(w)$ the *fraction* of employed workers with health status h working on jobs with insurance status x whose wage is below w , and let $g_h^x(w)$ be the corresponding density of $G_h^x(w)$. Thus, $e_h^x g_h^x(w)$ is the density of employed workers with health status h whose compensation package is (w, x) .

These objects would have to satisfy the steady state conditions for unemployment and for the allocations of workers across firms with different productivity. First, let us consider the steady state condition for unemployment. The inflow into unemployment with health status h is given by

$$\begin{aligned} [u_h]^+ &\equiv (1 - \rho) [\delta(1 - \lambda_e) + \delta\lambda_e(F(\underline{w}_h^1, 1) + F(\underline{w}_h^0, 0))] \\ &\quad \times [e_h^0\pi_{hh}^0 + e_h^1\pi_{hh}^1 + e_{h'}^0\pi_{hh'}^0 + e_{h'}^1\pi_{hh'}^1] \end{aligned} \quad (1.15a)$$

$$+ (1 - \rho) u_{h'} \pi_{hh'}^0 [1 - \lambda_u(1 - F(\underline{w}_h^1, 1) - F(\underline{w}_h^0, 0))] \quad (1.15b)$$

$$\begin{aligned} &+ (1 - \rho)(1 - \delta) e_{h'}^1 \pi_{hh'}^1 G_{h'}^1(\underline{w}_h^1) \\ &\quad \times [1 - \lambda_e(1 - F(\underline{w}_h^1, 1) - F(\underline{w}_h^0, 0))] \end{aligned} \quad (1.15c)$$

$$\begin{aligned} &+ (1 - \rho)(1 - \delta) e_{h'}^0 \pi_{hh'}^0 G_{h'}^0(\underline{w}_h^0) \\ &\quad \times [1 - \lambda_e(1 - F(\underline{w}_h^1, 1) - F(\underline{w}_h^0, 0))] \end{aligned} \quad (1.15d)$$

$$+ M\rho\mu_h. \quad (1.15e)$$

In the above expression, the term on line (1.15a) is the measure of employed workers who had health status h this period, did not leave the labor market but had their jobs terminated exogenously, and did not subsequently find a job that was better than being unemployed. The term on line (1.15b) is the measure of workers whose health status was

h' last period but transitioned to h this period and who did not leave for employment. The terms on lines (1.15c) and (1.15d) are the measures of workers currently working on jobs with and without health insurance, respectively, quitting into unemployment. To understand these expressions, consider the term on line (1.15c). First, quitting into unemployment only applies to workers who did not leave the labor market and whose job did not get terminated (i.e., $(1 - \rho)(1 - \delta)$ measure of them). Second, note that quitting into unemployment at health status h this period is possible only if the worker's health status was h' in the previous period then transitioned to h this period, because otherwise the worker would have quit already in the previous period, and moreover only if his wage was lower than \underline{w}_h^1 defined in (1.10); these are captured by the term $e_{h'}^1 \pi_{hh'}^1 G_{h'}^1(\underline{w}_h^1)$. Third, those who quit into unemployment will remain in the unemployment pool only if the offer they may have received is not acceptable, which happens with probability $[1 - \lambda_e(1 - F(\underline{w}_h^1, 1) - F(\underline{w}_h^0, 0))]$. Finally, the term on line (1.15e) is the measure of new workers born into health status h .

The outflow from unemployment with health status h is given by:

$$[u_h]^- \equiv u_h \left\{ \rho + (1 - \rho) \left[\pi_{h'h}^0 + \pi_{hh}^0 \lambda_u (1 - F(\underline{w}_h^1, 1) - F(\underline{w}_h^0, 0)) \right] \right\}. \quad (1.16)$$

It states that a ρ fraction of the unemployed with health status h will die and the remainder $(1 - \rho)$ will either change to health status h' (with probability $\pi_{h'h}^0$), or if their health does not change (with probability π_{hh}^0) they may become employed with probability $\lambda_u(1 - F(\underline{w}_h^1, 1) - F(\underline{w}_h^0, 0))$. Then, in a steady-state we must have

$$[u_h]^+ = [u_h]^-, h \in \{U, H\}. \quad (1.17)$$

Now we provide the steady state equation for workers employed on jobs (w, x) with health status h . The inflow of workers with health status h to jobs $(w, 1)$, denoted by

$[e_h^1(w)]^+$, is given as follows. If $w > \underline{w}_h^1$,

$$[e_h^1(w)]^+ \equiv (1 - \rho)f(w, 1)\lambda_u (u_h\pi_{hh}^0 + u_{h'}\pi_{hh'}^0) \quad (1.18a)$$

$$+ (1 - \rho)f(w, 1)\delta\lambda_e \times (\pi_{hh}^0 e_h^0 + \pi_{hh'}^0 e_{h'}^0 + \pi_{hh}^1 e_h^1 + \pi_{hh'}^1 e_{h'}^1) \quad (1.18b)$$

$$+ (1 - \rho)f(w, 1)(1 - \delta)\lambda_e \times \left[\begin{array}{c} \pi_{hh}^0 e_h^0 G_h^0(\underline{s}_h^0(w, 1)) + \pi_{hh'}^0 e_{h'}^0 G_{h'}^0(\underline{s}_h^0(w, 1)) \\ + \pi_{hh}^1 e_h^1 G_h^1(w) + \pi_{hh'}^1 e_{h'}^1 G_{h'}^1(w) \end{array} \right] \quad (1.18c)$$

$$+ (1 - \rho)(1 - \delta)\pi_{hh'}^1 e_{h'}^1 g_{h'}^1(w) \left[1 - \lambda_e \left(1 - \tilde{F}_h(w, 1) \right) \right], \quad (1.18d)$$

where $h' \neq h$ and $\tilde{F}_h(w, 1)$ is defined by

$$\tilde{F}_h(w, 1) = F(w, 1) + F(\underline{s}_h^0(w, 1), 0); \quad (1.19)$$

and $[e_h^1(w)]^+ = 0$ if $w < \underline{w}_h^1$. To understand expression (1.18), note that line (1.18a) presents the inflows from unemployed workers with health status h to job $(w, 1)$; line (1.18b) represents the inflow from those whose current match was destroyed but transitions to job $(w, 1)$ without experiencing an unemployment spell (recall our timing assumption 3(e) and 3(f) in Section 1.3.2); line (1.18c) represents inflows from workers who were employed on other jobs to job $(w, 1)$; and finally line ((1.18d) is the inflow from workers who were employed on the same job but has experienced a health transition from h' to h and yet did not transition to other better jobs, which occurs with probability $\left[1 - \lambda_e \left(1 - \tilde{F}_h(w, 1) \right) \right]$.

Denote the outflow of workers with health status h from jobs $(w, 1)$ by $[e_h^1(w)]^-$, and it is given by

$$[e_h^1(w)]^- \equiv e_h^1 g_h^1(w) \left\{ \begin{array}{c} [\rho + (1 - \rho)\pi_{hh}^1 \delta] + (1 - \rho)\pi_{h'h}^1 \\ + (1 - \rho)\pi_{hh}^1 \lambda_e (1 - \delta) \left[1 - \tilde{F}_h(w, 1) \right] \end{array} \right\}. \quad (1.20)$$

The outflow consists of job losses due to death and exogenous termination represented by the term $e_h^1 g_h^1(w) [\rho + (1 - \rho)\pi_{hh}^1 \delta]$, changes in current workers' health status represented by the term $e_h^1 g_h^1(w)(1 - \rho)\pi_{h'h}^1$, and transitions to other jobs represented by the term $e_h^1 g_h^1(w) (1 - \rho)\pi_{hh}^1 \lambda_e (1 - \delta) [1 - \tilde{F}_h(w, 1)]$. The steady state condition requires that

$$[e_h^1(w)]^+ = [e_h^1(w)]^- \text{ for } h \in \{U, H\} \text{ and for all } w \text{ in the support of } F(w, 1). \quad (1.21)$$

Similarly, the inflow of workers with health status h into jobs $(w, 0)$, denoted by $[e_h^0(w)]^+$, is as follows. If $w > \underline{w}_h^0$,

$$\begin{aligned} [e_h^0(w)]^+ &= f(w, 0)(1 - \rho)\lambda_u (u_h \pi_{hh}^0 + u_{h'} \pi_{hh'}^0) \\ &\quad + f(w, 0)(1 - \rho)\delta \lambda_e \\ &\quad \times (\pi_{hh}^1 e_h^1 + \pi_{hh'}^1 e_{h'}^1 + \pi_{hh}^0 e_h^0 + \pi_{hh'}^0 e_{h'}^0) \\ &\quad + f(w, 0)(1 - \rho)\lambda_e (1 - \delta) \\ &\quad \times \left[\begin{aligned} &\pi_{hh}^1 e_h^1 G_h^1(\underline{s}_h^1(w, 0)) + \pi_{hh'}^1 e_{h'}^1 G_{h'}^1(\underline{s}_h^1(w, 0)) \\ &+ \pi_{hh}^0 e_h^0 G_h^0(w) + \pi_{hh'}^0 e_{h'}^0 G_{h'}^0(w) \end{aligned} \right] \\ &\quad + (1 - \rho)(1 - \delta)\pi_{hh'}^0 e_{h'}^0 g_{h'}^0(w) \left[1 - \lambda_e (1 - \tilde{F}_h(w, 0)) \right], \end{aligned} \quad (1.22)$$

where $h \neq h'$ and $\tilde{F}_h(w, 0)$ is defined by

$$\tilde{F}_h(w, 0) = F(w, 0) + F(\underline{s}_h^1(w, 0), 1); \quad (1.23)$$

and $[e_h^0(w)]^+ = 0$ if $w < \underline{w}_h^0$. The outflow of workers with health status h from jobs $(w, 0)$, denoted by $[e_h^0(w)]^-$, is given by:

$$[e_h^0(w)]^- = e_h^0 g_h^0(w) \left\{ \rho + (1 - \rho) \left[\pi_{h'h}^0 + \pi_{hh}^0 (\delta + (1 - \delta)\lambda_e (1 - \tilde{F}_h(w, 0))) \right] \right\}. \quad (1.24)$$

The steady state condition thus requires that

$$[e_h^0(w)]^+ = [e_h^0(w)]^- \text{ for } h \in \{H, U\} \text{ and for all } w \text{ in the support of } F(w, 0). \quad (1.25)$$

From the four employment densities, $\langle e_h^x g_h^x(w) : h \in \{U, H\}, x \in \{0, 1\} \rangle$, we can define a few important terms related to *firm size*. First, given $e_h^x g_h^x(w)$ for $h \in \{U, H\}$ and $x \in \{0, 1\}$, the number of employees with health status h if a firm offers (w, x) is simply given by

$$n_h(w, x) = \frac{e_h^x g_h^x(w)}{f(w, x)}, \quad (1.26)$$

where the numerator is the total density of workers with health status h on the job (w, x) and the denominator is the total density of firms offering compensation package (w, x) . Of course, the total size of a firm that offers compensation package (w, x) is

$$n(w, x) = \sum_{h \in \{U, H\}} n_h(w, x) = \sum_{h \in \{U, H\}} \frac{e_h^x g_h^x(w)}{f(w, x)}. \quad (1.27)$$

Expressions (1.26) and (1.27) allow us to connect the firm sizes in steady state as a function of the entire distribution of employed workers $\{0, 1\}$.

Firm's Optimization Problem

A firm with a given productivity p decides what compensation package (w, x) to offer, taken as given the aggregate distribution of compensation packages $F(w, x)$. We assume that, before the firms make this decision, they each receive an i.i.d draw of $\sigma_f \epsilon$ where ϵ has a Type-I extreme value distribution and σ_f is a scale parameter. We interpret $\sigma_f \epsilon$ as an employer's idiosyncratic preference for offering health insurance. We assume that the $\sigma_f \epsilon$ shock a firm receives is persistent over time and it is separable from firm profits.²³

Given the realization of ϵ , each firm chooses (w, x) to maximize the steady-state flow profit *inclusive* of the shocks. It is useful to think of the firm's problem as a two-stage

²³These shocks allow us to smooth the insurance provision decision of the firms.

problem. First, it decides on the wage that maximizes the deterministic part of the profits for a given insurance choice; and second, it maximizes over the insurance choices by comparing the shock-inclusive profits with or without offering health insurance. Specifically, the firm's problem is as follows:

$$\max\{\Pi_0(p), \Pi_1(p) + \sigma_f \epsilon\}, \quad (1.28)$$

where

$$\Pi_0(p) = \max_{\{w_0\}} \Pi(w_0, 0) \equiv \left[\begin{array}{l} (p - w_0) n_H(w_0, 0) \\ + (pd - w_0) n_U(w_0, 0) \end{array} \right]; \quad (1.29)$$

$$\Pi_1(p) = \max_{\{w_1\}} \Pi(w_1, 1) \equiv \left[\begin{array}{l} (p - w_1 - m_H^1) n_H(w_1, 1) \\ + (pd - w_1 - m_U^1) n_U(w_1, 1) \end{array} \right] - C. \quad (1.30)$$

To understand the expressions (1.29), note that $n_H(w_0, 0)$ and $n_U(w_0, 0)$ are respectively the measure of healthy and unhealthy workers the firm will have in the steady state as described by (1.26) if it offers compensation package $(w_0, 0)$. Thus, $(p - w_0) n_H(w_0, 0)$ is the firm's steady-state flow profit from the healthy workers and $(pd - w_0) n_U(w_0, 0)$ is the flow profit from the unhealthy workers. The expressions (2.10) can be similarly understood after recalling that m_h^1 is the expected medical expenditure of worker with health status h and health insurance as defined in (1.4). For future reference, we will denote the solutions to problems (1.29) and (1.30) respectively as $w_0(p)$ and $w_1(p)$.

Due to the assumption that ϵ is drawn from i.i.d. Type-I extreme value distribution, firms' optimization problem (1.28) thus implies that the probability that a firm with productivity p offers health insurance to its workers is

$$\Delta(p) = \frac{\exp\left(\frac{\Pi_1(p)}{\sigma_f}\right)}{\exp\left(\frac{\Pi_1(p)}{\sigma_f}\right) + \exp\left(\frac{\Pi_0(p)}{\sigma_f}\right)}, \quad (1.31)$$

where $\Pi_0(p)$ and $\Pi_1(p)$ are respectively defined in (1.29) and (1.30).

1.3.4 Steady State Equilibrium

A *steady state equilibrium* is a list

$$\left\langle \left(\underline{w}_h^x, \underline{s}_h^x(\cdot, \cdot), \underline{q}_h^x \right), (u_h, e_h^x, G_h^x(w)), (w_x(p), \Delta(p)), F(w, x) \right\rangle \quad (1.32)$$

such that the following conditions hold:

- **(Worker Optimization)** Given $F(w, x)$, for each $(h, x) \in \{U, H\} \times \{0, 1\}$,
 - \underline{w}_h^x solves the unemployed workers' problem as described by (1.10);
 - $\underline{s}_h^x(\cdot, \cdot)$ solves the job-to-job switching problem for currently employed workers as described by (1.11);
 - \underline{q}_h^x describes the optimal strategy for currently employed workers regarding whether to quit into unemployment as described by (1.13);
- **(Steady State Worker Distribution)** Given workers' optimizing behavior described by $\left(\underline{w}_h^x, \underline{s}_h^x(\cdot, \cdot), \underline{q}_h^x \right)$ and $F(w, x), (u_h, e_h^x, G_h^x(w))$ satisfy the steady state conditions described by (1.14), (1.17), (1.21) and (1.25);
- **(Firm Optimization)** Given $F(w, x)$ and the steady state employee sizes implied by $(u_h, e_h^x, G_h^x(w))$, a firm with productivity p chooses to offer health insurance with probability $\Delta(p)$ where $\Delta(p)$ is given by (1.31). Moreover, conditional on insurance choice x , the firm offers a wage $w_x(p)$ that solves (1.29) and (1.30) respectively for $x \in \{0, 1\}$.
- **(Equilibrium Consistency)** The postulated distributions of offered compensation packages are consistent with the firms' optimizing behavior $(w_x(p), \Delta(p))$.

Specifically, $F(w, x)$ must satisfy:

$$F(w, 1) = \int_0^\infty \mathbf{1}(w_1(p) < w) \Delta(p) d\Gamma(p), \quad (1.33)$$

$$F(w, 0) = \int_0^\infty \mathbf{1}(w_0(p) < w) [1 - \Delta(p)] d\Gamma(p). \quad (1.34)$$

1.4 Qualitative Assessment of the Model

The complexity of the model precludes an analytical characterization of the equilibrium, thus we solve the equilibrium numerically.²⁴ The complexity of our model also prevents us from proving the existence and uniqueness of the equilibrium, but, throughout extensive numerical simulations, we always find a unique equilibrium for our baseline model based on our algorithm. We then present numerical simulation results using parameter estimates that we will report in Section 1.7 to illustrate how our model can generate the positive correlations among wage, health insurance and firm size we discussed in the introduction. We also use the numerical simulations to provide informal arguments about how some of key parameters of model are identified.

1.4.1 Numerical Simulations

In Column (1), labeled “Benchmark,” of Table 1.1, we report the main implications obtained from our benchmark model using parameter estimates that we report in Section 1.7. It shows that our baseline model is able to replicate the positive correlations among health insurance coverage rate, average wage, and employer size. Moreover, it also generates the empirically consistent prediction that the average health status of employees at firms offering health insurance is relatively better than those at firms not offering health insurance.

In Table 1.2, we use the estimates from Section 1.7 to shed light on the detailed

²⁴The details of our numerical algorithm are provided in Appendix A.1.

Statistics	Benchmark				
	(1)	(2)	(3)	(4)	(5)
Fraction of Firms Offering Health Insurance	0.5581	0.5651	0.5081	0.5401	0.5198
... if firm size is less than 50	0.5200	0.5275	0.5021	0.5029	0.5015
...if firm size is at least 50	0.9113	0.9135	0.5623	0.8837	0.6947
Fraction of Firms with Less than 50 Workers	0.9026	0.9026	0.8999	0.9023	0.8868
Average (4-month) Wages of Employed Workers	0.8501	0.8502	0.8688	0.8532	0.8868
... for insured employees	0.8986	0.8977	0.8680	0.9074	0.9125
... for uninsured employees	0.6211	0.6206	0.8698	0.6437	0.8400
Fraction of Healthy Workers	0.9494	0.9497	0.9644	0.9471	0.9363
... among the uninsured workers	0.9017	0.9019	0.9646	0.8999	0.8967
... among the insured workers	0.9600	0.9600	0.9641	0.9599	0.9586

Table 1.1: Predictions of the Baseline Model: Benchmark and Comparative Statistics.

Notes: (1). The benchmark predictions are based on the parameter estimates reported in Section 1.7. (2). The average wages are in units of \$10,000.

mechanisms for why in our model more productive firms have stronger incentives to offer health insurance than less productive firms. For this purpose, we simulate the health composition of the workforce for the firms with the bottom 20 and the top 20 productivity levels in our discretized productivity distribution. Row 1 of Table 1.2 shows that, in the steady state, the fraction of unhealthy workers in low and high productivity firms that offer health insurance are respectively 4.9% and 3.7%; in contrast, the fraction of unhealthy workers in low and high productivity firms that do not offer health insurance are respectively 9.6% and 10.7%. Offering health insurance seems to improve the health composition of workers over not offering health insurance for high-productivity firms, more so than for the low productivity firms. In Panels A-C, we disentangle the advantage of high-productivity firms relative to low-productivity firms in offering health insurance into three components.

In Panel A (or Row 2), we show that, in the low-productivity firms, the fraction of unhealthy among the new hires – including those hired directly from unemployment pool and those poached from other firms (i.e., job-to-job switchers) – is 8.0% if they offer health insurance and 7.4% if they do not; in contrast, in the high-productivity firms the fraction of unhealthy is 5.06% if they offer health insurance and 5.05% if they do not. Thus, the new hires attracted to firms that offer health insurance are indeed somewhat unhealthier, which is manifestation of adverse selection; but importantly, the new hires to high-productivity firms are significantly healthier than those to the low-productivity firms. This reflects the fact that, a high-productivity firm offering health insurance can poach workers from a much wider range of firms, including a larger fraction of workers from firms that already offer insurance and thus are healthier; in contrast, a low-productivity firm offering health insurance can only poach workers from firms with even lower productivity, most of which do not offer health insurance and thus have less healthy workers.

In Panel B, we show that any adverse selection effect that a firm offering health insurance suffers in terms of the health composition of their new hires is quickly remedied

Statistics	Low-Prod. Firms		High-Prod. Firms	
	HI	No HI	HI	No HI
[1] Frac. of Unhealthy Workers in Steady State	0.049	0.096	0.037	0.107
[2] Frac. of Unhealthy Among New Hires	Panel A: Adverse Selection Effect			
	0.080	0.074	0.0506	0.0505
	Panel B: Health Insurance Effect on Health			
[3] One-Period Ahead Frac. of Unhealthy Among New Hires	0.067	0.084	0.046	0.067
[4] Nine-Period Ahead Frac. of Unhealthy Among New Hires	0.038	0.109	0.037	0.107
	Panel C: Retention Effect			
[5] Job-to-Job Transition Rate for Healthy Workers	0.109	0.126	8.29E-9	4.03E-14
[6] Job-to-Job Transition Rate for Unhealthy Workers	0.104	0.126	8.29E-9	5.91E-5

Table 1.2: Understanding Why High-Productivity Firms Are More Likely to Offer Health Insurance than Low Productivity Firms.

Notes: For the simulations reported in this table, the low-productivity and high productivity firms are the firms with the bottom 20 and top 20 values of productivity in our discretized productivity support. See Footnote 1 in Appendix A.1.

by the positive effect of health insurance on health. In Row 3, we show that, just one-period later, the new hires' health composition is already in favor of firms that offer health insurance. For low-productivity firms, the fraction of unhealthy workers among those hired a period (4-months) ago, *were they not to leave*, is 6.7% and 8.4% respectively in those offering health insurance and those not offering health insurance. Similarly, for high-productivity firms, the fraction of unhealthy workers among those hired a period ago is 4.6% and 6.7% respectively in those offering health insurance and those not offering health insurance. In Row 4 we show that if the new hires from nine-periods (3 years) ago *were not to leave*, the fraction of unhealthy among them would be only 3.807% in low-productivity firms that offer health insurance, but it would be 10.9% in low-productivity firms that do not offer health insurance. Similarly, among high-productivity firms, the fraction of unhealthy workers among those hired nine periods ago, if they were not to leave, would be 3.7% and 10.7% respectively in those offering health insurance and those not offering health insurance.

Finally, in Panel C we show that the positive effect of health insurance on health, which leads to increased productivity of the workers, is better captured by high productivity firms. It shows that the job-to-job transition rates for workers in high-productivity firms, regardless of their health status, is significantly lower than that in low-productivity firms.

Thus in our model, high-productivity firms enjoy several advantages in offering health insurance to their workers relative to low-productivity firms: first, they face less severe adverse selection problem among the new hires; second, they are more likely to retain their healthy workers, which allows them to capture the increased productivity from the health improvement effect of health insurance as well as reduce the health care cost.

1.4.2 Comparative Statics

In Columns (2)-(5) of Table 1.1 we also present some comparative statics result to shed light on the effects of different parameters on the equilibrium features of our model. These

shed light on how different parameters may be identified in our empirical estimation.

Fixed Administrative Cost of Offering Health Insurance. In Column (2) of Table 1.1, we investigate the effect of the fixed administrative cost C on health insurance offering rate, by setting it to 0 as supposed to the estimated value of $C = 0.0730$ (i.e., \$730 per 4 months) as reported in Table 1.8. Comparing the results in Column (2) with the benchmark results in Column (1), we find that lowering the fixed administrative cost of offering health insurance affects mainly the coverage rate for small firms; and its effect on the insurance offering rate of large firms is much smaller. Moreover, it does not affect much of the other outcomes. Although we still have a positive correlation between firm size and health insurance offering rate, the offering rate for small firms is around 52.75% if $C = 0$.

Health Insurance Effect on Health. In Column (3), we shut down the effect of health insurance on the dynamics of health status by assuming that health transition process for the uninsured is the same as that of the insured, $\widehat{\pi}_{h'h}^0 = \pi_{h'h}^1$.²⁵ Column (3) of Table 1.1 shows that the fraction of large firms offering health insurance decrease significantly when $\widehat{\pi}_{h'h}^0$ is set to be equal to $\pi_{h'h}^1$. Moreover, this change significantly reduces the positive correlation between wage and health insurance. Therefore, the health insurance effect on health substantially affects the relationship among insurance offering rates, wages, and employer size in our model.

The reason why large firms decide not to offer health insurance when $\widehat{\pi}_{h'h}^0 = \pi_{h'h}^1$ can be understood as follows. When $\widehat{\pi}_{h'h}^0 = \pi_{h'h}^1$, i.e., when health insurance provision does not influence the dynamics of worker's health status, the health composition of a firm's workforce is fully determined by health composition of the workers at the time they accept the offer. The bottom two cells in Column (3) show that health composition of firms offering health insurance is worse than that of firms who do not, because health

²⁵We also obtain similar qualitative result in the opposite scenario, where health transition of the insured is set to be equal to that estimated for the uninsured, i.e., $\widehat{\pi}_{h'h}^1 = \pi_{h'h}^0$.

insurance provision attract more unhealthy workers. This creates an adverse selection problem which is not subsequently overcome as in Panel B of Table 1.2, thus leading to some firms not to provide coverage.

Risk Aversion. In Column (4) of Table 1.1 we simulate the effect on the equilibrium when we decrease the CARA coefficient from the estimated value of 0.4915 in Table 1.8 to 0.1. A reduction in CARA coefficient leads to a modest reduction in the health insurance offering rate for the large firms. It also increases the average wages in firms without health insurance.

Productivity Effect of Health. In Column (5) of Table 1.1 we investigate the productivity effect of health by changing d from 0.3386 in Table 1.8 to 1.00. This eliminates the negative productivity effect of bad health. Column (5) shows that the absence of the negative productivity effect of bad health leads to a substantial reduction of the coverage rate for the large employers relative to the benchmark. The reason is that, in the benchmark when bad health reduces productivity, the large firms, which tend to retain workers longer as shown in Panel C of Table 1.2, have stronger incentive than smaller firms to improve the health of their workforce in order to raise the expected flow profit. Moreover, an increase in d increases firms' wage offers in general due to the productivity improvement.

1.4.3 Identification of γ , d , C and σ_f

As shown in Columns (2) and (4)-(5) in Table 1.1, the CARA coefficient γ , the productivity effect of health d , and the fixed administrative cost of offering health insurance C , all have important effects on the firms' incentives to provide health insurance. How are they separately identified? Here we provide some "heuristic" discussion.

As we detail in Section 1.6, in our estimation we use both worker-side data which has information about workers' labor market dynamics and firm-side data that has information

about firm size, wages and health insurance offering. While it is true that the CARA coefficient γ affects the firms' incentives to provide insurance as shown in Column (4) of Table 1.1, it also affects the workers' job-to-job transitions. In particular, if γ is larger (i.e. when workers are more risk averse), we would expect to observe *more frequent* transitions of workers from jobs *without* health insurance to a job *with* health insurance, especially after a deterioration of health status, and even if the transition involves a reduction in wages. Moreover, the magnitude of the wage cut a worker is willing to tolerate in order to switch from a job without health insurance to a job with health insurance increases with the risk aversion parameter γ . These effects are not shown in Table 1.1, but will be incorporated in our estimation via the likelihood function of the workers' labor market transition dynamics.

As shown in Columns (4) and (5) in Table 1.1, both the productivity effect of health d and risk aversion γ affect the relationship between the probability of offering health insurance and firm size. Of course, the scale parameter σ_f in (1.31) also affects the relationship between the probability of offering health insurance and firm productivity (and thus firm size). These three parameters are separately identified for the following reasons. First, the risk aversion parameter γ is disciplined by the worker-side job-to-job transition information as we described above; second, even though the parameter d and the scale parameter σ_f both affect the *slope* between the firm size and insurance offering probability from the firm-side data, the parameter d has an additional effect on the differences in wages for firms depending on whether they offer health insurance. Finally, the administrative cost C is identified from the the probability (in *level*) of small firms offering health insurance.

1.5 Data Sets

In this section, we describe our data sets and sample selection. In order to estimate the model, it is ideal to use employee-employer matched dataset which contains information

about worker’s labor market outcome and its dynamics, health, medical expenditure, and health insurance, and firm’s insurance coverage rate and size. Unfortunately, such a data set does not exist in the U.S. Instead, we combine three separate data sets for our estimation: (1) Survey of Income and Program participation; (2) Medical Expenditure Panel Survey; and (3) Robert Wood Johnson Employer Health Insurance Survey.

1.5.1 Survey of Income and Program Participation

Our main dataset for individual labor market outcome, health, and health insurance is 1996 Panel of Survey of Income Program Participation (hereafter, SIPP 1996).²⁶ SIPP 1996 interviews individuals *every four months* up to twelve times, so that an individual may be interviewed over a four-year period. It consists of two parts: (1) core module, and (2) topical module. The core module, which is based on interviews in each wave, contains detailed *monthly* information regarding individuals’ demographic characteristics and labor force activity, including earnings, number of weeks worked, average hours worked, employment status, as well as whether the individual changed jobs during each month of the survey period. In addition, information for health insurance status is recorded in each wave; it also specifies the source of insurance so we know whether it is an employment-based insurance, a private individual insurance, or Medicaid, and we also know whether it is obtained through the individual’s own or the spouse’s employer. The topical module contains *yearly* information about the worker and his/her family member’s self reported health status and *out-of-pocket* medical expenditure at interview waves 3, 6, 9 and 12.²⁷

Sample Selection Criterion. The total sample size after matching the topical module and the core module is 115,981. In order to have an estimation sample that is somewhat homogeneous in skills as we assume in our model, we restrict our sample to men (drop-

²⁶SIPP 1996 Panel is available at: http://www.census.gov/sipp/core_content/1996/1996.html

²⁷In both SIPP and MEPS, we use the self-reported health status to construct whether the individual is healthy or unhealthy. The self-reported health status has five categories. We categorize “Excellent”, “Very Good” and “Good” as *Healthy* and “Fair” and “Poor” as *Unhealthy*.

ping 59,846 female individuals) whose ages are between 26-46 (dropping an additional 38,016 individuals). In addition, we only keep individuals who are not in school, not self-employed, do not work in the public sector, are not engaged in the military, and do not participate in any government welfare program (dropping an additional 6,995 individuals in total). We also require that our sample be covered either by an employer-based health insurance in his own name or is uninsured (dropping an additional 1,948 individuals). We restrict our samples to individuals who are at most high school graduates (dropping 3,060 individuals). Finally we drop top and bottom 3% of salaried workers (dropping an additional 817 individuals). Our final estimation sample that meets all of the above selection criterion consists of a total of 5,309 individuals.

1.5.2 Medical Expenditure Panel Survey (MEPS)

The weakness of using SIPP data for our research is the lack of information for total medical expenditure. To obtain the information, we use Medical Expenditure Panel Survey (hereafter, MEPS) 1997-1999. We use its Household Component (HC), which interviews individuals *every half year* up to five times, so that an individual may be interviewed over a two-and-a-half-year period.²⁸ Medical expenditure is recorded at annual frequency. Several health status related variables are recorded in each wave. Moreover, health insurance status is recorded at monthly level. We use the same sample selection criteria as SIPP 1996. The sample size is 4,815.

1.5.3 Robert Wood Johnson Foundation Employer Health Insurance Survey

In addition, we also need information for employer size and associated health insurance offering rate, which is not available from the worker-side data. The data source we use is 1997 Robert Wood Johnson Foundation Employer Health Insurance Survey (hereafter,

²⁸MEPS HC is publicly available at <http://www.meps.ahrq.gov>.

RWJ-EHI).²⁹ It is a nationally representative survey of public and private establishments conducted in 1996 and 1997. It contains information about employer’s characteristics such as industry, firm size, and employees’ demographics, as well as information about health insurance offering, health insurance plans, employees’ eligibility and enrollment in health plans, and the plan type.

We restrict the sample to establishments which belong to the private sector and have at least three employees. The final sample size is 19,089.

1.5.4 Summary Statistics

Table 1.3 reports the summary statistics of the key variables in the 1996 SIPP data. About 76% of the employed workers receive health insurance from their employers; the average 4-month wage for employed workers with health insurance is about \$9,240, higher than that for those without health insurance which is about \$6,187. The unemployment rate for our selected sample is about 3.18%, lower than the overall unemployment rate in the U.S. in 1996 (which was about 5.4%).³⁰ About 95.11% of our sample reported their health to be healthy (i.e. either “Good”, “Very Good”, or “Excellent”). Moreover, it is important to note that 95.36% of the workers with insurance and 93.89% of those without insurance reported healthy.

In Table 1.4 we report the comparison of summary statistics for the individuals in MEPS 1997-1999 and those in SIPP 1996. Both the fraction of healthy workers and the fraction of employed workers who own health insurance are somewhat lower in MEPS than in SIPP. By using the mean expenditure given health and health insurance in MEPS, we also impute the annual average medical expenditure based on SIPP’s health and health insurance composition for the SIPP sample. It shows that annual medical expenditures are similar in the two samples.

Finally, in Table 1.5 we provide the summary statistics for our firm side data based

²⁹It is publicly available at <http://www.icpsr.umich.edu/icpsrweb/HMCA/studies/2935>

³⁰See U.S. Department of Labor, Bureau of Labor Statistics at website: <http://stats.bls.gov>.

Variable	Mean	Std. Dev.
Fraction of Insured Among Employed Workers	0.7619	0.4260
Average (4-Month) Wages for Employed Workers	0.8538	0.3532
... for insured employees	0.9240	0.3462
... for uninsured employees	0.6187	0.2750
Fraction of Unemployed Workers	0.0318	0.1758
Fraction of Healthy Workers	0.9511	0.2177
... among insured workers	0.9536	0.2103
... among uninsured workers	0.9389	0.2398

Table 1.3: Summary Statistics: SIPP 1996.

Notes: The average wages are in units of \$10,000.

Variable	MEPS		SIPP
	(1)	(2)	(2)
Fraction of Healthy Workers	0.912 (0.284)	0.951 (0.218)	
Fraction of Insured Among Employed Workers	0.651 (0.477)	0.762 (0.426)	
Annual Medical Expenditure	0.077 (0.337)	0.079*	
... for those with health insurance and who are healthy	0.078 (0.265)	-	
... for those without health insurance and who are healthy	0.043 (0.366)	-	
... for those with health insurance and who are unhealthy	0.293 (0.603)	-	
... for those without health insurance and who are unhealthy	0.133 (0.451)	-	

Table 1.4: Summary Statistics: Comparison between MEPS 1997-1999 and SIPP 1996.

Notes: (1). The average wages are in units of \$10,000. (2). Standard deviations are in parentheses. (3). The annual medical expenditure for SIPP is imputed based on the average annual medical expenditures for workers of different health insurance and health status combinations computed from MEPS, reported in Column (1), using the fractions of the workers of the four health insurance and health status combinations that can be calculated from Table 1.3.

on RWJ-EHI 1997. In general, firms that tend to offer health insurance have large size in employment and provide higher wage. Moreover, wages, both unconditional and conditional on insurance status, are very close to the one reported for the 1996 SIPP sample. Therefore, although we restrict samples to relatively unskilled workers in SIPP, the compensation patterns seem to be quite consistent in the worker-side and employer-side data sets.

1.6 Estimation Strategy

In this section we present our strategy to structurally estimate our baseline model using the datasets we described above.³¹ We estimate parameters regarding health transitions and medical expenditure distribution without using the model. The remaining parameters are estimated via a minimum-distance estimator which follows Imbens and Lancaster (1994) and Petrin (2002). They consider the situation where moments come from different data sources. In this study we construct worker-side moments from the likelihood of individuals' labor market transition, as in Bontemps, Robin, and Van den Berg (1999, 2000) and Shephard (2012). Then, we construct firm-side moments such as firm size distribution and firm's coverage rate conditional on their size from employer-side data. Loosely speaking, the parameters are chosen to best fit the data from *both* sides of labor markets. This is the main difference from the existing estimation procedure used in Bontemps, Robin, and Van den Berg (1999, 2000) and Shephard (2012), where model parameters are chosen to fit worker side data alone.³² As a result, we assume a parametric specification

³¹The details of the numerical estimation procedure are available in Appendix A.4.

³²Consequently they can estimate productivity distribution nonparametrically so that the model's prediction of workers' wage distribution perfectly fits with the data. Specifically, in Bontemps, Robin, and Van den Berg (1999, 2000) and Shephard (2012), worker-side parameters are estimated from the likelihood function of individual labor market transitions. Then, firm productivity distribution is estimated to perfectly fit wage distribution observed from the worker side by utilizing the theoretical relationship between wage offer and firm productivity implied from the model. Note that one can still apply semiparametric multi-step estimation to fit both worker and employer side moments if one has access to employee-employer matched panel data. For example, Postel-Vinay and Robin (2002) and Cahuc, Postel-Vinay, and Robin (2006) nonparametrically estimate worker's sampling distribution of job offer from each firm to match observed wage distribution. Given the estimated sampling distribution, they

Variable Name	Mean	Std. Dev.
Average Establishment Size	19.92	133.40
... for those that offer health insurance	30.08	177.24
... for those that do not offer health insurance	6.95	11.03
Health Insurance Coverage Rate	0.56	0.50
... for those with less than 50 workers	0.53	0.50
... for those with 50 or more workers	0.95	0.23
Average Annual Wage Compensation, in \$10,000	2.53	2.44
... for those that offer health insurance	2.92	2.50
... for those that do not offer health insurance	2.03	2.27

Table 1.5: Summary Statistics: RWJ-EHI 1997.

Note: Standard deviations are in parentheses.

of the productivity distribution and it is estimated, jointly with other parameters, to fit both the wage and firm size distributions. Specifically, as we mentioned in Section 1.3.1, we specify that the productivity distribution is given by a lognormal distribution with mean μ_p and variance σ_p^2 .

In our empirical application, the model period is set to be four months, driven by the fact that we can only observe the transition of health insurance status at four-month intervals in the SIPP data. In this study, we do not try to estimate β but set $\beta = 0.99$ so that annual interest rate is about 3%.³³ Moreover, we set the exogenous death rate ρ to be 0.001.³⁴ Finally, the after-tax income schedule (1.6) is set to be that estimated in Kaplan (2012), i.e., $\tau_0 = 0.0056$, $\tau_1 = 0.6377$ and $\tau_2 = -0.1362$.

1.6.1 First Step

In Step 1 we estimate parameters $(\alpha_0, \beta_0, \gamma_0)$ in the probability of receiving a medical shock in (1.2) and the parameters $(\alpha_m, \beta_m, \gamma_m, \sigma_{hx} : h \in \{H, U\}, x \in \{0, 1\})$ in the distribution of medical expenditures as specified in (1.3), as well as the health transitions π^x as in (1.5) without explicitly using the model. They are estimated by GMM using the MEPS data. A total of twelve sample target moments are used, including the mean and variance of the medical expenditure conditional on health and health insurance status (eight moments) and the fraction of individuals with zero medical expenditures conditional on health and health insurance status (four moments). For simplicity, we estimate these parameters using a subsample of individuals whose health and health insurance status are *unchanged* throughout the year. The *annual* theoretical moments conditional on health insurance and health status are constructed from parameters which are defined for our model with a four-month period.³⁵

then estimate productivity distribution of firms to perfectly fit the employer-size distribution.

³³ It is known from Flinn and Heckman (1982) that it is difficult to separately identify the discount factor β from the flow unemployed income b in standard search models.

³⁴This roughly matches the average 4-month death rate for men in the age range of 26-46, which is the sample of individuals we include in our estimation.

³⁵We use the identity weighting matrix for simplicity.

Because we are assuming that the effect of health insurance and health status on medical shocks and medical expenditures are exogenous, our restriction to the subsample of individuals whose health and health insurance status are *unchanged* throughout the year does not create a biased sample for our estimation purpose. However, it is useful to recognize that this subsample differs from the overall MEPS sample. Table 1.6 provides the analogous summary statistics of the MEPS subsample we used in our first step estimation.³⁶ The comparison of Tables 1.6 and 1.4 shows that, not surprisingly, the magnitudes of medical expenditure are substantially lower in this subsample than those in the overall sample.

We estimate the parameters in health transition matrix $(\pi_{HH}^1, \pi_{UU}^1, \pi_{HH}^0, \pi_{UU}^0)$ using the 1996 SIPP data based on maximum likelihood. The key issue we need to deal with is that our model period is 4 months; and while we can observe health insurance status each period (every four months), we observe health status only every *three* periods (a year). We deal with this issues as follows. Let $x_t \in \{0, 1\}$ be worker's insurance status at period t , and let $h_t \in \{H, U\}$ and $h_{t+3} \in \{H, U\}$ denote respectively the worker's health status in period t and $t+3$ (when it is next measured), the likelihood of observing $h_{t+3} \in \{H, U\}$ conditional on x_t, x_{t+1}, x_{t+2} and $h_t \in \{H, U\}$ can be written out explicitly using the Law of Total Probability. For example, if $h_{t+3} = H$ and $h_t = H$, we have:³⁷

$$\begin{aligned} \Pr(h_{t+3} = H | x_t, x_{t+1}, x_{t+2}, h_t = H) &= \pi_{HH}^{x_t} \pi_{HH}^{x_{t+1}} \pi_{HH}^{x_{t+2}} \\ &\quad + \pi_{HH}^{x_t} (1 - \pi_{HH}^{x_{t+1}}) (1 - \pi_{UU}^{x_{t+2}}) \\ &\quad + (1 - \pi_{HH}^{x_t}) (1 - \pi_{UU}^{x_{t+1}}) \pi_{HH}^{x_{t+2}} \\ &\quad + (1 - \pi_{HH}^{x_t}) \pi_{UU}^{x_{t+1}} (1 - \pi_{UU}^{x_{t+2}}). \end{aligned}$$

We use them to formulate the log-likelihood of observed data, which records the health transition every three periods, as a function of one-period health transition parameters

³⁶The sample size is 2,892.

³⁷The formulae for the other cases are analogous and are available in Appendix A.2.

Variable	Mean	Std. Dev.
Fraction of Healthy Workers	0.9723	0.1640
Fraction of Insured Among Employed Workers	0.6677	0.4711
Annual Medical Expenditure, in \$10,000	0.0621	0.3033
... for those with health insurance and who are healthy <i>throughout the year</i>	0.0675	0.1987
... for those without health insurance and who are healthy <i>throughout the year</i>	0.0365	0.4003
... for those with health insurance and who are unhealthy <i>throughout the year</i>	0.4804	0.9299
... for those without health insurance and who are unhealthy <i>throughout the year</i>	0.1249	0.2955

Table 1.6: Summary Statistics of the Subsample of the MEPS 1997-1999 Used in the Estimation of Medical Expenditure Distributions in the First Step.

as captured by π^x , for $x \in \{0, 1\}$, as in (1.5) in our model.

1.6.2 Second Step

In the second step, we estimate the remaining parameters $\theta = (\theta_1, \theta_2)$ where $\theta_1 = (\lambda_u, \lambda_e, \delta, \gamma, \mu, b)$ are parameters that affect worker-side dynamics and their preferences to work and to have health insurance, and $\theta_2 = (C, d, M, \mu_p, \sigma_p)$ are the additional parameters that are mostly relevant to the firm-side moments. Our objective function is based on the optimal GMM which consists of two types of moments. The first set of moments are derived from the worker-side data in SIPP in the form of the log-likelihood of the observed labor market dynamics of the workers, which we aim to maximize by requiring that the first derivatives should be equal to zero, following Imbens and Lancaster (1994). The second set of moments come from the firm-side data RWJ-EHI.

Specifically, let the targeted moments be

$$g(\theta) = \begin{bmatrix} \frac{\sum_i \partial \log(L_i(\theta))}{\partial \theta} \\ \mathbf{s} - \mathbb{E}[\mathbf{s}; \theta] \end{bmatrix}, \quad (1.35)$$

where $L_i(\theta)$ is individual i 's contribution to the labor market dynamics likelihood, which we discuss in details in Section 1.6.3; and \mathbf{s} is a vector of firm-side moments we describe in Section 1.6.4. Then, we construct an objective function as

$$\min_{\{\theta\}} g(\theta)' \Omega g(\theta), \quad (1.36)$$

where the weighting matrix Ω is chosen as a consistent estimator of $\mathbb{E}[g(\theta)'g(\theta)]^{-1}$, which is obtained using $\tilde{\theta}$, a preliminary consistent estimate of θ . As in Petrin (2002), we first assume that $\mathbb{E}[g(\theta)'g(\theta)]$ takes block diagonal matrix because different moments come from different sampling processes. Let $G(\theta) = \mathbb{E}[\frac{\partial g(\theta)}{\partial \theta}]$, the gradient of the moments with respect to the parameters evaluated at the true parameter values. The asymptotic

variance of $\sqrt{n}(\hat{\theta} - \theta)$ is then given by

$$[G(\theta)' \Omega G(\theta)]^{-1},$$

which we use to calculate the standard error of parameter estimates.

1.6.3 Deriving the Likelihood Functions of Workers' Labor Market Dynamics

Here we derive the likelihood functions of workers' labor market dynamics similar to those in Bontemps, Robin, and Van den Berg (1999, 2000). Let $F(w, x)$ denote the distribution of (w, x) in the labor market.

We will first derive the likelihood contribution of the labor market transitions of *unemployed workers*. Consider an unemployed worker at period 1 with health status is h_1 , who experiences an unemployment spell of duration l and in period $l + 1$ transitions to a job (\tilde{w}, x) . To ease exposition, let us first *suppose* that health history between $j = 1$ to $l + 1$ for this worker, $(h_1, h_2, \dots, h_{l+1})$, is observed. The likelihood contribution of observing such a transition is:

$$\frac{u_{h_1}}{M} \times \prod_{j=2}^l \left\{ \begin{array}{l} \Pr(h_j | h_{j-1}, x_{j-1} = 0) \\ \times \left[(1 - \lambda_u) + \lambda_u \left(F(\underline{w}_{h_j}^1, 1) + F(\underline{w}_{h_j}^0, 0) \right) \right] \end{array} \right\} \quad (1.37a)$$

$$\times \Pr(h_{l+1} | h_l, x_l = 0) \times [\lambda_u f(\tilde{w}, 1)]^{\mathbf{1}(x=1)} \times [\lambda_u f(\tilde{w}, 0)]^{\mathbf{1}(x=0)} \quad (1.37b)$$

where $\mathbf{1}(x = 1)$ is an indicator function taking value one if we observe a transition to employment with $(\tilde{w}, 1)$ at period $l + 1$, and similarly $\mathbf{1}(x = 0)$ is an indicator function taking value one if we observe a transition to employment with $(\tilde{w}, 0)$ at period $l + 1$. To understand (1.37), note that the first term in line (1.37a), u_h/M , reflects the assumption that the initial condition of individuals is drawn from the steady state worker distribution because u_h/M the probability that an unemployed worker with health status h is sampled.

The second term in line (1.37a) is the probability that individual experiences l periods of unemployment with health status transitions (h_2, \dots, h_l) during the process; note that the term $\left[(1 - \lambda_u) + \lambda_u (F(\underline{w}_{h_j}^1, 1) + F(\underline{w}_{h_j}^0, 0)) \right]$ is the probability that the individual does not receive an offer or receives an offer that is lower than the relevant reservation wages $\underline{w}_{h_j}^1$ or $\underline{w}_{h_j}^0$. The term on line (1.37b) is the probability that his health transitions from h_l to h_{l+1} in period $l + 1$ and receive an acceptable offer (\tilde{w}, x) from the relevant density function $f(\tilde{w}, x)$.

Now as we described earlier in Section 1.5, SIPP data we observe the workers' self-reported health status only annually (at interview waves 3, 6, 9 and 12); as a result, we do not always observe workers' health history in-between labor market transitions. However, since we already estimated the health transitions conditional on health insurance in Step 1, we can integrate out the unobserved health status.³⁸

We can similarly derive the likelihood contribution of the job dynamics of *employed workers*. Consider an employed worker in period 1 with health status h_1 working on a job with compensation package (w, x) . Suppose that the worker experiences a job status changes in period $l + 1$. For an employed worker, there are four possible job status changes:

- [Event “*Job Loss*”]: The individual experienced a job loss at period $l + 1$;
- [Event “*Switch 1*”]: The individual transitioned to a job (\tilde{w}, x') such that $x' = x$ and the accepted wage is $\tilde{w} > w$;
- [Event “*Switch 2*”]: The individual transitioned to a job (\tilde{w}, x') such that $x' = x$ and the accepted wage is $\tilde{w} < w$;
- [Event “*Switch 3*”]: The individual transitioned to a job (\tilde{w}, x') such that $x' \neq x$ and the accepted wage is \tilde{w} .

Again, suppose that the health history between $j = 1$ to $l + 1$ for this worker,

³⁸Details for the likelihood functions when the health history in-between labor market transitions are not observable are provided in Appendix A.3.

$(h_1, h_2, \dots, h_{l+1})$, is observed, then the likelihood contribution is given by:

$$\frac{e_h^x g_h^x(w)}{M} \times \prod_{j=2}^l \left\{ \begin{array}{c} \Pr(h_j|h_{j-1}, x)(1 - \delta) \\ \times \left[(1 - \lambda_e) + \lambda_e \left(F(w, x) + F(\underline{s}_{h_j}^{x'}(w, x), x') \right) \right] \end{array} \right\} \quad (1.38a)$$

$$\times \Pr(h_{l+1}|h_l, x) \times \left\{ \begin{array}{ll} \delta \left[(1 - \lambda_e) + \lambda_e \sum_{\tilde{x}} F(\underline{w}_{h_{l+1}}^{\tilde{x}}, \tilde{x}) \right] & \text{if Event is "Job Loss"} \\ \lambda_e f(\tilde{w}, x) & \text{if Event is "Switch 1"} \\ \delta \lambda_e f(\tilde{w}, x) & \text{if Event is "Switch 2"} \\ (1 - \delta) \lambda_e f(\tilde{w}, x') + \delta \lambda_e f(\tilde{w}, x') & \text{if Event is "Switch 3"} \end{array} \right. \quad (1.38b)$$

To understand (1.38), note that similar to that in (1.37), the first term in line (1.38a), $e_h^x g_h^x(w)/M$, is the probability of sampling an employed worker with health status h working on a job (w, x) ; the second term in line (1.38a) is the probability that individual stays with the job (w, x) for l periods with health status transitions (h_2, \dots, h_l) during the process. Line (1.38b) expresses the likelihood of observing health transition from h_l to h_{l+1} in period $l+1$ and one of the four job status change events. For example, the event “Job Loss” is observed in period $l+1$ with probability $\delta \left[(1 - \lambda_e) + \lambda_e \sum_{\tilde{x}} F(\underline{w}_{h_{l+1}}^{\tilde{x}}, \tilde{x}) \right]$ because in order for a job loss to occur, the worker has to experience an exogenous shock that destroys the current match (which occurs with probability δ), and then he does not get matched to another acceptable job (which occurs with probability $(1 - \lambda_e) + \lambda_e \sum_{\tilde{x}} F(\underline{w}_{h_{l+1}}^{\tilde{x}}, \tilde{x})$). To understand the probability of event “Switch 2”, we note that in order for a worker to switch to a job (\tilde{w}, x') with $x' = x$ but $\tilde{w} < w$, the worker must have experienced a job separation (which occurs with probability δ), but is then lucky enough to find an acceptable job immediately, which happens with probability $\lambda_e f(\tilde{w}, x)$. The probability of the other job switch events are derived similarly.

1.6.4 Employer-Side Moments

In our estimation, we also require that our model’s predictions match the following employer-side moments calculated from the RWJ-EHI data. These moments correspond

to the vector \mathbf{s} in expression (1.35):

- Mean firm size;
- Fraction of firms less than 50 workers;
- Mean size of firms that offer health insurance;
- Mean size of firms that do not offer health insurance;
- Health insurance coverage rate;
- Health insurance coverage rate among firms with more than 50 workers;
- Health insurance coverage rate among firms with less than 50 workers;
- Average wages of firms with less than 50 workers;
- Average wages of firms with more than 50 workers.

1.7 Estimation Results

1.7.1 Parameter Estimates

Parameters Estimated in the First Step. Table 1.7 reports the parameter estimates from step 1. The estimated coefficients imply that unhealthy individuals and individuals with health insurance are more likely to experience medical shocks; moreover, conditional on experiencing medical shocks, unhealthy and insured individuals are more likely to incur larger medical expenditures. The finding that health insurance increases both the incidence and magnitude of medical expenditures captures the moral hazard effect of health insurance. As we report in Table 1.9, our model fits the means and variances of medical expenditures by health and healthy insurance status in the data well.

In Panel B of Table 1.7, we find that there is a significant health insurance effect on the dynamics of health since $\pi_{HH}^1 > \pi_{HH}^0$ and $\pi_{UU}^1 < \pi_{UU}^0$, implying that not having health insurance increases the probability that the next period health status is unhealthy. It is also interesting to note that our estimates indicate that health insurance has a higher marginal effect on health for a currently unhealthy worker than for a currently healthy

Parameter	Estimate	Std. Err.
Panel A: Medical Expenditure Parameters in Eq. (1.4)		
α_0	-1.0909	(0.0446)
β_0	0.5247	(0.0723)
γ_0	0.5787	(0.0747)
α_m	-4.4222	(0.3099)
β_m	1.6262	(0.3268)
γ_m	0.7227	(0.3867)
α_p	-1.0909	(0.0446)
σ_{H1}	1.4783	(0.0662)
σ_{H0}	1.9895	(0.1235)
σ_{U1}	1.3584	(0.0919)
σ_{U0}	1.3193	(0.0173)
Panel B: Health Transition Parameters in Eq. (1.5)		
π_{HH}^1	0.9865	(0.0023)
π_{HH}^0	0.9689	(0.0058)
π_{UU}^1	0.7294	(0.0310)
π_{UU}^0	0.7587	(0.0365)

Table 1.7: Parameter Estimate from Step 1.

footnotesizeNote: Standard errors are in parentheses. The unit of medical expenditure is \$10,000.

worker, that is, $\pi_{HU}^1 - \pi_{HU}^0 > \pi_{HH}^1 - \pi_{HH}^0$.

Parameters Estimated in the Second Step. Table 1.8 reports the parameter estimates from step 2. Panel A shows that our estimate of CARA coefficient is about 0.4915E-4 (recalling that our unit is in \$10,000). Using the four-month average wages for employed workers reported in Table 1.3, which is about \$8,538, our estimated CARA coefficient implies a relative risk aversion of about 0.42. These are squarely in the range of estimates of CARA and Relative Risk Aversion coefficients in the literature (see Cohen and Einav (2007) for a summary of such estimates).

We find that the offer arrival rate for an employed worker, λ_e , is about 0.268, which implies that on average it takes about 15 months for a currently employed worker to receive an outside offer; we also find that the offer arrival rate for an unemployed worker, λ_u , is 0.434, implying that on average it takes about 9 months for an unemployed to receive an offer.³⁹ Our estimate of the unemployment income b is small, about \$137,

³⁹Dey and Flinn (2005) estimated that the mean wait between contacts for the unemployed is about

Parameter	Estimates	Std. Err.
Panel A: Parameters in $\theta_1 \equiv (\lambda_u, \lambda_e, \delta, \gamma, \mu, b)$		
γ	0.4915	(0.0051)
b	0.0137	(0.0002)
λ_u	0.4340	(0.0112)
λ_e	0.2680	(0.0038)
δ	0.0179	(0.0003)
μ_H	0.9930	(0.0156)
Panel B: Parameters in $\theta_2 \equiv (C, d, M, \mu_p, \sigma_p, \sigma_f)$		
d	0.3386	(0.0063)
C	0.0730	(0.0063)
M	18.8920	(8.7940)
μ_p	-0.5680	(0.0031)
σ_p	0.4043	(0.0036)
σ_f	0.2397	(0.0025)

Table 1.8: Parameter Estimate from Step 2.

reflecting that a large fraction of the UI benefits are probably expensed for job search costs. In Panel A, we also report that our estimate of the probability of exogenous job destruction, δ , is about 1.79% in a four-month period; and the fraction of newly arrived workers who are healthy is about 99.30%.

Panel B reports our estimates of parameters $\theta_2 \equiv (C, d, M, \mu_p, \sigma_p)$. We find that the productivity of a worker in bad health, d , is only 0.3386, implying that there is a significant amount of productivity loss from bad health. This seems plausible because we categorize only those whose self-reported health is “Poor” or “Fair” as unhealthy. Moreover, we find that the fixed administration cost of offering health insurance is about \$730 per four month, i.e., about \$2,190 per year.

In order to fit the average firm size, our estimate of M , the ratio between workers and firms, is about 18.892. This estimate is smaller than the average establishment size of 19.923.25 months, while the a contact between a new potential employer and a currently employed individual occurs about every 19 months. The differences for the contact rate for the unemployed between our study and Dey and Flinn (2005) could be due to the fact that a period is four months in our study while it is a week in Dey and Flinn (2005). An unemployed individual in both the first month and the fifth month will be considered as being in a continuous unemployment spell, though at weekly frequency he could have been matched with some firms inbetween. This may lead us to a lower estimate for the contact rate for the unemployed. Another possibility is the differences in the sample selection: our sample includes only individuals with no more than high school degree, while Dey and Flinn (2005)’s sample has at least a high school degree.

reported in Table 1.5 because in our model some low-productivity firms do not attract any workers in equilibrium. We also estimated that the scale and shape parameters of the lognormal productivity distribution are respectively -0.5860 and 0.4043, which implies that the mean (4-month) productivity of firms is about 0.6149 (i.e. \$6149). The fact that the mean accepted four-month wage in our sample is 0.8538 (see Table 1.3) is largely due to the fact more productive firms attract more workers in the steady state as our model implies, but also due to the fact that a fraction of the low-productivity firms are not able to attract any workers in equilibrium i.e., they are inactive.

1.7.2 Within-Sample Goodness of Fit

In this section, we examine the within-sample goodness of fit of our estimates by simulating the equilibrium of our estimated model and compare the model predictions with their data counterparts.

Worker-Side Goodness of Fit. Table 1.9 reports the model fits for medical expenditure in the first step. It shows that our parameter estimates fit the data on the means (in Panel A) and variances (in Panel B) of medical expenditure conditional on health and health insurance status very well; moreover, in Panel C we show that we accurately replicate the fraction of individuals with zero medical expenditures conditional on health and health insurance status.

Table 1.10 reports the model fit for the worker-side moments. It shows that the model fits really well for cross section worker distribution in terms of health, health status, health insurance, wage, and employment distribution. Note that these moments are not directly targeted in our estimation.

Figure 1.1 plots the distribution of workers' accepted wages by health insurance status. It shows that our model is able to capture the overall patterns reasonably well, but it predicts a much more concentrated wage distribution than what is in the data, especially

	Data	Model
Panel A: Mean Annual Medical Expenditure		
Healthy & insured	0.0672	0.0673
Healthy & uninsured	0.0365	0.0359
Unhealthy & insured	0.4804	0.4794
Unhealthy & uninsured	0.1249	0.1249
Panel B: Variance of Annual Medical Expenditure		
Healthy & insured	0.0393	0.0392
Healthy & uninsured	0.1601	0.1601
Unhealthy & insured	0.8084	0.8084
Unhealthy & uninsured	0.0856	0.0856
Panel C: Fraction with Zero Medical Expenditure		
Healthy & insured	0.3324	0.3368
Healthy & uninsured	0.6458	0.6413
Unhealthy & insured	0.1290	0.1213
Unhealthy & uninsured	0.3600	0.3646

Table 1.9: Fit for Medical Expenditure Distributions: Model vs. Data.

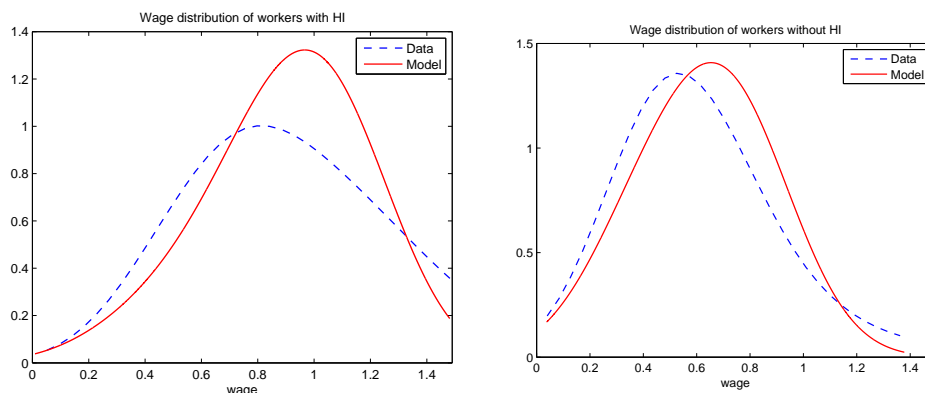


Figure 1.1: The Distributions of Workers’ Accepted Wages by Health Insurance Status: Model vs. Data.

among workers who have health insurance from their employers.

Employer-Side Goodness of Fit. Table 1.11 compares the model’s predictions of the targeted employer-side moments listed in Section 1.6.4 with those in the data. With the exception of the average wage of firms with less than 50 workers, our model fits all the other moments, including mean firm size, fraction of firms with less than 50 workers, and health insurance coverage rate (overall and by firm size).

Moments		Data	Model
Fraction of individuals who are unemployed and healthy		0.0314	0.0301
Fraction of individuals who are unemployed and unhealthy		0.0040	0.0021
Fraction of individuals who are employed, healthy and have health insurance		0.7009	0.7667
Fraction of individuals who are employed, unhealthy and have health insurance		0.0340	0.0319
Fraction of individuals who are employed, healthy and do not have health insurance		0.2156	0.1525
Fraction of individuals who are employed, unhealthy and do not have health insurance		0.0140	0.0167
Mean wage (\$10,000)		0.8538	0.8501
Mean wage with health insurance (\$10,000)		0.9240	0.8986
Mean wage without health insurance (\$10,000)		0.6187	0.6211
Mean medical expenditure (\$10,000)		0.0266	0.0253

Table 1.10: Worker-Side Moments in the Labor Market: Model vs. Data.

Moments	Data	Model
Mean firm size	19.92	18.5239
Fraction of firms less than 50 workers	0.93	0.9026
Mean size of firms that offer health insurance	30.08	27.0368
Mean size of firms that do not offer health insurance	6.95	7.2363
Health insurance coverage rate	0.56	0.5581
Health insurance coverage rate among firms with less than 50 workers	0.53	0.5200
Health insurance coverage rate among firms with more than 50 workers	0.95	0.9113
Average wages of firms with less than 50 workers	0.84	0.4129
Average wages of firms with more than 50 workers	0.92	0.9563

Table 1.11: Employer-Side Moments: Model vs. Data.

Figure 1.2 plots the size distribution of the firms both in the data and that implied by our model estimates. Figure 1.3 shows the size distributions of firms by health insurance offering status. Both figures show that our model is able to capture the size distribution of firms overall and by health insurance status reasonably well.

We should point out that even though our model qualitatively predicts the positive correlation between wage and firm size, it generates a much steeper relationship between them than what is in the data. Moreover, as we showed in Figure 1.1, our fit of workers' wage distribution conditional on health insurance status is still not ideal. Because firm productivity is positively correlated with wage offer in our model, in order to fit worker's wage distribution which is very dispersed, we need to have a relatively large variance of firm productivity. However, since firm size and wage are positively correlated in our model, a larger variance of firm productivity distribution leads to a steeper relationship between wage and firm size. The difficulty of simultaneously fitting firm size distribution from firm-side data and wage distribution from the worker-side data is known from Postel-Vinay and Robin (2002) and Cahuc, Postel-Vinay, and Robin (2006), who proposed to address the issue by introducing a wedge between worker's sampling distribution of firms and firm's productivity distribution.⁴⁰

1.8 Counterfactual Experiments

In this section, we use our estimated model to examine the impact of the Affordable Care Act, its various components, and alternative policy designs. For the ACA, we consider a stylized version which incorporates its main components as mentioned in the introduction: first, all individuals are required to have health insurance or have to pay a penalty; second, all firms with more than 50 workers are required to offer health insurance, or have to pay a penalty; third, we introduce a health insurance exchange where individuals

⁴⁰We believe that incorporating richer worker heterogeneity in their productivity, beside health status, may be a more important direction for further research.

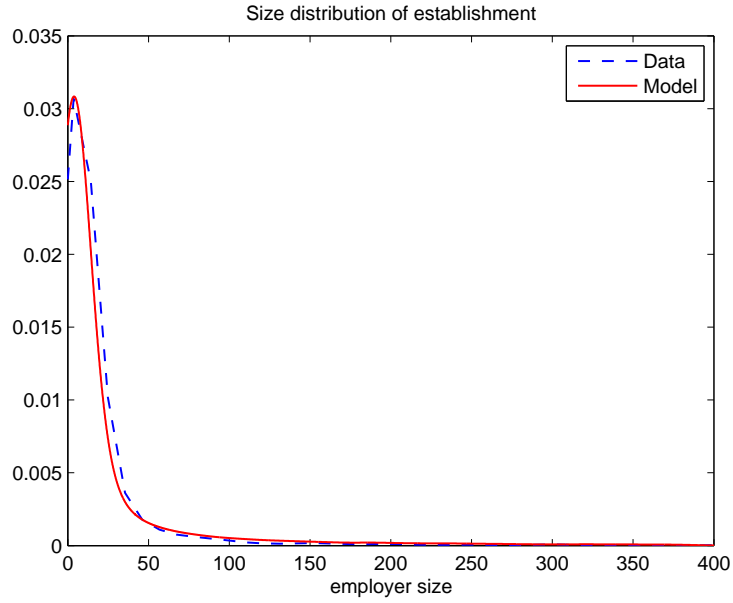


Figure 1.2: Size Distribution of Firms: Model vs. Data.

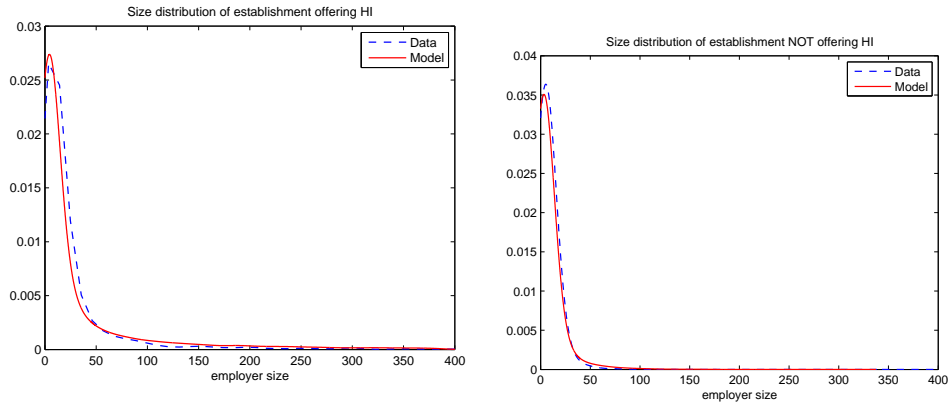


Figure 1.3: Size Distribution of Firms by Insurance Offering Status: Model vs. Data.

can purchase health insurance at community rated premium; fourth, the participants in health insurance exchange can obtain income-based subsidies.

The introduction of health insurance exchange represents a substantial departure from our benchmark model because premium in exchange needs to be endogenously determined. As a result, we will first describe how we extend and analyze our benchmark model to incorporate the health insurance exchange.

1.8.1 Model for the Counterfactual Experiments

We provide a brief explanation of the main changes in the economic environment for the model used in our counterfactual experiments.

The Main Change in Individuals' Environment. We now assume individuals who are not offered health insurance by their employers and those who are unemployed can purchase individual health insurance from the health insurance exchange. We assume that the insurance purchased from the exchange is identical to those offered by the employers in that it also fully insures medical expenditure risk. Thus in the extended model, an individual's insurance status x is defined as

$$x = \begin{cases} 0 & \text{if uninsured} \\ 1 & \text{if insured through employer} \\ 2 & \text{if insured through exchange.} \end{cases}$$

We assume that the effect on health for health insurance purchased from the exchange, denoted by π^2 analogously defined as (1.5), is identical to that for employer-sponsored health insurance, i.e., $\pi^2 = \pi^1$.

We also incorporate the premium subsidies to the individuals and penalties if uninsured into the model. Let $S(y, R^{EX})$ denote income based subsidies to an individual with income y who purchase health insurance from the exchange where R^{EX} is the premium in exchange, which is to be determined in equilibrium. Similarly, let $P_W(y)$ denote the penalty to individuals who remain uninsured, which also depends on income level.

Worker's problem. Under this extension, the expected flow utility $v_h(y, x)$ in the counterfactual is defined as:

$$v_h(y, x) = \begin{cases} \mathbb{E}_{\tilde{m}_h^0} u(T(y) - \tilde{m}_h^0 - P_W(y)) & \text{if } x = 0 \\ u(T(y)) & \text{if } x = 1 \\ u(T(y) + S(y, R^{EX}) - R^{EX}) & \text{if } x = 2 \end{cases} \quad (1.39)$$

The value function of an unemployed individual with health insurance status $x \in \{0, 2\}$ becomes

$$\frac{U_h(x)}{1 - \rho} = v_h(b, x) + \beta \mathbb{E}_{h'|h, x} \left[\begin{aligned} & \lambda_u \int \max \{V_{h'}(w, 1), U_{h'}(x_{h'}^*)\} dF(w, 1) \\ & + \lambda_u \int \max_{x' \in \{0, 2\}} \{V_{h'}(w, x'), U_{h'}(x_{h'}^*)\} dF(w, 0) \\ & + (1 - \lambda_u) U_{h'}(x_{h'}^*), \end{aligned} \right] \quad (1.40)$$

where

$$x_{h'}^* = \arg \max_{x' \in \{0, 2\}} U_{h'}(x'). \quad (1.41)$$

Similarly, the value function of an employed worker with health status h working on a job with insurance status (w, x) , $V_h(w, x)$, is as follows. If $x = 1$,

$$\begin{aligned} \frac{V_h(w, 1)}{1 - \rho} = v_h(w, 1) & + \beta \lambda_e \left\{ \begin{aligned} & (1 - \delta) \mathbb{E}_{h'|h, 1} \Omega_E^1(h', w, 1) \\ & + \delta \mathbb{E}_{h'|h, 1} \Omega_U(h') \end{aligned} \right\} \\ & + \beta(1 - \lambda_e) \\ & \times \left\{ \begin{aligned} & (1 - \delta) \mathbb{E}_{h'|h, 1} \max \left\{ \begin{aligned} & U_{h'}(x_{h'}^*), \\ & V_{h'}(w, 1) \end{aligned} \right\} \\ & + \delta \mathbb{E}_{h'|h, 1} U_{h'}(x_{h'}^*) \end{aligned} \right\}, \end{aligned} \quad (1.42)$$

and if $x \in \{0, 2\}$,

$$\begin{aligned} \frac{V_h(w, x)}{1 - \rho} = & v_h(w, x) + \beta \lambda_e \left\{ \begin{array}{l} (1 - \delta) \mathbf{E}_{h'|h,1} \Omega_E^1(h', w, x) \\ + \delta \mathbf{E}_{h'|h,1} \Omega_U(h') \end{array} \right\} \\ & + \beta(1 - \lambda_e) \left\{ \begin{array}{l} (1 - \delta) \mathbf{E}_{h'|h,1} \max \left\{ \begin{array}{l} U_{h'}(x_{h'}^*), \\ V_{h'}^0(w, x_{h'}^*(\tilde{w})) \end{array} \right\} \\ + \delta \mathbf{E}_{h'|h,1} U_{h'}(x_{h'}^*) \end{array} \right\}, \end{aligned} \quad (1.43)$$

where in (1.42) and (1.43),

$$\Omega_E^2(h', w, 1) = \left[\begin{array}{l} \int \max\{V_{h'}(\tilde{w}, 1), V_{h'}(w, x_{h'}^*(\tilde{w})), U_{h'}(x_{h'}^*)\} dF(\tilde{w}, 1) \\ + \int \max\{V_{h'}(\tilde{w}, x_{h'}^*(\tilde{w})), V_{h'}(w, x_{h'}^*(\tilde{w})), U_{h'}(x_{h'}^*)\} dF(\tilde{w}, 0) \end{array} \right] \quad (1.44)$$

$$\Omega_E^1(h', w, 1) = \left[\begin{array}{l} \int \max\{V_{h'}(\tilde{w}, 1), V_{h'}(w, 1), U_{h'}(x_{h'}^*)\} dF(\tilde{w}, 1) \\ + \int \max\{V_{h'}^0(\tilde{w}, x_{h'}^*(\tilde{w})), V_{h'}(w, 1), U_{h'}(x_{h'}^*)\} dF(\tilde{w}, 0) \end{array} \right] \quad (1.45)$$

$$\Omega_U(h') = \left[\begin{array}{l} \int \max\{U_{h'}(x_{h'}^*), V_{h'}(\tilde{w}, 1)\} dF(\tilde{w}, 1) \\ + \int \max\{U_{h'}(x_{h'}^*), V_{h'}(\tilde{w}, x_{h'}^*(\tilde{w}))\} dF(\tilde{w}, 0) \end{array} \right] \quad (1.46)$$

$$x_{h'}^*(\tilde{w}) = \arg \max_{x \in \{0, 2\}} V_{h'}(\tilde{w}, x). \quad (1.47)$$

We characterize the individuals' optimal job acceptance strategies, and their optimal decision regarding whether to purchase insurance from the exchanges when they are unemployed or when their employers do not offer health insurance similar to those for the benchmark model. We also characterize the steady state worker distribution among firms $\langle e_h^x, G_h^x(w) \rangle$ for $x \in \{0, 1, 2\}$ when the two additional terms, e_h^2 and $G_h^2(w)$, are now respectively the measure of employed workers with health status h who purchase insurance from the exchange, and the distribution of wages among them.⁴¹

⁴¹Details for the derivation of steady state employment distribution used in our counterfactual policy experiments are provided in Appendix A.5.3.

Firms' Problem. Firms with more than 50 workers now face a penalty if they do not offer health insurance. Let $P_E(n)$ denote the the amount of the penalty, which depends on the firm size n .

There are two important changes to the firms' problem. The first one is how firm size is determined. Because of the insurance exchange, some of their workforce may be insured even if they do not offer health insurance. Specifically, $n(w, 0)$, the size of firms not offering health insurance, becomes

$$n(w, 0) = \sum_{h \in \{U, H\}} n_h(w, 0) = \sum_{h \in \{U, H\}} \frac{e_h^0 g_h^0(w) + e_h^2 g_h^2(w)}{f(w, 0)},$$

and the expression for $n(w, 1)$ remains the same as before.

Second, because of the employer mandate, firm's profit maximization problem will change. It now becomes

$$\max\{\Pi_0(p), \Pi_1(p) + \sigma_f \epsilon\},$$

where:

$$\Pi_0(p) = \max_{w_0} \Pi(w_0, 0) \equiv \left[\begin{array}{l} (p - w_0) n_H(w_0, 0) \\ + (pd - w_0) n_U(w_0, 0) \end{array} \right] - P_E(n(w, 0)), \quad (1.48)$$

$$\Pi_1(p) = \max_{w_1} \Pi(w_1, 1) \equiv \left[\begin{array}{l} (p - w_1 - m_H) n_H(w_1, 1) \\ + (pd - w_1 - m_U) n_U(w_1, 1) \end{array} \right] - C \quad (1.49)$$

where the term $P_E(n(w, 0))$ in the expression for $\Pi_0(p)$ reflects the possible penalty to employers for not offering employer-sponsored health insurance to their workers.

Insurance Exchange. The premium in the insurance exchange, R^{EX} , is determined based on the average medical expenditures of all participants in the health insurance exchange, multiplied by $1 + \xi$, where $\xi > 0$ is loading factor for health insurance exchange;

specifically,

$$R^{EX} = (1 + \xi) \frac{\sum_{h \in \{H, U\}} m_h^2 [u_h^2 + \int e_h^2(w) g_h^2(w) dw]}{\sum_{h \in \{H, U\}} [u_h^2 + \int e_h^2(w) g_h^2(w) dw]} \quad (1.50)$$

where m_h^2 is expected medical expenditure of individual with health status h for individuals with insurances purchased from the exchange which, due to our assumption that the insurances in the exchange are identical to those from the firms, is exactly the same as m_h^1 described by (1.4); u_h^2 is the measure of unemployed workers participating insurance exchange with health status h ; and $e_h^2(w) g_h^2(w)$ is the density for employed workers not being offered health insurance from employers but participating insurance exchange with health status h .

The steady state equilibrium for the post-reform economy can be defined analogous to that for our benchmark model in Section 1.3.4 and is provided in Appendix A.5.1.

Numerical Algorithm to Solve the Equilibrium. We use numerical methods to solve the equilibrium. The basic iteration procedure to solve the equilibrium for the counterfactual environment remains the same as that described in the Appendix A.1, but there are two important changes. First, we need to find the fixed point of not only $(w_0(p), w_1(p), \Delta(p))$ but also R^{EX} , the premium in insurance exchange. Second, because the penalty associated with employer mandate depends on size of the firm, for example, the threshold under the ACA for firms to pay penalty if they do not offer health insurance is 50; as a result we need to modify the algorithm to allow for a potential mass point of employers just to the left of 50 when we derive optimal wage policy $w_0(p)$.⁴²

Finally, the establishment of the health insurance exchange with community rating may result in multiple equilibria under some counterfactual policy experiments. In our numerical simulations, we sometimes find multiple equilibria and we will report them.

⁴²The details of the modified numerical algorithm are provided in Appendix A.5.2.

1.8.2 Parameterization of the Counterfactual Policies

Before we conduct counterfactual experiments to evaluate the effect of ACA and its components, we need to address several issues regarding how to introduce the specifics of ACA provisions, such as penalty associated with individual mandate, employer mandate and the premium subsidies, into our model. First, we estimated our model using data sets in 1996, while the ACA policy parameters are chosen to suit the economy in 2011. However, the U.S. health care sector has very different growth rate than that of the overall GDP; in particular, there are substantial increases in medical care costs relative to GDP in the last 15 years. Thus we need to appropriately adjust the policy parameters in the ACA to make them more in line with the U.S. economy around 1996. Second, the amount of penalties and subsidies are defined as annual level, while our model period is four months. We simply divide all monetary units in the ACA by three to obtain the applicable number for a four-month period. Third, we need to decide on the magnitude of the loading factor ξ that appeared in (1.50) that is applicable in the insurance exchange. We calibrate ξ based on the ACA requirement that all insurance sold in the exchange must satisfy the ACA regulation that the medical loss ratio must be at least 80%. This implies that $\xi = 0.25$.⁴³

We now describe how we translate the ACA provisions for 2011 into applicable formulas for our 1996 economy.

Penalties Associated with Individual Mandate. The exact stipulation of the penalty in ACA if an individual does not show proof of insurance (from 2016 when the law is fully implemented) is that individuals without health insurance coverage pay a tax penalty of the greater of \$695 per year or 2.5% of the taxable income above the Tax Filing Threshold (TFT), which can be written as:

$$P_W^{ACA}(y) = \max \{0.025 \times (y - \text{TFT}_{2011}), \$695\} \quad (1.51)$$

⁴³The medical loss ratio is the ratio of the total claim costs the insurance company incurs to total insurance premium collected from participants. The medical loss ratio implied by (1.50) is simply $1/(1 + \xi)$, thus an 80% medical loss ratio corresponds to $\xi = 0.25$. ACA requires that $\xi \leq 0.25$.

where y is annual income.

We adjust the above formula in several dimensions. First, the \$695 amount is adjusted by the ratio of the 1996 Medical Care CPI (CPI_Med_1996) relative to the 2011 Medical Care CPI (CPI_Med_2011); this is appropriate if we believe that the amount \$695 is chosen to be proportional to the 2011 medical expenditures. We then multiply it by 1/3 to reflect our period-length of fourth months instead of a year. Second, we need to adjust the TFT_2011 by the ratio of 1996 CPI of all goods (CPI_All_1996) relative to the 2011 CPI of all goods (CPI_All_2011) and also multiply it by 1/3 to reflect that our income is the four-month income.⁴⁴ Finally, we need to adjust the percentage 2.5% by the differential growth rate of medical care and GDP, i.e., multiply it by the relative ratio of $\frac{\text{CPI_Med_1996}}{\text{CPI_All_1996}}$ and $\frac{\text{CPI_Med_2011}}{\text{CPI_All_2011}}$. With these adjustments, we specify the adjusted penalty associated with individual mandate appropriate for the 1996 economy as:

$$\begin{aligned}
 P_W(y) &= \max \left\{ \begin{array}{l} 0.025 \times \left(\frac{\text{CPI_Med_1996}}{\text{CPI_All_1996}} \right) / \left(\frac{\text{CPI_Med_2011}}{\text{CPI_All_2011}} \right) \\ \times \left(y - \frac{1}{3} \text{TFT_2011} \times \frac{\text{CPI_All_1996}}{\text{CPI_All_2011}} \right), \\ \frac{1}{3} \times \$695 \times \frac{\text{CPI_Med_1996}}{\text{CPI_Med_2011}} \end{array} \right\} \\
 &\approx \max \left\{ \frac{0.025}{1.42} \times (y - 2,323), \$119 \right\}, \tag{1.52}
 \end{aligned}$$

where y is four-month income in dollars.

Penalties Associated with Employer Mandate. ACA stipulates that employers with 50 or more full-time employees that do not offer health insurance coverage will be assessed each year a penalty of \$2,000 per full-time employee, excluding the first 30 employees from the assessment. That is,

$$P_E^{ACA}(n) = (n - 30) \times \$2,000. \tag{1.53}$$

We adjust the above formula by first scaling the \$2,000 per-worker penalty using the

⁴⁴We obtain CPI data for medical care and all goods both from Bureau of Labor Statistics website: <http://www.bls.gov/cpi/data.htm>.

ratio of the 1996 Medical Care CPI relative to the 2011 Medical Care CPI, and then multiply it by 1/3 to reflect our period-length of four months instead of a year, i.e.,

$$\begin{aligned}
 P_E(n) &= \frac{1}{3} \left[(n - 30) \times \$2,000 \times \frac{\text{CPI}_{\text{Med.1996}}}{\text{CPI}_{\text{Med.2011}}} \right] \\
 &= 342.45 (n - 30).
 \end{aligned} \tag{1.54}$$

Income-Based Premium Subsidies. ACA stipulates that premium subsidies for purchasing health insurance from the exchange are available if an individual's income is less than 400% of Federal Poverty Level (FPL), denoted by FPL400.⁴⁵ The premium subsidies are set on a sliding scale such that the premium contributions are limited to a certain percentage of income for specified income levels. If an individual's income is at 133% of the FPL, denoted by FPL133, premium subsidies will be provided so that the individual's contribution to the premium is equal to 3.5% of his income; when an individual's income is at FPL400, his premium contribution is set to be 9.5% of the income. When his income is below FPL133, he will receive insurance with zero premium contribution. If his income is above FPL400, he is no longer eligible for premium subsidies. Note that the premium support rule as described in ACA creates a discontinuity at FPL133: individuals with income below FPL133 receives free Medicaid, but those at or slightly above FPL133 have to contribute at least 3.5% of his income to health insurance purchase from the exchange. To avoid this discontinuity issue, we instead adopt a slightly modified premium support formula as follows:

$$S(y, R^{EX}) = \begin{cases} \max \left\{ \begin{aligned} &R^{EX} - \left[0.0350 + 0.060 \frac{(3y - \text{FPL133})}{\text{FPL400} - \text{FPL133}} \right] y \\ &\times \frac{\text{CPI}_{\text{Med.1996}}}{\text{CPI}_{\text{Med.2011}}}, 0 \end{aligned} \right\} & \text{if } y < \frac{\text{FPL400}}{3} \\ R^{EX} & \text{if unemployed} \\ 0, & \text{otherwise,} \end{cases} \tag{1.55}$$

when y is four-month income. According to (1.55) the individual contribution to insurance premium increases linearly from 3.5% of his income when his income is at 133% of the

⁴⁵We assume that FPL is defined as single person. In 1996, it is \$7,730 annually.

FPL to 9.5% when his income is at 400% of the FPL.

1.8.3 Main Result

In Table 1.12, we report results from several counterfactual policy experiments and contrast the outcomes under these counterfactual policies with the benchmark.

Benchmark. In Column (1) of Table 1.12, we report in Panel A predictions related to the firm side using our baseline model using the parameter estimates reported in Tables 1.7 and 1.8. It shows that our benchmark model predicts that 55.81% of the *active* firms offer health insurance to their workers, but the health insurance offering rate is 91.13% if the firm size is more than 50 and only 52.00% if it has fewer than 50 workers; moreover, our model predicts that 90.26% of the firms have fewer than 50 workers. We find that 98.72% of the firms are active in the benchmark environment; and the average labor productivity, taking into account the productivity loss from unhealthy workers, is \$11,300 per four months.⁴⁶

In Panel B, we report predictions related to worker side. We find that our model predicts that in the benchmark environment, 20.12% of the population would have no health insurance; 3.22% would be unemployed. We also find that the average wage of all workers is about \$8,501 per four months, but it is about \$8,986 for workers with health insurance from their employers and \$6,211 for those without. We find that the fraction of healthy workers in the economy overall is 94.94%, but it is 96% among those insured by ESHI is 96.00% in contrast to 90.17% among those who are uninsured. We find that 61.42% of the employed workers work in firms with 50 or more workers. Finally, we find that the consumption equivalent valuation (CEV) of workers' lifetime welfare is about \$6,152 in the benchmark economy.

⁴⁶Recall that in our model, some low-productivity firms would not be able to attract any workers and they are considered *non-active* firms. The set of non-active firms is affected by the counterfactual policies. Thus our model allows for an extensive margin on the firm side.

	Benchmark (1)	ACA (2)	EX+Sub+EM (3)	EX+Sub+IM (4)
Panel A: Effects on the Firm Side				
Frac. of firms offering HI	0.5581	0.5486	0.5494	0.5531
...if firm size is less than 50	0.5200	0.5039	0.5012	0.5111
...if firm size is 50 or more	0.9113	0.9993	0.9988	0.9506
Frac. of firms with less than 50 workers	0.9026	0.9097	0.9031	0.9043
Average labor productivity	1.1300	1.1299	1.1309	1.1349
Firm's profit	0.4717	0.4937	0.4809	0.4765
Frac. of firms in operation	0.9872	0.9872	0.9872	0.9872
Panel B: Effects on the Worker Side				
Uninsured rate	0.2012	0.0727	0.1218	0.0644
Frac. of emp. workers with HI from ESHI	0.8253	0.8284	0.8259	0.8364
Frac. of emp. workers with HI from EX	-	0.0965	0.0482	0.0971
Average wage	0.8501	0.8449	0.8482	0.8526
... with health insurance	0.8986	0.8934	0.9002	0.9019
... without health insurance	0.6211	0.6132	0.6021	0.6014
Unemployment rate	0.0322	0.0320	0.0322	0.0322
Frac. of healthy workers	0.9494	0.9592	0.9558	0.9598
... among uninsured	0.9017	1.0000	1.0000	1.0000
... among insured through ESHI	0.9600	0.9636	0.9628	0.9636
... among insured through EX	-	0.8890	0.7170	0.8984
Frac. of emp. workers in firms with 50+ workers	0.6142	0.5940	0.6129	0.6096
Average worker utility (CEV)	0.6152	0.6133	0.6164	0.6184
Panel C: Effects on Expenditures				
Average tax expenditure to ESHI	0.0084	0.0083	0.0083	0.0084
Subsidies to exchange purchases	-	0.0034	0.0038	0.0032
Revenue from penalties	-	0.0010	0.00002	0.0009
Average health expenditure	0.0253	0.0273	0.0272	0.0273
Average premium in ESHI	0.0306	0.0301	0.0302	0.0300
Premium in exchange	-	0.0439	0.0595	0.0427

Table 1.12: Counterfactual Policy Experiments: Evaluation of the ACA and its Two Variations.

In Panel C, we report statistics related to expenditures. We find that the tax exemption of ESHI premium leads to a tax expenditure of about \$84 *per capita* every four months. We also find that the per capita four-month medical expenditure is about \$253; and the average insurance premium *among those insured* by ESHI is about \$306 every four months.

Affordable Care Act. In Column (2), we report the simulation results when we introduce the ACA, including insurance exchange (EX), Individual Mandate (IM), Employer Mandate (EM) and Premium Subsidies (Sub), as parameterized in Section 1.8.2.

The important finding from Column (2) is that, under the ACA, our model predicts that there would be significant reduction in the uninsured rate relative to the benchmark: the uninsured rate under ACA is predicted to be about 7.27% in contrast to 20.12% under the benchmark. While this is certainly a significant reduction in the uninsured rate, it should be noted that it is still far from universal coverage. It is also interesting to note that the 7.27% uninsured population are all healthy and are employed workers who belong to middle income level so that premium subsidies to them are relatively small. They apparently prefer paying the penalty associated with not meeting the individual mandate than purchasing insurance from the exchange.

A major reason for the reduction of uninsured rate under the ACA is that 9.65% of the employed workers purchase insurance from the exchange. This, coupled with the fact that in our counterfactual model all unemployed workers (3.20%) will purchase from the exchange as well because they will receive the insurance for free due to their low income at unemployment, accounts for almost all the reductions in the uninsured rate. Interestingly, we find in Panel A that the ACA would slightly *reduce* the fraction of active firms that offer health insurance from 55.81% to 54.86%. However, we find that, as a result of employer mandate penalty for firms with more than 50 workers, such firms significantly increase the probability that they offer health insurance to their workers from 91.13% in the benchmark to 99.93% under the ACA; in contrast, the probability that firms with less

than 50 workers offer health insurance is reduced from 52% in the benchmark to 50.39% under the ACA.

We find that ACA leads to a reduction in average labor productivity and worker wages. This is related to the changes in the size distribution of firms: fewer workers are employed in larger, more productive, firms under the ACA. In Panel B we also report that ACA improves the health of the workers overall: the fraction of healthy workers increases from 94.94% to 95.92%. We find that the fraction of employed workers in firms with 50 or more workers decreases from 61.42% in the benchmark to 59.4% under the ACA. We also find that workers' CEV is \$6,133 under the ACA, lower than the benchmark of \$6,152.

In Panel C, we report the effect of ACA on expenditure related variables. First, because a smaller fraction of firms offer health insurance under the ACA, and because health insurance premium is not subject to income taxation, the tax expenditure due to ESHI premium exemption is reduced somewhat from \$84 per capita in the benchmark to \$83 under ACA. The government also incurs on average \$34 per capita subsidies to health insurance purchases from the exchange; however, the revenue from penalties from individuals who decide to go without insurance or firms with 50 or more workers which do not offer health insurance is also about \$10 per capita. Because of the reduction in the uninsured rate, there is about a 8% increase in the average medical expenditure.

Interestingly, the average four-month premium in ESHI is slightly reduced from \$306 in the benchmark to \$301 under the ACA; this reduction is partly due to the improved health of the population under the ACA. We also find that the four-month premium in the exchange is about \$439, which is significantly higher than the average ESHI premium due to the severe adverse selection problem in the exchange.

ACA without the Individual Mandate. In Column (3), we report simulation results from a hypothetical environment of ACA without the individual mandate, i.e. only EX, Sub and EM components of ACA are implemented. This would correspond to the case had the Supreme Court ruled against the constitutionality of the individual mandate.

Surprisingly, we find that ACA without the individual mandate would also have achieved significant reduction in the uninsured rate. In Panel B, we show that the uninsured rate under “EX+Sub+EM” would be about 12.18%, which is 4.91 percentage points higher than under the ACA, but still represent close to 39% reduction from the 20.12% uninsured rate predicted in the benchmark. The reason for the sizeable reduction in the uninsured rate despite the absence of individual mandate is the premium subsidies. Individuals are risk averse so they would like to purchase insurance if the amount of premium they need to pay out of pocket is sufficiently small, which is true for many workers in low-wage firms that do not offer health insurance. Those workers who work in firms with medium-wages but do not offer health insurance turn out to be those workers who decide to pay the penalty and go without health insurance, if they are healthy. Notice that the fraction of employed workers who purchase health insurance from the exchange is significantly lower in “EX+Sub+EM” (Column 3) than under the full ACA (Column 2): only 4.82% of the employed workers have insurance from the exchange in “EX+Sub+EM” while 9.65% do so under the ACA. This further intensifies the adverse selection problem in the exchange, leading to a substantial increases in the premium in the exchange (from \$439 under the ACA to \$595 in “EX+Sub+EM”). Interestingly, because of the premium increase, the per capita premium subsidy is \$38, higher than the \$34 per capita amount under the ACA, despite the decrease in the number of participants in exchange.

Note that the average worker lifetime utility in CEV is \$6,164 under ACA without individual mandate, which is higher than that under the full ACA and that under the benchmark.

ACA without Employer Mandate. In Column (4), we report the result from a hypothetical environment of ACA *without* the employer mandate. This would roughly correspond to a health care system in the spirit of what is implemented in Netherlands and Switzerland where individuals are mandated to purchase insurance from the private insurance market, employers are not required to offer health insurance to their workers,

and government subsidizes health care for the poor on a graduated basis.⁴⁷

We find that, surprisingly, such a system without employer mandate would actually lead to an even lower uninsured rate than the full version of ACA. We find that the uninsured rate under this “EX+Sub+IM” system would be about 6.44%, lower than the 7.27% uninsured rate predicted under the full ACA. Because there is no size-dependent employer mandates, firms do not have to reduce their size to less than 50 in order to avoid paying penalties.⁴⁸ Indeed, the fraction of employed workers in firms with 50 or more workers is 60.96% under this “EX+Sub+IM” system, somewhat higher than 59.40% under the ACA.

We also find that more individuals obtain health insurance from insurance exchange. The premium in the exchange decreases from \$439 under the full ACA to \$427 under this “EX+Sub+IM” system. Furthermore, the workers’ average lifetime utility measured by CEV is higher than under the full ACA.

To understand why the employer mandate on large firms might increase the uninsured rate, it is important to recognize that the employer mandate has the following two effects. First, since the mandate increases the health insurance offer rate of large firms, it improves the overall health composition of the population. However, through labor market transitions, the workers hired at large firms may return to unemployment (due to exogenous separation) and subsequently find jobs at small firms that may not offer health insurance. Since these workers are healthier, the health composition of potential entrants in insurance exchange improves, which alleviates the adverse selection problem in the insurance exchange. This externality from employer mandate contributes to an increase of the participants in exchange, lowering uninsured rate.

However, there is a second countervailing effect. The size-dependent employer man-

⁴⁷Strictly speaking, the Swiss health care system expressly forbids employers from providing basic social health insurance as a benefit of employment, though employers can provide supplemental health insurance to their workers. See Fijolek (2012, p.8) for a description.

⁴⁸Indeed, we find a probability mass of firms (2.12%) with size just below the mandate threshold of 50 under the ACA. The probability mass disappears under the “EX+Sub+IM” system.

date indeed increases the health insurance offering by large and high-productivity firms. However, small firms' incentive to offer health insurance may be reduced. The reason is that small firms anticipate that their workers will benefit less from being offered health insurance. In our model, workers demand health insurance because it not only provides insurance against the health expenditure shocks in the current period, but also it reduces future health expenditure risks since health insurance improves the realization of future health. If these workers anticipate that they will move to high-productivity firms offering health insurance with higher probability, the incentives to purchase health insurance in the current period may be lower. This channel may also reduce the incentives of healthy uninsured workers to participate in insurance exchange. This phenomena, known as *dynamic inefficiency* in the literature of insurance markets, may therefore lead small firms not to offer health insurance, and also lead workers not offered insurance by their employers to forgo purchasing health insurance from the exchange. Both can lead to higher uninsured rate. Therefore, the quantitative impact of the size-dependent employer mandate on overall uninsured rate is nontrivial and can be highly dependent on policy parameters.

Finally, workers' lifetime utility is higher under "EX+Sub+IM" system than under the full ACA for the following reason. Because of the employer mandate in the ACA, large firms not offering health insurance need to pay penalty. One option to finance the penalties is to reduce its wage offer. It then has an equilibrium impact on firms offering health insurance so that they can reduce wage offer to attract workers. Indeed, in Columns (3) and (4) we see that the average wages of workers with and without health insurance are both lower under the ACA than under the "EX+Sub+IM" system. This overall wage decline under the ACA is responsible for the lower worker utility relative to the "EX+Sub+IM" system.

	EX	EX+Sub	EX+IM		
	(1)	(2)	(3)	(4)	(5)
Panel A: Effects on the Firm Side					
Frac. of firms offering HI	0.5591	0.5610	0.5476	0.5683	0.5721
...if firm size is less than 50	0.5210	0.5202	0.5105	0.5272	0.5294
...if firm size is 50 or more	0.9146	0.9498	0.8919	0.9568	0.9795
Frac. of firms with less than 50 workers	0.9031	0.9049	0.9028	0.9044	0.9052
Average labor productivity	1.1297	1.1300	1.1323	1.1308	1.1312
Firm's profit	0.4702	0.4712	0.4725	0.4717	0.4732
Frac. of firms in operation	0.9919	0.9919	0.9872	0.9919	0.9919
Panel B: Effects on the Worker Side					
Uninsured rate	0.2002	0.1919	0.1320	0.1806	0.1690
Frac. of emp. workers with HI from ESHI	0.8261	0.8347	0.8082	0.8464	0.8583
Frac. of emp. workers with HI from EX	0.0000	0.0000	0.0554	0.0000	0.0000
Average wage	0.8495	0.8490	0.8529	0.8490	0.8483
... with health insurance	0.8989	0.8983	0.9037	0.8950	0.8930
... without health insurance	0.6151	0.6011	0.6391	0.5952	0.5769
Unemployment rate	0.0320	0.0320	0.0322	0.0318	0.0318
Frac. of healthy workers	0.9495	0.9500	0.9550	0.9509	0.9518
... among uninsured	0.9020	0.9022	1.0000	0.9030	0.9037
... among insured through ESHI	0.9600	0.9601	0.9628	0.9602	0.9603
... among insured through exchange	-	-	0.7280	-	-
Frac. of emp. workers in firms with 50+ workers	0.6139	0.6091	0.6136	0.6102	0.6081
Average worker utility (CEV)	0.6149	0.6149	0.6188	0.6127	0.6127
Panel C: Effects on Expenditures					
Tax expenditure to ESHI	0.0084	0.0085	0.0082	0.0086	0.0087
Subsidies to exchange purchases	-	-	0.0042	-	-
Tax revenue from penalties	-	-	-	0.0023	0.0021
Average health expenditure	0.0253	0.0254	0.0272	0.0255	0.0257
Average premium in ESHI	0.0306	0.0306	0.0302	0.0306	0.0305
Premium in exchange	0.1156	0.1997	0.0598	0.1156	0.1997

Table 1.13: Counterfactual Policy Experiments: Evaluation of Various Components of the ACA (1).

	EX+EM		EX+IM+EM	
	(1)	(2)	(3)	(4)
Panel A: Effects on the Firm Side				
Frac. of firms offering HI	0.5613	0.5618	0.5708	0.5704
...if firm size is less than 50	0.5139	0.5143	0.5236	0.5233
...if firm size is 50 or more	0.9996	0.9996	0.9999	0.9999
Frac. of firms with less than 50 workers	0.9024	0.9021	0.9010	0.9012
Average labor productivity	1.1279	1.1280	1.1282	1.1283
Firm's profit	0.4781	0.4783	0.4802	0.4799
Frac. of firms in operation	0.9919	0.9919	0.9919	0.9919
Panel B: Effects on the Worker Side				
Uninsured rate	0.1882	0.1867	0.1714	0.1728
Frac. of emp. workers with HI from ESHI	0.8383	0.8402	0.8559	0.8543
Frac. of emp. workers with HI from EX	0.0000	0.0000	0.0000	0.0000
Average wage	0.8453	0.8453	0.8448	0.8449
... with health insurance	0.8940	0.8937	0.8885	0.8887
... without health insurance	0.5943	0.5919	0.5869	0.5898
Unemployment rate	0.0320	0.0320	0.0318	0.0318
Frac. of healthy workers	0.9504	0.9505	0.9516	0.9515
... among uninsured	0.9021	0.9021	0.9030	0.9032
... among insured through ESHI	0.9602	0.9602	0.9604	0.9603
... among insured through exchange	-	-	-	-
Frac. of emp. workers in firms with 50+ workers	0.6160	0.6167	0.6194	0.6188
Average worker utility (CEV)	0.6127	0.6128	0.6106	0.6106
Panel C: Effects on Expenditures				
Tax expenditure to ESHI	0.0085	0.0085	0.0087	0.0087
Subsidies to exchange purchases	-	-	-	-
Tax revenue from penalties	6.4330E-6	6.9414E-6	0.0021	0.0022
Average health expenditure	0.0254	0.0254	0.0256	0.0256
Average premium in ESHI	0.0306	0.0306	0.0305	0.0305
Premium in exchange	0.0900	0.1997	0.1208	0.1997

Table 1.14: Counterfactual Policy Experiments: Evaluation of Various Components of the ACA (2).

1.8.4 Assessing the Effects of Components of the ACA

In Tables 1.13 and 1.14, we report the results from several counterfactual experiments that are aimed to understand the effects of several components of the ACA. In Columns (1) and (2) in Table 1.13, we report two equilibria when we introduce only the insurance exchange to the benchmark economy. In both equilibria, having an exchange alone does little to the uninsured rate in equilibrium: the equilibrium uninsured rate under this counterfactual is only slightly lower relative to the benchmark economy. In fact, the exchange will not have any participants at all due to the adverse selection problem. However, the presence of the exchange still causes small changes to the labor market, both on the firm side and on the worker side, because the exchange affects the outside options of the workers' and thus affects the firms' decisions regarding wage and health insurance offering decisions in equilibrium.

In Column (3) in Table 1.13, we report the results when we introduce health insurance exchange *and* health insurance premium subsidies. It shows that the introduction of premium subsidies and exchange leads to a sizable reduction in the uninsured rate to about 13.20%. The exchange is quite active with all the unemployed and 5.54% of the employed workers purchasing insurance from the exchange. However, without employer mandate, the introduction of exchange and premium subsidies also lead to a reduction in the probabilities of firms offering health insurance to their workers.

In Columns (4) and (5) in Table 1.13, we report the two equilibria when we introduce health insurance exchange and individual mandate. As in the case of Columns (1) and (2), the exchange will not have any participants. This indicates that the proposed individual mandate alone is not sufficiently large enough to solve adverse selection problem in the insurance exchange. Instead, the individual mandate leads more employers to offer health insurance. As a result, uninsured rate is 18.06% in Column (4) and 16.90% in Column (5). Because there still exists the sizable fraction of uninsured, introducing individual mandate lowers worker utility relative to an economy with insurance exchange only.

In Columns (1) and (2) in Table 1.14, we report the results when we introduce the health insurance exchange and employer mandate into the benchmark economy. We again find that the exchange is not active. There is a reduction of the uninsured rate, from 20.12% in the benchmark to 18.82% in Column (1) and to 18.67% in Column (2), but the declines of the uninsured rate are mostly due to the increased probability of offering health insurance by firms with 50 or more workers.

In Columns (3) and (4) in Table 1.14, we report the results when we introduce the ACA without that the income-based premium subsidies. Relative to the full ACA results reported in Column 2 of Table 2.17, the uninsured rate is about twice as large, 17.14% in Column (3) and 17.28% in Column (4). Moreover, worker utility is much lower than the experiments described from Column (1) to Column (5) in Table 1.13 and from Column (1) and Column (2) in Table 1.14. These results demonstrate that the proposed premium subsidies are crucial to solve adverse selection problem in the insurance exchange and contribute importantly to the substantial reduction of uninsured rate achieved under the full ACA.

1.8.5 Role of Tax Exemption for ESHI Premium

In this section, we describe the results from counterfactual experiments where the tax exemption status of employer-sponsored health insurance premium is eliminated, both under the benchmark model and under the ACA. We are interested in these counterfactual experiments because, given the growing federal deficits in the United States, reducing tax expenditures - tax exemption for ESHI premium being one of the major tax expenditure categories – has been mentioned in several prominent reports.⁴⁹

Columns (1) and (3) of Table 1.15 report the same simulation results for the benchmark and the ACA as reported in Table 1.12 under the current tax exemption status for ESHI premium. In Column (2), we remove the tax exemption for ESHI under the benchmark

⁴⁹See, for example, National Commission on Fiscal Responsibility and Reform (2010).

	Benchmark		ACA	
	Exempt (1)	No Exempt (2)	Exempt (3)	No Exempt (4)
Panel A: Effects on the Firm Side				
Frac. of firms offering HI	0.5581	0.5419	0.5486	0.5290
...if firm size is less than 50	0.5200	0.5053	0.5039	0.4889
...if firm size is 50 or more	0.9113	0.8834	0.9993	0.9837
Frac. of firms with less than 50 workers	0.9026	0.9031	0.9097	0.9188
Average labor productivity	1.1300	1.1287	1.1299	1.1133
Firm's profit	0.4717	0.4692	0.4937	0.4995
Frac. of firms in operation	0.9872	0.9872	0.9872	0.9919
Panel B: Effects on the Worker Side				
Uninsured rate	0.2012	0.2339	0.0727	0.0915
Frac. of emp. workers with HI from ESHI	0.8253	0.7916	0.8284	0.7871
Frac. of emp. workers with HI from EX	-	-	0.0965	0.1170
Average wage	0.8501	0.8510	0.8449	0.8401
... with health insurance	0.8986	0.9072	0.8934	0.8983
... without health insurance	0.6211	0.6374	0.6132	0.6336
Unemployment rate	0.0322	0.0320	0.0320	0.0353
Frac. of healthy workers	0.9494	0.9470	0.9592	0.9557
... among uninsured	0.9017	0.9007	1.0000	1.0000
... among insured through ESHI	0.9600	0.9597	0.9636	0.9627
... among insured through exchange	-	-	0.8890	0.8798
Frac. of emp. workers in firms with 50+ workers	0.6142	0.6127	0.5940	0.5698
Average worker utility (CEV)	0.6152	0.6077	0.6133	0.6028
Panel C: Effects on Expenditures				
Tax expenditure to ESHI	0.0084	-	0.0083	-
Subsidies to exchange purchases	-	-	0.0034	0.0044
Tax revenue from penalties	-	-	0.0010	0.0014
Average health expenditure	0.0253	0.0249	0.0273	0.0275
Average premium in ESHI	0.0306	0.0307	0.0301	0.0302
Premium in Exchange	-	-	0.0439	0.0466

Table 1.15: Counterfactual Policy Experiments: Evaluating the Effects of Eliminating the Tax Exemption for EHI Premium under the Benchmark and the ACA.

economy. We find that removing the tax exemption increase the uninsured rate from 20.12% to 23.39%. It leads to an increase in average wage for workers, and a deterioration of workers' health.

In Column (4), we remove the tax exemption for ESHI under ACA. We find that removing the tax exemption increase the uninsured rate from 7.27% to 9.15%. These differences are driven by the fact that under the ACA, workers who do not receive health insurance from their employers would have to purchase health insurance from the exchange or pay a penalty. Overall, our findings show that eliminating the tax exemption status for ESHI premium will increase the uninsured rate, both under the benchmark and under the ACA, but the impact is not sufficient to lead to the collapse of the ESHI.

In fact, in Table 1.15, we report that even without the tax exemption for ESHI premium, a substantial fraction of the firms will choose to offer health insurance to their workers, both in the benchmark economy and under the ACA. In the benchmark economy, we find that 54.19% of the firms will offer health insurance to their workers when ESHI premium is no longer exempt from income taxation; this is only slightly lower than 55.58% when ESHI premium is exempt from income taxation. Similarly, 52.90% of the firms will offer health insurance to their workers under the ACA when ESHI premium is not exempt from income taxation, which is again only slightly lower than 54.86% with exemption. There are several reasons that firms have strong incentives to offer health insurance to their workers in our economy. First, workers are risk averse and firms are risk neutral; thus firms can enjoy the risk premium by offering health insurance to their workers. Second, health insurance improves health and healthy workers are more productive. Thus firms, particularly those with higher productivity, will have incentives to offer health insurance to their workers so that their workforce will be healthier and thus more productive. This mechanism is illustrated in Table 1.2.

1.8.6 Other Counterfactual Experiments

In this section, we describe the results from several additional counterfactual policy experiments.

The Role of Individual Mandate Penalty

Recall that in Columns (4) and (5) in Table 1.13 we considered a counterfactual economy where we introduce only insurance exchange and individual mandate to the benchmark economy and we found that the uninsured rate would be 16.90% or 18.82% depending on equilibrium selection. Here we investigate how high the penalties for not having health insurance need to be in order to achieve universal coverage with only exchange and individual mandate. This experiment allows us to understand the impact of the strictness of individual mandate on the uninsured rate. To do so, we modify the formula of individual mandate to:

$$P_W(y) = \max \left\{ \begin{array}{l} 0.025 \times \frac{\left(\frac{\text{CPI_Med_1996}}{\text{CPI_All_1996}} \right)}{\left(\frac{\text{CPI_Med_2011}}{\text{CPI_All_2011}} \right)} \\ \times \left(y - \frac{1}{3} \text{TFT_2011} \times \frac{\text{CPI_All_1996}}{\text{CPI_All_2011}} \right), P^* \end{array} \right\}$$

where P^* is the amount of penalty under which the economy achieves universal coverage. Note that in this economy we might encounter multiple equilibria; and in this exercise, we search for the minimum level of P^* that yields a *unique* equilibrium with full coverage.

We find that if P^* is set to

$$\begin{aligned} P^* &= 15 \times \frac{1}{3} \times \$695 \times \frac{\text{CPI_Med_1996}}{\text{CPI_Med_2011}} \\ &\approx \$1785, \end{aligned}$$

which is 15 times as large as the proposed penalty in the ACA, the economy achieves the universal coverage.

The main results for this counterfactual economy (IM + EX) with an individual man-

	EX+ IM (1)	Sub+EX +EM (2)	MA Reform (3)	No ESHI	
				Sub+EX +IM (4)	Sub+EX +IM (5)
Panel A: Effects on the Firm Side					
Frac. of firms offering HI	0.5293	0.5286	0.5509	-	-
...if firm size is less than 50	0.5071	0.4779	0.5083	-	-
...if firm size is 50 or more	0.7400	0.9972	0.9529	-	-
Frac. of firms with less than 50 workers	0.9047	0.9025	0.9041	0.5192	0.6958
Average labor productivity	1.1388	1.1212	1.1280	1.2185	1.2169
Firm's profit	0.4689	0.4995	0.4893	3.0564	1.6809
Frac. of firms in operation	0.9919	0.9898	0.9872	0.2065	0.3264
Panel B: Effects on the Worker Side					
Uninsured rate	0.0000	0.0752	0.0529	0.5896	0.0000
Frac. of emp. workers with HI from ESHI	0.6948	0.8061	0.8470	-	-
Frac. of emp. workers with HI from EX	0.3052	0.1162	0.0983	0.2929	1.0000
Average wage	0.8633	0.8426	0.8449	0.8179	0.8909
... with health insurance	0.9019	0.9006	0.8949	-	-
... without health insurance	0.7755	0.6121	0.5787	-	0.8909
Unemployment rate	0.0318	0.0320	0.0322	0.1662	0.1090
Frac. of healthy workers	0.9644	0.9591	0.9606	0.9229	0.9644
... among uninsured	-	1.0000	1.0000	1.0000	-
... among insured through ESHI	0.9643	0.9636	0.9638	-	-
... among insured through EX	0.9644	0.8994	0.9106	0.7146	0.9643
Frac. of emp. workers in firms with 50+ workers	0.6093	0.6150	0.6103	0.7827	0.7080
Average worker utility (CEV)	0.6175	0.6142	0.6146	0.4866	0.5630
Panel C: Effects on Expenditures					
Average tax expenditure to ESHI	0.0070	0.0081	0.0085	-	-
Subsidies to exchange purchases	-	0.0051	0.0029	0.0179	0.0057
Revenue from penalties	0.0000	0.00004	0.0011	0.0096	0.0000
Average health expenditure	0.0273	0.0273	0.0273	0.0269	0.0273
Average premium in ESHI	0.0304	0.0300	0.0300	-	-
Premium in exchange	0.0341	0.0429	0.0411	0.0603	0.0342

Table 1.16: Counterfactual Policy Experiments: Evaluation of Alternative Policy Arrangements.

Notes: Column (1) is the case where IM leads to zero uninsured rate; Column (2) is the case where there is no IM but flat income Sub; Column (3) is the MA reform case; Column (4) and (5) are the cases of no ESHI.

date penalty 15 times as large as in the ACA are reported in Column (1) of Table 1.16. We find a substantial reduction of coverage rate by large firms relative to the benchmark economy; instead, 30.52% of employed workers obtain health insurance from insurance exchange, which is much higher than under the full ACA. While the premium in the exchange is still higher than the average ESHI premium, some firms choose not offer health insurance because the strong presence of the insurance exchange increases the workers' outside option and thus limits the degree in which firms can extract the risk premium from workers' wages by offering health insurance. Moreover, because all individuals are insured, the fraction of healthy workers increases, and as a result the average labor productivity increases. Also, because workers are now all insured and their wages are higher, worker utility increases relative to the benchmark economy.

ACA without the Individual Mandate: How Much Additional Subsidies Are Needed?

One of the major criticisms to the ACA has been the presence of individual mandate. As we demonstrate in Columns (2) and (3) of Table 1.12, eliminating the individual mandate from the ACA leads to an uninsured rate that is 4.91 percentage points higher than that under the ACA. An interesting question is, how much additional premium subsidies are needed to achieve the uninsured rate similar to under the ACA? To answer the question, we consider that, in addition to the proposed premium subsidies, the government provides a flat subsidy to any employed workers who choose to participate in the insurance exchange.

We find that if the government offers a four-month subsidy of \$135, the uninsured rate decreases to 7.52%, close to that under the full ACA. However, the per capita subsidies to exchange purchases increase from \$38 to \$51, which is higher than the \$34 amount under the ACA. All the results are in Column (2) of Table 1.16.

Massachusetts Health Care Reform

Next, we examine Massachusetts (MA) Health Care Reform implemented in 2006. It is well known that the ACA is based on the MA reform and there are strong similarities between them. However, employer mandate is implemented somewhat differently from the ACA, so is the premium subsidy. In this section, we investigate what happens if the government follows exactly the same reform as that in the MA.

To parametrize the MA reform, we consider the following stylized version of the reform as described in Kolstad and Kowalski (2012c).

Individual Mandates. We assume that it is the same as the ACA.⁵⁰

Employer Mandates. Firms with more than 10 workers are subject to the penalty tax if they do not offer health insurance. The amount of penalty is equal to \$295 times the number of full time employees. By using the same argument for the parameterization in the ACA, we parameterize it as follows: for firms with more than 10 workers, the amount of penalty, $P_E^{MA}(n)$, is

$$P_E^{MA}(n) = \frac{1}{3} \left[n \times \$295 \times \frac{\text{CPI_Med_1996}}{\text{CPI_Med_2011}} \right].$$

Premium Subsidies to Exchange Participants. As in the ACA, the income based subsidies are available to individuals participating in insurance exchange. However, it is available to individuals whose income is less than 300% FPL (FPL300). Therefore, we

⁵⁰Note that the actual policy taken in MA was that penalty is equal to a half of premium of the least generous qualifying plan.

parameterize it as:

$$S(y, R^{EX}) = \begin{cases} \max \left\{ \begin{aligned} &R^{EX} - \left[0.0350 + 0.060 \frac{(3y - \text{FPL133})}{\text{FPL300} - \text{FPL133}} \right] y \\ &\times \frac{\text{CPLMed}_{1996}}{\text{CPLMed}_{2011}}, 0 \end{aligned} \right\} & \text{if } y < \frac{\text{FPL300}}{3} \\ R^{EX} & \text{if unemployed} \\ 0, & \text{otherwise,} \end{cases} \quad (1.56)$$

Findings. The main result is Column (3) in Table 1.16. We find that the uninsured rate is 5.29% under the MA reform, which is lower than the 7.27% under the ACA. Because employer mandate is imposed more uniformly across firms, the positive externalities from the health improvement of workers insured by their firms are larger than under the ACA, which leads to lower premium in the exchange (\$411 under the MA reform vs. \$439 under the ACA), as well as that in the ESHI market. Both contributed to a lower uninsured rate. Worker utility in CEV is \$6,146, again higher than under the ACA. These findings are qualitatively consistent with Kolstad and Kowalski (2012c).

No Employer Sponsored Health Insurance Market

In Columns (4) and (5), we investigate the effects of *eliminating* employer sponsored health insurance market. In Column (4), we report the results from an experiment where we *prohibit* firms from offering ESHI, but instead we introduce the health insurance exchange, individual mandate and premium subsidies as stipulated in the ACA. We find a rather drastic change in the outcomes. Only 20.65% of firms are in operation, compared with 98.72% in the benchmark economy. Unemployment rate also increase to 16.62% from 3.22% in the benchmark economy. Uninsured rate is 58.96%, which is more than twice as large as the one in the benchmark economy. Insurance premium in exchange is \$603, higher than under the full ACA. It thus indicates that if there is no employer sponsored health insurance market, the proposed subsidies and individual mandate are not large enough to solve adverse selection problem in insurance exchange. This result suggests

that ESHI allows low productive firms to be active in the market because they can exploit the workers' risk premium.

The next question we investigate is, is it possible to achieve universal coverage when ESHI is eliminated by increasing the individual mandate penalty amount? The results are reported in Column (5). We find that, if we set the penalty of not having insurance to be 2.5% of income, or \$1,390 (i.e. twice the amount of penalty in the ACA), it leads to the full coverage. The main outcomes are reported in Column (5) of 1.16. Relative to Column (4), it achieves much lower unemployment rate and higher worker utility. Labor productivity decreases because less productive firms enter more into the labor market. More interestingly, it also achieves the substantial reduction in premium subsidies. This is due to the reduction of the unemployed who receive the full premium subsidies, and the decrease of premium itself.

1.9 Conclusion

We present and empirically implement an equilibrium labor market search model where risk averse workers facing medical expenditure shocks are matched with employers making health insurance coverage decisions. The distributions of wages, health insurance provisions, employer size, employment and worker's health are all endogenously determined in equilibrium. We estimate our model using various micro data sources including the 1996 panel of the Survey of Income and Program Participation (SIPP), the Medical Expenditure Panel Survey (MEPS, 1997-1999) and the 1997 Robert Wood Johnson Foundation Employer Health Insurance Survey. The equilibrium of our estimated model is largely consistent with the dynamics of the workers' labor market experience, health, health insurance and medical expenditure, as well as the distributions of employer sizes in the data.

We use our estimated model to examine the impact of the key components of the 2010 Affordable Care Act (ACA), including the individual mandate, the employer mandate, the

insurance exchange and the income-based insurance premium subsidy, as well as various combinations of these ACA components.

We find that the implementation of the full version of the ACA would significantly reduced the uninsured rate from 20.12% in the benchmark economy to 7.27%. This large reduction of the uninsured rate is mainly driven by low-wage workers participating in the insurance exchange with their premium supported by the income-based subsidies. We find that, if the subsidies were removed from the ACA, the insurance exchange will suffer from severe adverse selection problem so it is not active at all, though the presence of the exchange still leads to a small reduction of the uninsured rate from 20.12% in the benchmark to 17.14-17.28% under “ACA without the subsidies.”

We find that the ACA would also have achieved significant reduction in the uninsured rate if its individual mandate component were removed. We find in our simulation that under “ACA without individual mandate”, the uninsured rate would be 12.18%, significantly lower than the 20.12% under the benchmark. The premium subsidy component of the ACA would have in itself drawn all the unemployed (healthy or unhealthy) and the low-wage employed (again both healthy and unhealthy) in the insurance exchange.

Interestingly, we find that the current version of ACA *without* employer mandate may be more efficient than the one with employer mandate. The latter achieves higher average productivity, higher worker’s average utility, higher average wage, and similar government spending.

We also simulate the effects of eliminating the tax exemption for employer-sponsored health insurance (ESHI) premium both under the benchmark and under the ACA. We find that, while the elimination of the tax exemption for ESHI premium would reduce the probability of all firms, especially the larger ones, offering health insurance to their workers, the overall effect on the uninsured rate is rather modest. We find that in the benchmark economy the uninsured rate would increase from 20.12% to 23.39% when the ESHI tax exemption is removed; and it will increase from 7.27% to 9.15% under the ACA.

Finally, we find that prohibiting firms from ESHI would lead to a huge reduction on the fraction of active firms in the labor market. This result suggests that ESHI allows low productive firms to be active in the market because they can exploit the workers' risk premium.

We should emphasize that our study is only a first step toward understanding the mechanism through which the ACA, and more generally any health insurance reform, may influence labor markets equilibrium. We estimated our model using a selected sample of individuals with relatively homogeneous skills (men with no more than high school graduation between ages 26-46), and thus our quantitative findings may only be valid for this population. Thus the quantitative results we present in this study should be understood with these qualifications in mind. However, we believe that the various channels we uncovered in this study through which components of ACA interact with the labor market and with each other are of importance even in richer models.

There are many areas for future research. First and foremost, it will be important to introduce richer worker heterogeneity in the equilibrium labor market model; it is also important to endogenize health care decisions, and incorporate workers' life-cycle considerations. Incorporating Medicaid, the free public health insurance for the poor, into a model with endogenous asset accumulation decisions is also an important direction. Finally, there are many additional channels through which firms and workers might have responded to individual mandates and employer mandates that we abstracted in this study. We plan to address those issues in our future research.

Chapter 2

Health Insurance Exchange Design in an Empirical Equilibrium Labor Market Model

2.1 Introduction

The Patient Protection and Affordable Care Act of 2010 (ACA) represents the most significant health care reform in the United States in the past 40 years. The proposal and passage of the ACA was driven by the fact that close to 20% of the U.S. population does not have health insurance. Under the pre-ACA system, individuals tend to be uninsured if their family heads do not work at firms offering health insurance through employer sponsored health insurance (ESHI).¹ Among the uninsured, 80% are employed, but most of them are not offered health insurance by their employer; the remaining 20% are not employed. Although many provisions of the ACA will take effect starting in 2014, one of the most important provisions is the establishment of health insurance exchanges (HIX). HIX are individual insurance markets where insurance plans cannot price or deny coverage

¹Roughly 60% of the non-elderly have ESHI. Also, 10% of worker compensation consists of ESHI premia (Kaiser Family Foundation and Health Research and Educational Trust (2009)).

based on preexisting conditions.² A core idea of the ACA is to let the uninsured purchase health insurance from HIX. To provide individuals with incentives to participate in HIX, the ACA includes features such as individual mandates (a tax penalty on the uninsured), premium subsidies and an age-based rating.^{3,4} Although the desirability of these features has been discussed extensively in policy debates, there have been few studies evaluating the current HIX system or examining the possibility of alternative designs that improve welfare.

To examine the welfare impact of HIX design, a number of issues need to be addressed. First, HIX may be subject to adverse selection due to the prohibition of pricing based on preexisting conditions. This prohibition, by attracting disproportionately more unhealthy individuals, may cause healthy individuals to remain uninsured and incur the tax penalty rather than pay an actuarially unfair (higher) premium. Second, HIX interacts with the labor market due to the presence of ESHI. Individual health insurance purchase decisions in HIX, the key decisions affecting the extent of adverse selection, depend on the availability of ESHI, an outcome endogenously determined in the labor market. HIX may also affect worker labor supply and job mobility decisions, as well as firm decisions about offering ESHI.

In this study, I evaluate the current HIX system and consider the question of its optimal design, accounting for adverse selection and equilibrium labor market interactions. First, I develop and empirically implement a life cycle equilibrium labor market search model integrated with the pre-ACA health insurance market. Various forms of individual heterogeneity are incorporated to understand the heterogeneous welfare consequences of HIX. Second, using the estimated model, I investigate the impact of HIX, as implemented under the ACA, on individuals with different characteristics and the aggregate

²Health insurance exchanges are also known as health insurance marketplaces.

³Other important components of the ACA affecting the uninsured rate are Medicaid expansion and employer mandate (a tax penalty to large firms not providing coverage); these will be incorporated in my analysis as well.

⁴In the text of the ACA, individual (employer) mandates are formally referred to as individual (employer) shared responsibility.

consequences. Understanding these responses is necessary to conduct the welfare analysis. Third, I conduct a normative analysis. I examine the optimal design of HIX by finding, subject to the ACA's government revenue constraint, the values of three major design components that maximize a utilitarian social welfare function—tax penalties on the uninsured, premium subsidies and age-based rating regulations.⁵

The benchmark model is designed to explain key patterns among health, health insurance and labor market outcomes observed in the pre-ACA economy. It builds on several strands of literature. First, it builds on a growing literature of empirical life cycle models of labor supply and health (for example, Rust and Phelan (1997); Blau and Gilleskie (2008); Khwaja (2010); French and Jones (2011), Papageorge (2012), among others). Second, it builds on the small literature on equilibrium search models with endogenous ESHI provision (Dey and Flinn (2005); Bruegemann and Manovskii (2010); Aizawa and Fang (2013)).

In the model, individuals make health care utilization, labor supply and job mobility decisions over the life cycle. In each period, health status affects current period utility, the distribution of latent medical expenditure shocks and labor productivity. Health status is influenced by health care utilization, which takes into account its impact on future health status and its cost. Medical expenditure risk is insured if individuals are covered by health insurance. Health insurance may be offered by firms as part of compensation. In the labor market, both non-employed and employed workers meet firms randomly and then decide whether to accept a job offer based on the compensation package, which consists of a wage and an ESHI offering. Employed workers accumulate work experience, which increases their stock of skills. Finally, individuals differ by education and unobserved type, the latter being a determinant of risk preference, initial labor market skill, and the evolution of health status.

⁵The rationale of these policies in insurance markets with adverse selection has been extensively studied in public economics literature. See Einav, Finkelstein, and Levin (2010b) and Einav and Finkelstein (2011) for the survey.

Firms, which differ by their productivity, determine compensation packages to maximize their steady state flow profit. To account for various anti-discrimination laws restricting their choice of compensation package, I assume that firms set their wage offers by choosing a skill price subject to the constraint that the health cannot be priced, and decide whether to offer health insurance to all employees or not.⁶ Therefore, their coverage decisions are made to take into account the impact on the composition of workers having different characteristics within the firm. Moreover, I assume that health insurance costs are tax-deductible. I characterize a steady-state equilibrium where economic decisions by workers and firms are simultaneously determined.

The model features various forms of individual heterogeneity, both observed and unobserved, to understand the welfare consequences of HIX. I incorporate life cycle decision-making and skill characteristics because the major components of HIX design (individual mandates, premium subsidies and age-based rating regulation) vary by age and income. I also incorporate heterogeneity in health risk and risk preference, because those characteristics are known to be important determinants of health insurance purchasing decisions.⁷ Moreover, all of these characteristics impact the worker's and the firm's responses to the design of HIX. Individuals of different ages and skills change their labor supply and job mobility decisions differently because premia and subsidies in HIX vary by a worker's age and income. Also, individual heterogeneity impacts the firm's decision about offering ESHI given that it is affected by the composition of the firm's workforce. Although incorporating various life cycle dimensions of individuals into an equilibrium labor market search model substantially complicates the analysis, the model is still numerically tractable under my approach to solve the firm's compensation package, using techniques

⁶Existing anti-discrimination laws in the U.S. prohibit compensation packages from being based on age and health. Starting in 2014, firms are also prohibited from offering different sets of health plans to full-time employees with different income levels. Also, if ESHI is offered, workers cannot obtain premium subsidies from HIX unless the premium contribution of the ESHI plan for singles exceeds 9.5% of annual income. While part-time workers can obtain premium subsidies regardless of the ESHI offering, this study considers a demographic group that consists of few part-time workers.

⁷See Chetty and Finkelstein (2012) for an extensive review of the findings in the literature of empirical insurance markets.

that are an extension of Barlevy (2008), Burdett, Carrillo-Tudela, and Coles (2011) and Bagger, Fontaine, Postel-Vinay, and Robin (2013).

I estimate the model using three data sources: the 2004 Survey of Income and Program Participation (SIPP), the 2004-2007 Medical Expenditure Panel Survey (MEPS), and the Kaiser Family 2004-2007 Employer Health Insurance Benefit Survey (Kaiser). The first two are panel data sets on worker-side labor market, health, health insurance, and medical expenditure, while the third is a cross-sectional firm-level data set containing information about firms' characteristics and health insurance coverage. Worker-side data shows that individual health insurance coverage status is positively correlated with wages, education status, and age. Firm-side data shows that large firms tend to offer ESHI. Estimation is carried out via the method of simulated moments. The model fits worker-side moments such as health insurance, wage, employment, health, medical expenditure and their transitions over age profiles and education status, as well as firm-side moments such as the coverage rate and the size distribution.

The model estimates show that high-productivity firms are more likely to offer ESHI, which can be understood as follows. In the model, although firms want to attract more productive workers, high productivity firms are more likely to attract them because they can offer greater compensation. These more productive workers are typically experienced workers, who tend to be older and thus have a higher demand for health insurance. They also are more educated workers. From my estimates, educated workers are more risk-averse and have a higher demand for health insurance than less educated workers. The consequence of sorting of workers with high health insurance demand leads high-productivity firms to offer ESHI. This mechanism, although intuitive but generally ignored in the literature, simultaneously explains positive correlations among worker age, wage, education status, and health insurance coverage and also between firm size and health insurance coverage.

In counterfactual experiments, I introduce HIX to the pre-ACA economy as a compet-

itive individual insurance market where individuals can purchase health insurance. Moreover, I incorporate other important features of the ACA, including employer mandate and Medicaid expansion. I first investigate the impact of the ACA. The key features of the ACA's design for HIX are as follows. First, the tax penalty imposed on the uninsured increases with their income. Second, premium subsidies decrease with income. Third, the age-based rating regulation is such that the maximum allowable premium ratio (MPR) between the oldest and the youngest is 3.⁸

I find that the ACA decreases the uninsured rate in my estimation sample from 23.6% to 7.8%, where the remaining uninsured are mainly young employed workers who are healthy. Compared with them, the population of individuals purchasing health insurance from HIX consists of the sicker and older individuals. It therefore indicates the presence of adverse selection. The decrease in the uninsured rate leads to an increase in the fraction of healthy individuals, by 1 percentage point, as insured individuals tend to take health care utilization whenever they are hit by medical expenditure shocks relative to the uninsured.

Although the fraction of healthy individuals increases, and healthy individuals are more productive, I find that the steady state level of aggregate labor productivity decreases by 0.6%. This result arises because more workers are allocated to low productivity firms. The main channel is the inability of firms to make ESHI offerings separately for each employee and to provide the opportunity for employees to choose the source of coverage (i.e., from HIX or ESHI). If firms offer ESHI, their employees lose the opportunity to purchase subsidized health insurance from HIX. This makes accepting a job offer from a firm offering ESHI less attractive to older and less skilled workers, who can purchase health insurance at the lower cost given the ACA's design of premium subsidies and rating regulations. Those workers prefer not to move to high productivity jobs offering ESHI unless these jobs offer a high wage. As a result, more workers are allocated to low productivity firms. This decline of labor productivity contributes to a decrease in output

⁸The exact specification of HIX designs and other components in the ACA (employer mandate and Medicaid expansion) modeled in the study is in section 2.6.

per capita. Although the ACA increases the employment rate by 0.4 percentage point due to the increase in the fraction of healthy individuals, output per capita decreases by 0.2%.

Next, I investigate the optimal design of HIX, which maximizes social welfare subject to the ACA's government revenue constraint. I allow the government to optimally set the age-based rating regulation, which determines the maximum premium ratio (MPR) between the oldest and the youngest, and the premium subsidies and tax penalties to the uninsured as nonlinear functions of age and income. I find that the optimal combination of these policies increases aggregate labor productivity by 0.5% relative to the ACA while achieving a slightly lower level of the uninsured rate, 7.6% under the optimal HIX. To achieve the same welfare under HIX implemented by the ACA, the government needs to provide an annual lump-sum transfer to individuals amounting to \$195 per year, which corresponds to 7.6% of medical expenditure. The optimal structure makes it less beneficial for old workers relative to young workers to purchase health insurance from HIX by setting larger MPR and subsidies that decrease with age, rather than age-independent subsidies as in the ACA. In this structure, the adverse selection problem among the young is partially resolved. Moreover, it gives older workers an incentive to work at firms offering ESHI. This increases the allocation of workers from low to high productivity firms, raising aggregate labor productivity.

Finally, I assess the importance of modeling equilibrium labor market interactions to evaluate the design of HIX. I assume that firm's compensation package is exogenously determined and the same as the one under the pre-ACA economy. Then, I evaluate the impact of each component of HIX, as well as the optimal design of HIX. I find that both are qualitatively and quantitatively very different from the one with endogenous compensation packages. This finding suggests the importance of modeling equilibrium labor market interactions to evaluate the design of HIX.

The related literature. This study is related to several strands of literature. First of all, it belongs to a new and growing literature evaluating the impact of the ACA. Ericson and Starc (2012), Hackmann, Kolstad, and Kowalski (2013) and Handel et al. (2013) develop an equilibrium model of HIX with adverse selection and examine HIX designs.⁹ Ericson and Starc (2012) study the role of imperfect competition in the Massachusetts (MA) unsubsidized health insurance exchanges established due to the 2006 MA health care reform which has many features in common with the ACA, while Hackmann, Kolstad, and Kowalski (2013) investigate the extent of adverse selection and the optimal individual mandate in those exchanges. Handel et al. (2013) examine equilibria when multiple insurance products can be traded in the competitive insurance market. As in this study, these studies investigate the efficacy of individual mandates, premium subsidies, and age-based rating regulations. The common assumption of these studies is that individuals have access to HIX only and therefore they do not consider equilibrium labor market interactions.

Aizawa and Fang (2013) study the labor market effects of the ACA, arguing that explicit modeling of the labor market equilibrium is crucial to evaluate the ACA.¹⁰ They examine the impact of the ACA under the setting where workers are infinitely lived and homogeneous except in health. I advance this research agenda in a number of directions. First and most importantly, I shift the focus to the normative analysis, especially the optimal design of HIX, which requires modeling individual life cycle decisions and heterogeneity. Although incorporating these features into an equilibrium labor market search model substantially complicates the framework, I show how to maintain the tractability of analyses. Second, I show how equilibrium effects of HIX interact with individual heterogeneity. In particular, because the current HIX system forces a partial pooling between young and old individuals, the impact of the ACA on individuals with different ages differs

⁹See Bundorf, Levin, and Mahoney (2012a) and Geruso (2012) for studying premium rating within employers.

¹⁰See Pohl (2012) and Hai (2013) for an evaluation of Medicaid expansion.

substantially. Third, I show that the optimal structure of HIX consists of age-dependent policies.

The welfare analysis conducted in this chapter is also related to Kolstad and Kowalski (2012c), who evaluate the welfare impact of the 2006 MA health care reform by focusing on labor market distortions.¹¹ This study has a number of important differences from theirs: First, theirs is based on the sufficient statistic approach which allows to conduct a welfare analysis using difference-in-difference identification strategy. This approach requires both pre- and post-reform data, the latter being not available for the ACA at present. Second, while they consider labor market distortions in terms of the level of employment and wages, my study also takes into account distortions arising from the misallocation of workers. Third, while they examine welfare costs from the labor market distortion alone, my welfare analysis considers both the labor market and HIX together.¹²

My focus on the general equilibrium effects of HIX design is also related to studies assessing the macroeconomic impacts of ACA (Bruegemann and Manovskii (2010), Ozkan (2011), Cole, Kim, and Krueger (2012), Hansen, Hsu, and Lee (2012), Pashchenko and Porapakkarm (2013), and others). Although their models are richer than mine in certain dimensions (e.g., saving and capital accumulation), my study complements theirs by endogenizing firm coverage decisions and allowing both individual and firm heterogeneity, which play an important role in characterizing the optimal design of HIX.

Second, this study is also related to a large literature investigating the link between health insurance systems and labor markets, early contributions of which are reviewed by Currie and Madrian (1999) and Gruber (2000). A main topic in this literature has been whether the existing employer based health insurance system leads to an inefficient sorting of workers across firms, known as job lock and job push problems. Among them, Dey

¹¹See also Finkelstein, Taubman, Wright, Bernstein, Gruber, Newhouse, Allen, Baicker, and the Oregon Health Study Group (2012), Baicker, Finkelstein, Song, and Taubman (2013) and Garthwaite, Gross, and Notowidigdo (2013) for an evaluation of the impact of Medicaid on health and labor supply by focusing on recent Oregon Medicaid experiments and Tennessee Medicaid reform.

¹²See also Mulligan (2013b) and Mulligan (2013a) who argue that the ACA may distort labor markets because it imposes large marginal tax rates.

and Flinn (2005) is the closest to my study.¹³ They develop a search-matching-bargaining model with endogenous ESHI provisions and quantify the extent of job lock and job push in the pre-ACA economy. This study extends their approach and shows that HIX causes additional inefficiency in worker allocation.

Finally, this study is methodologically related to several branches of the labor and public finance literature. The empirical approach is closely related to the literature on structural estimation of equilibrium search allowing worker heterogeneity. Recently, Bagger, Fontaine, Postel-Vinay, and Robin (2013) estimate an equilibrium search model with two sided heterogeneity, allowing for individual human capital accumulation.¹⁴ This study adds to the literature by fitting various additional worker life cycle economic events and decisions within an empirical equilibrium labor market search model. My focus on how a certain institution affects labor productivity through worker reallocation is also related to Gourio and Roys (2012) and Garicano, LeLarge, and Van Reenen (2013), who empirically study the impact of regulations in France that are dependent on firm-size on worker reallocation and labor productivity. My empirical welfare analysis is related to the literature studying optimal insurance and taxation policies, including Einav et al. (2010c) for optimal mandates in annuity markets, and also Conesa, Kitao, and Krueger (2009) and Blundell and Shephard (2012) for optimal taxation. Finally, my welfare analysis about age- and income-based pricing and subsidies is related to the optimal taxation literature studying the role of tagging (e.g., Akerlof (1978); Michelacci and Ruffo (2011); Weinzierl (2011); Farhi and Werning (2013)).

The rest of this chapter is organized as follows. Section 2.2 presents the pre-ACA model of this study; Section 2.3 describes the data sets; Section 2.4 explains my estimation strategy; Section 2.5 presents my estimation results; Section 2.6 describes the main results

¹³See Fang and Gavazza (2011), a study of the link between ESHI systems and life-cycle medical expenditure. Lang (2010) and Dizioli and Pinheiro (2012) theoretically study incentives of ESHI provisions.

¹⁴Shephard (2012) also estimates a two-sided heterogenous version of Burdett and Mortensen (1998) model, but worker heterogeneity is assumed to be a permanent characteristic. See also Yamaguchi (2010) for an estimation of search-matching-bargaining model with endogenous human capital accumulation.

from evaluating the ACA and the components of HIX; Section 2.7 shows the main results from evaluating the ACA and the current HIX system; Section 2.8 concludes and discuss directions for future research.

2.2 Model

2.2.1 Environment

In this section, I first lay out an economic environment for the benchmark model, i.e., the model of the pre-ACA economy, which I use for estimation. Time is discrete and measured in periods of four months.¹⁵ Consider an economy populated with a continuum of workers with a measure $M > 0$ and a continuum of employers with a measure normalized to 1. They are randomly matched in a frictional labor market. Each worker lives for a finite horizon $t = t_0, \dots, T$, while employers exist infinitely. I use $\beta \in (0, 1)$ to denote the discount factor. Each worker makes health care, labor supply and job mobility decisions up to the period T . Then, they exogenously retire from the labor market and are replaced by newborn workers. Upon entering the labor market, the new workers are initially heterogeneous with respect to their education status ed which is either college graduate (C) or non-college graduate (NC) and with respect to their time-invariant type $type \in \{type1, type2\}$, the latter of which is a determinant of individual preference, labor market skills, and health transitions.¹⁶

Individual preference

Each individual has time-separable, expected-utility preferences, which are defined over consumption C_{it} ; employment status $P_{it} \in \{0, 1\}$, which takes a value of 1 if he is employed and 0 if he is not employed; and health status $h_{it} \in \{H, U\}$, which takes on value H if

¹⁵My choice of four months as a unit of time is motivated by the construction of data sets. See Section 2.3 for details.

¹⁶As is common with many empirical life cycle labor supply models, $type$ will be treated as heterogeneous and unobserved by the econometrician in the empirical part of this study.

he is healthy and U if he is unhealthy. Let $U_t(C_{it}, P_{it}, h_{it}; type)$ be the period utility for individuals with time-invariant type $type$, which takes the following functional form:

$$\begin{aligned}
U_t(C_t, P_t, h_t; type) &= \exp(-\gamma_{type}C_t) - \eta_{pt}P_t \\
&\quad - \eta_h^{type}\mathbf{I}(h_t = U) - \eta_{hpt}P_t\mathbf{I}(h_t = U)
\end{aligned}$$

where γ_{type} is the CARA coefficient, η_{pt} is the disutility from working which varies with individual age; η_h^{type} is the disutility from being unhealthy U ; and η_{hpt} is the disutility of work for an unhealthy individual which varies with individual age. The last term is incorporated to fit the relationship between health and employment status. I assume that individuals can neither save nor borrow. The budget constraint of the individual is then given by

$$C_t = \tau_w(w_t)P_t + (1 - P_t)b - OOP^{INS}(x_t m_t) \quad (2.1)$$

where $\tau_w(w_t)$ is after-tax labor income, b is non-employed income, and $OOP^{INS}(x_t m_t)$ is out-of-pocket expenditure. Out-of-pocket expenditure is a function of period t 's medical expenditure and health insurance status: $x_t \in \{0, 1\}$ is health care choice, m_t is latent medical expenditure shocks, and $INS \in \{0, 1\}$ is health insurance status, where $INS = 0$ if the individual is uninsured and $INS = 1$ if the individual is insured through ESHI. Although it is plausible to allow heterogeneous characteristics of insurance plans, given the limitation of the data, I assume that insurance is a homogeneous product in the pre-ACA model. Moreover, I assume that health insurance provides full insurance to individuals.

In addition, I specify that after period T , individuals receive the terminal value, which is merely a function of health status. The formal specification of the terminal value is described in Section (2.2.2).

Health shock and health transition

In each period, a worker is hit by latent medical expenditure shock m_t which is a function of health h_t , age (measured in four-month intervals) t , and idiosyncratic shock ϵ_t . I specify its functional form as follows:

$$\begin{aligned} m_t &= \max\{m_t^* - \kappa_{h_t}, 0\} \\ m_t^* &= \exp(\omega_1^{h_t} + \omega_2^{h_t}t + \omega_3^{h_t}t^2 + \epsilon_t^m), \\ \epsilon_t^m | h_t &\sim i.i.d.N(0, \sigma_{h_t}^2), \\ h_t &\in \{H, U\} \end{aligned}$$

where m_t^* is a latent health shock which is distributed over the log normal distribution; and ϵ_t is i.i.d. idiosyncratic shock which is conditionally heteroskedastic with respect to health status h_t . In this specification, I incorporate a parameter $\kappa_{h_t} > 0$ which is used to capture the possibility that individuals do not report any positive medical expense because they are not hit by a positive amount of latent medical expenditure shocks, i.e., they are not hit by health shocks at all. Moreover, note that all of the parameters in the latent medical expenditure distribution differ according to the beginning of period health status h_t .

Conditional on latent medical expenditure shocks m_t , the worker chooses health care utilization $x_t \in \{0, 1\}$, which affects the realization of the next period health status. The transition to next period health status is determined by

$$\begin{aligned}
\Pr[h_{t+1} = k | x_t, h_t, \epsilon_t^m, t, type] &= \frac{\exp\left(\begin{aligned} &\phi_{1k}m_t + \phi_{2k}x_t + \phi_{3k}x_tm_t + \phi_{4k}t \\ &+ \sum \phi_{5k}^i \mathbf{1}(h_t = i) + \sum \phi_{6k}^j \mathbf{1}(type = j) \end{aligned}\right)}{\sum_{k'} \exp\left(\begin{aligned} &\phi_{1k'}m_t + \phi_{2k'}x_t + \phi_{3k'}x_tm_t + \phi_{4k'}t \\ &+ \sum \phi_{5k'}^i \mathbf{1}(h_t = i) + \sum \phi_{6k'}^j \mathbf{1}(type = j) \end{aligned}\right)}
\end{aligned}$$

where $\mathbf{1}(type = j)$ is the dummy variable for individual unobserved type j . It is important to allow an interaction term between x_t and m_t . Otherwise, to avoid reducing consumption, individuals tend to choose no health care utilization when they are hit by large latent medical expenditure shocks.

Health insurance market

In the baseline model intended to capture the pre-ACA U.S. health insurance market, I assume that workers can obtain health insurance only if their employers offer it. This is a simplifying assumption meant to capture the fact that the individual private insurance market is very small in the U.S.¹⁷ In my counterfactual experiment, I will introduce a competitive individual insurance market which I call the health insurance exchange (HIX).

Individual labor productivity

Each individual possesses labor productivity which affects the size of their compensation. I assume that an individual produces output $e_{\bar{x}}(p)$ as a function of (1) a vector of individual

¹⁷Indeed, the fraction of individuals with individual insurance among the whole sample is 2%, and the fraction of such individuals among the sum of individuals with individual insurance and the uninsured is just 10% in my data set. Moreover, even for those who are covered by individual insurance, insurance products typically do not cover pre-existing conditions. In addition, one in seven applicants for health insurance are rejected (Hendren (2013)). Hendren (2013) provides both theoretical and empirical analyses showing that adverse selection leads to a collapse of many other individual insurance markets.

characteristics $\tilde{X} = (ed, type, E_t, h_t)$ where ed is education, $type$ is individual permanent type, E_t is labor market experience, and h_t is health status, and (2) the permanent productivity of the firm the individual is currently matched with, denoted by p . The log of output is specified as

$$\begin{aligned}
\ln(e_{\tilde{X}}(p)) &= e_w^*(ed, type, E_t) + e_h^*(h_t) + p \\
&= \sum \alpha_1^e \mathbf{1}(ed = e) + \sum \alpha_1^k \mathbf{1}(type = k) \\
&+ \left(\sum \alpha_4^e \mathbf{1}(ed = e) + \sum \alpha_4^k \mathbf{1}(type = k) \right) E_t \\
&+ \left(\sum \alpha_5^e \mathbf{1}(ed = e) + \sum \alpha_5^k \mathbf{1}(type = k) \right) E_t^2 \\
&+ \alpha_h \mathbf{1}(h_t = U) + p.
\end{aligned} \tag{2.2}$$

where $e_w^*(ed, type, E_t)$ is the worker skill explained by $(ed, type, E_t)$ and $e_h^*(h_t)$ is the worker skill explained by h_t . I assume that output is multiplicatively separable in $(ed, type, E_t)$ and h_t . The separability of health is assumed to maintain tractability when I characterize firm's optimal wage policy.

Individual labor market experience is accumulated as long as the individual is employed. That is,

$$E_{t+1} = \begin{cases} E_t + 1 & \text{if } P_t = 1 \\ E_t & \text{if } P_t = 0 \end{cases}.$$

Firm

Firms are heterogeneous with respect to their permanent productivity. In the population of firms, the distribution of productivity is denoted by $\Gamma(\cdot)$ which has a density function $d\Gamma$ that is continuous and positive everywhere. In my empirical application, I specify Γ to be lognormal with mean μ_p and variance σ_p^2 , i.e., $p \sim \ln N(\mu_p, \sigma_p^2)$.

Firms have access to a constant return to scale production function. In each period, they offer a package of wage and health insurance provision to maximize their steady state profit flow. If they offer health insurance, they incur the cost of health insurance provision,

which is equal to the sum of the total expected medical expenditure of their workforce and a fixed administrative cost ξ_{ESHI} . The health insurance costs are tax-exempt.

Compensation package

Health insurance provision and the wage offer are determined as a solution of the firm's profit maximization problem. Given the laws prohibiting discrimination in compensation packages based upon health or age, firms cannot offer separate compensation packages to each individual. Moreover, given this constraint and dimensionality of individual characteristics, it is hard to solve the optimal compensation contract for each worker. Therefore, I reduce the dimension of the potential contract space so that the model is still tractable, but captures the most important patterns of the data.

Specifically, I assume that firms post a skill price θ^{ed} for each skill group ed subject to the constraint that health h_t cannot be priced. Moreover, I assume that firms decide whether to offer health insurance to all of their workforce $INS \in \{0, 1\}$. As a result, a worker with characteristics $\tilde{X} = (ed, type, E_t, h_t)$ in a firm offering a compensation package (θ, INS) receives a wage offer which is equal to

$$w_{\tilde{X}}^{INS}(\theta) = \theta_{INS}^{ed} \exp(e_w^*(ed, type, E_t)). \quad (2.3)$$

Given the wage offer, the flow profit of a firm with productivity p from hiring a worker with \tilde{X} is $\exp(e_w^*(ed, type, E_t))(\exp(p)e_h^*(h_t) - \theta)$ where $g_h^*(h_t)$ is the productivity effect of health.¹⁸ Finally, I assume that wages are subject to classical measurement error, with errors following a log normal distribution.¹⁹

¹⁸I implicitly assume that employer contribution of insurance premium is 100% in this baseline model. That is, if workers want to opt out the coverage, they cannot receive an additional wage to compensate. This is not an unrealistic choice because the current U.S. average is 85% for single workers' premiums.

¹⁹One could instead model these errors as transitory skill shock. My treatment of measurement error follows the work of Low, Meghir, and Pistaferri (2010) which estimates the wage process of a life-cycle search model using the same data source (SIPP).

Labor market

The labor market is frictional and workers and employers randomly meet. A non-employed worker receives a new job offer from an employer with probability λ_u^{ed} and an employed worker receives an offer from an employer with probability λ_e^{ed} where ed is individual education status. The compensation is drawn from the offer distribution $F^{ed}(\theta, INS)$. Upon receiving the job offer, the worker decides whether to accept it.

In addition to changing jobs, employed workers are allowed to quit and become non-employed. Furthermore, they are hit by an exogenous job destruction shock with probability δ^{ed} , upon which workers lose their current jobs. Because the model period is a relatively long, I allow the exogenous job destruction shock and the arrival of a new job offer to occur within the same period with probability $\delta^{ed}\lambda_e^{ed} > 0$. Moreover, I allow an additively separable preference shock to being non-employed ϵ_t^n which follows a Type-I extreme value distribution with scale parameter σ_n .²⁰

Timing in a period

At the beginning of each period, individuals, who are heterogeneous in their education, unobserved type, age, and health status, are either unemployed or working for employers offering different wage and health insurance packages. I now describe the explicit timing assumptions for a period that I use in the derivation of the value functions in Section 2.2.2. These particular timing assumptions simplify our derivation, but they are not crucial.

1. An employed individual produces output and accumulates labor market experience.
2. Idiosyncratic health shock ϵ_t^m is realized.
3. An individual makes health care decision x_t .

²⁰This preference shock is incorporated to smooth the labor supply function with respect to wage offers, which is useful when I solve the model numerically. An alternative is to add time-invariant continuous heterogeneous flow utility of being non-employed, as done by Bontemps, Robin, and Van den Berg (1999). Because of the nonstationary nature of the individual problem in my model, I need to allow for dynamics of unobserved flow utility, which substantially complicates analysis.

4. The next period health status is realized.
5. An employed worker is hit by an exogenous job destruction shock with probability δ .
6. An individual draws a preference shock for being non-employed.
7. A non-employed worker receives a job offer with probability λ_u and decides whether to accept the job offer. An employed worker receives a job offer with probability λ_e and chooses to accept the offer, to stay at the current job, or to quit and become non-employed. The employed worker who does not receive the offer decides whether to stay at the current job or quit and become non-employed.

Initial condition

I assume that non-college graduates enter the labor market at age 18, while college graduates enter at age 22. I assume that upon entry into the labor market, they have no labor market experience and start their career as non-employed. Notice that the return to education is captured by labor productivity. Workers are assigned an unobserved type, the distribution of which is specific to education. Moreover, I assume that all newborn workers are healthy.²¹ Finally, I assume that the population of individuals and firms grows at the constant rate n in each period. Let μ_t^τ denote the fraction of a cohort size of age t in the total population of individuals at period τ . I consider a steady state economy in which μ_t^τ is constant over time, i.e., $\mu_t^\tau = \mu_t$.

2.2.2 Analysis of the Model

In this section, I define the steady state equilibrium of the model. To do so, I first consider the individual life cycle optimization problem. The solution to the individual optimization problem is then used to determine the steady state distribution of individuals. Next, I

²¹Note, however, that their health evolution is affected by their permanent type *type* and therefore they face different health risks.

formulate the firm's optimization problem and characterize their optimal decisions for posting compensation package.

Individual optimization problem over the life cycle

At the beginning of a period, the state space of an individual at age t consists of a six dimensional vectors: $(\tilde{X}_t, \theta, INS)$: four dimensional vector of individual characteristics $\tilde{X}_t = (ed, type, E_t, h_t)$ and the compensation package (θ, INS) . Note that $INS_t = 0$ for any non-employed worker. Consider that all individuals face offer distribution of compensation package, which is denoted by $F^{ed}(\theta, INS)$. Denote $V_0^t(\tilde{X}_t)$ as the value function of a non-employed worker with age t who is in a state \tilde{X} and $V_1^t(\tilde{X}_t, \theta, INS)$ is the value function of an employed worker with age t who is in a state $(\tilde{X}_t, \theta, INS)$.

Consider a non-employed worker having characteristics $\tilde{X}_t = (ed, type, E_t, h_t)$. His value function is defined by

$$V_0^t(\tilde{X}_t) = \mathbb{E}_{\epsilon_t^m} \left[\begin{array}{l} \max_{x_t} U_t(C_t(x_t, \epsilon_t^m, \tilde{X}_t), 0, h_t) \\ + \beta \sum_{\hat{h}} \Pr(h_{t+1} = \hat{h} | \tilde{X}_t, x_t, \epsilon_t^m) \lambda_u^{ed} \tilde{V}_0^t(\tilde{X}_{t+1}) \\ + \beta \sum_{\hat{h}} \Pr(h_{t+1} = \hat{h} | \tilde{X}_t, x_t, \epsilon_t^m) (1 - \lambda_u^{ed}) V_0^{t+1}(\tilde{X}_{t+1}) \end{array} \right] \quad (2.4)$$

subject to budget constraint (2.1) where \tilde{X}_{t+1} is the next period's individual characteristics, $\tilde{X}_{t+1} = (ed, type, E_t, \hat{h})$. $\tilde{V}_0^t(\tilde{X}_{t+1})$ is defined as

$$\tilde{V}_0^t(\tilde{X}_{t+1}) = \int \mathbb{E}_{\epsilon_t^n} [\max\{V_0^{t+1}(\tilde{X}_{t+1}) + \epsilon_t^n, V_1^{t+1}(\tilde{X}_{t+1}, \theta, INS)\}] dF^{ed}(\theta, INS)$$

The first term is the flow utility from the current period's consumption which is affected by the health care utilization choice x_t and the realization of an i.i.d. latent medical expenditure shock ϵ_t^m . The second term is the expected value from receiving a job offer, denoted by $\tilde{V}_0^t(\tilde{X}_{t+1})$, multiplied by the health transition probability which is affected by x_t and the realization of the i.i.d. latent medical expenditure shock ϵ_t^m . Conditional on

receiving the offer, the individual decides whether to accept it, which depends on the value from staying as unemployed $V_0^{t+1}(\tilde{X}_{t+1})$ and the value from accepting the job offer with compensation package (θ, INS) denoted $V_1^{t+1}(\tilde{X}_{t+1}, \theta, INS)$. Note that the decision is subject to the preference shock of working, denoted by ϵ_t^n . The third term is the expected value from not receiving a job offer.

Similarly, consider an employed worker who is in a state $(\tilde{X}_t, \theta, INS)$. The value function of such an individual is defined by

$$V_1^t(\tilde{X}_t, \theta, INS) = \mathbb{E}_{\epsilon_t^m} \left[\begin{array}{l} \max_{x_t} U_t(C_t(x_t, \epsilon_t^m, \tilde{X}_t, \theta, INS), 1, h_t) \\ + \beta \sum_{\hat{h}} \Pr(h_{t+1} = \hat{h} | \tilde{X}_t, x_t, \epsilon_t^m) \delta^{ed} \tilde{V}_1^t(\tilde{X}_{t+1}) \\ + \beta \sum_{\hat{h}} \Pr(h_{t+1} = \hat{h} | \tilde{X}_t, x_t, \epsilon_t^m) (1 - \delta^{ed}) \hat{V}_1^t(\tilde{X}_{t+1}, \theta, INS) \end{array} \right] \quad (2.5)$$

subject to budget constraint (2.1) where $\tilde{X}_{t+1} = (ed, type, E_t + 1, \hat{h})$ is the next period individual characteristics and

$$\begin{aligned} \tilde{V}_1^t(\tilde{X}_{t+1}) &= (1 - \lambda_e^{ed}) V_n^{t+1}(\tilde{X}_{t+1}) \\ &+ \lambda_e^{ed} \int \mathbb{E}_{\epsilon_t^n} \max \left\{ \begin{array}{l} V_0^{t+1}(\tilde{X}_{t+1}) + \epsilon_t^n, \\ V_1^{t+1}(\tilde{X}_{t+1}, \theta', INS') \end{array} \right\} dF^{ed}(\theta', INS') \end{aligned}$$

and

$$\begin{aligned} \hat{V}_1^t(\tilde{X}_{t+1}, \theta, INS) &= (1 - \lambda_e^{ed}) \mathbb{E}_{\epsilon_t^n} \max \left\{ \begin{array}{l} V_0^{t+1}(\tilde{X}_{t+1}) + \epsilon_t^n, \\ V_1^{t+1}(\tilde{X}_{t+1}, \theta, INS) \end{array} \right\} \\ &+ \lambda_e^{ed} \int \mathbb{E}_{\epsilon_t^n} \max \left\{ \begin{array}{l} V_0^{t+1}(\tilde{X}_{t+1}) + \epsilon_t^n, \\ V_1^{t+1}(\tilde{X}_{t+1}, \theta, INS), \\ V_1^{t+1}(\tilde{X}_{t+1}, \theta', INS') \end{array} \right\} dF^{ed}(\theta', INS'). \end{aligned}$$

As before, the first term is the flow utility from the current period's consumption, which is affected by health care utilization choice x_t and the realization of the i.i.d. latent

medical expenditure shock ϵ_t^m . The second term is the expected value from being hit by an exogenous job destruction shock, $\tilde{V}_1^t(\tilde{X}_{t+1})$, which consists of two parts, expressed in the following square bracket. The first part is the gain from not receiving a job offer. In this case, the worker is forced to be non-employed in the next period. The second part is the gain from receiving a job offer. In this case, the worker decides whether to accept the job offer or remain non-employed. The third term is the expected value from being hit by an exogenous job destruction shock, denoted by $\hat{V}_1^t(\tilde{X}_{t+1}, \theta, INS)$. Again, it consists of two parts, expressed in the following square bracket. The first part is the gain from not receiving a job offer from another firm. In this case, the worker decides whether to stay in the current job or to become non-employed. The second part is the case in which the worker receives a job offer. If the worker receives a job offer with compensation package (θ', INS') , he decides whether to move to a new firm, stay in his current job, or quit and become non-employed.

Because individuals live for a finite period, we need to specify the value of the terminal period. The terminal value function of an individual with state S_t is simply given by

$$V_i^T(S_T) = v_h^T \mathbf{I}(h_T = U), \quad (2.6)$$

where $i \in \{0, 1\}$ is his employment status at period T .

It is easy to see that the decisions of non-employed individuals about accepting a job offer follow the standard reservation rule strategy. Here, notice that the strategy is nonstationary because my model uses a life cycle environment. It is easy to see that for the employed worker working at a firm offering a compensation package of skill price and health insurance provision status (θ, INS) and receiving a job offer from a firm offering a compensation package (θ', INS') , his job acceptance decision is purely determined by comparing these two compensation packages and is not affected by firm productivity per se. This is because individual wages are fully determined by skill prices given the skill level, as seen from (2.3).

Moreover, it is clear that individual decisions to choose health care utilization x_t take into account its impact on the current period's consumption and the dynamics of health status. Because health enters into an individual's direct utility, which allows interactions with disutility from working, an individual's health care decisions affect his ability to work in the future.

Steady state worker distribution

The steady state distribution of workers is given by workers' optimal choices. I assume that at $t = 1$, new workers enter the population as non-employed with exogenously determined health, education, and type. Because workers live for a finite period, the steady state distribution of individuals includes age as state variables. Denote $g_t(\tilde{X}, \theta, INS)$ as a steady state measure of employed workers with age t and characteristics $\tilde{X}_t = (ed, type, E, \hat{h})$ receiving compensation packages (θ, INS) . Similarly, $u_t(\tilde{X})$ is defined as a steady state measure of the non-employed with age t and characteristics \tilde{X} . Note that g_t and u_t are fully determined by the inflow from the distribution of g_{t-1} and u_{t-1} . The determinants of $g_t(\tilde{X}, INS, \theta_{INS}^{ed})$ and $u_t(\tilde{X})$ are described as follows. Given the offer distribution of

compensation package $F^{ed}(\theta, INS)$:

$$\begin{aligned}
\frac{g_t(\tilde{X}_t, \theta, INS)}{1+n} &= \sum_{h_{t-1}} g_{t-1}(\tilde{X}_{-1}^A, \theta, INS) \\
&\quad \times \mathbb{E}_{\epsilon_t^m} \left[\Pr(h_t = \hat{h} | \tilde{X}_{-1}^A, \theta, INS, \epsilon_t^m) \right] \Psi_{t-1}(\tilde{X}_t, \theta, INS) \\
&\quad + \sum_{h_{t-1}} u_{t-1}(\tilde{X}_{-1}^B) \\
&\quad \times \mathbb{E}_{\epsilon_t^m} \left[\Pr(h_t = \hat{h} | \tilde{X}_{-1}^B, \epsilon_t^m) \right] \Pr(\Omega_1^E(\tilde{X}_t, \theta, INS)) \\
&\quad \times \lambda_u^{ed} f^{ed}(\theta, INS) \\
&\quad + \sum_{h_{t-1}} \sum_{INS'} \int g_{t-1}(\tilde{X}_{-1}^A, \theta', INS') \\
&\quad \times \mathbb{E}_{\epsilon_t^m} \left[\Pr(h_t = \hat{h} | \tilde{X}_{-1}^A, \theta', INS', \epsilon_t^m) \right] \mathbf{1}(\theta, INS, \theta', INS') d\theta' \\
&\quad \times (1 - \delta^{ed}) \lambda_e^{ed} f^{ed}(\theta, INS) \\
&\quad + \sum_{h_{t-1}} \sum_{INS'} \int g_{t-1}(\tilde{X}_{-1}^A, \theta', INS') \\
&\quad \times \mathbb{E}_{\epsilon_t^m} \left[\Pr(h_t = \hat{h} | \tilde{X}_{-1}^A, \theta', INS', \epsilon_t^m) \right] \Pr(\Omega_1^E(\tilde{X}, \theta, INS)) d\theta' \\
&\quad \times \delta^{ed} \lambda_e^{ed} f^{ed}(\theta, INS) \tag{2.7}
\end{aligned}$$

where $\tilde{X}_{-1}^A = (ed, type, E_{t-1}, h_{t-1})$ and $\tilde{X}_{-1}^B = (ed, type, E_t, h_{t-1})$ are individual characteristics in the last period for the employed and non-employed, respectively, which can turn into \tilde{X} in this period. $\Psi_{t-1}(\tilde{X}_t, \theta, INS)$ is the probability of staying at the same firm

$$\begin{aligned}
\Psi_{t-1}(\tilde{X}_t, \theta, INS) &= (1 - \delta^{ed})(1 - \lambda_e^{ed}) \Pr(\Omega_1^E(\tilde{X}_t, \theta, INS)) \\
&\quad + (1 - \delta^{ed}) \lambda_e^{ed} (\Pr(\Omega_2^E(\tilde{X}_t, \theta, INS)) \Pr(\Omega_1(\tilde{X}_t, \theta, INS))).
\end{aligned}$$

(Note that I will provide an intuition for each term of Ψ_{t-1} in the next paragraph). $\Pr(\Omega_1^E(\tilde{X}_t, \theta, INS))$ is the probability that individuals with \tilde{X} prefer to work a job with compensation package (θ, INS) over being non-employed, where the set $\Omega_1^E(\tilde{X}_t, \theta, INS)$

is formally defined as

$$\Omega_1^E(\tilde{X}_t, \theta, INS) = V_0^t(\tilde{X}_t) + \epsilon_t^n < V_1^t(\tilde{X}_t, \theta, INS)$$

where ϵ_t^n is a preference shock from not working. Because it follows a type-I extreme value distribution with scale parameter σ_n , $\Pr(\Omega_1^E(\tilde{X}_t, \theta, INS))$ is characterized as

$$\Pr(\Omega_1^E(\tilde{X}_t, \theta, INS)) = \frac{\exp\left(\frac{V_1^t(\tilde{X}_t, \theta, INS)}{\sigma_n}\right)}{\exp\left(\frac{V_0^t(\tilde{X}_t)}{\sigma_n}\right) + \exp\left(\frac{V_1^t(\tilde{X}_t, \theta, INS)}{\sigma_n}\right)}.$$

Next, $\Pr(\Omega_2^E(\theta, INS))$ is the probability that individuals who receive a job offer from other firms decide to stay in the current job:

$$\Pr(\Omega_2^E(\tilde{X}_t, \theta, INS)) = F^{ed}(\theta, INS) + F^{ed}(\tilde{\theta}_{\widehat{INS}}(\tilde{X}_t, \theta), \widehat{INS})$$

for $\widehat{INS} \neq INS$ where $\tilde{\theta}_{\widehat{INS}}(\tilde{X}_t)$ is the threshold skill price which can be defined as

$$V_1^t(\tilde{X}_t, \theta, INS) = V_1^t(\tilde{X}_t, \tilde{\theta}_{\widehat{INS}}(\tilde{X}_t, \theta), \widehat{INS}).$$

$\mathbf{1}(\theta, INS, \theta', INS')$ is the indicator function such that individuals prefer to take an offer from (θ, INS) over (θ', INS') :

$$\mathbf{1}(\theta, INS, \theta', INS') = \begin{cases} 1 & \text{if } V_1^t(\tilde{X}_t, \theta, INS) > V_1^t(\tilde{X}_t, \theta', INS') \\ 0 & \text{otherwise} \end{cases}$$

While the expression of how g_t is determined looks rather complicated, it can be understood fairly easily. The first term is the inflow from the workers who work in firms offering the same contract (θ, INS) , transition to the health status \hat{h} this period, and decide to stay at the same firm. Notice that the health transition probability is denoted by $\Pr(h_t = \hat{h} | \tilde{X}_{-1}^A, \theta, INS', \epsilon_t^m)$ which does not include health care decision x_t , as the op-

timal health care decision is a function of a vector of $(\tilde{X}_{-1}^A, \theta, INS', \epsilon_t^m)$. The probability of staying at the same firm, denoted by $\Psi_{t-1}(\theta, INS, \tilde{X})$, consists of two terms. The first term is the probability that the worker does not receive any offer at all and prefers to stay at the same firm over quitting and becoming non-employed. The second term is the probability that the worker receives an offer from another firm but turns it down and stays at the same firm. The second term of the right-hand side of (2.7) is the inflow from the non-employed. This happens if a non-employed individual gets an offer with probability $\lambda_u^{ed} f^{ed}(\theta, INS)$ and decides to accept it. The third and fourth terms of the right-hand side of (2.7) are the inflow from the currently employed working in other firms but receiving and accepting offers with compensation (θ, INS) . The difference between the third and fourth terms derives from whether they are hit by an exogenous job destruction shock, which affects the probability of accepting a job offer. Finally, in order to determine the size of g_t , I need to take into account that population size grows at a constant rate n . Therefore, the population at age t , g_t , should be divided by $1 + n$.

Similarly, one can express the determinant of $u_t(\tilde{X}_t)$ as

$$\begin{aligned}
\frac{u_t(\tilde{X}_t)}{1+n} &= \sum_{h_{t-1}} u_{t-1}(\tilde{X}_{-1}^B) \\
&\quad \times \mathbb{E}_{\epsilon_t^m} \left[\Pr(h_t = \hat{h} | \tilde{X}_{-1}^B, \epsilon_t^m) \right] (1 - \lambda_u^{ed}) \\
&+ \sum_{h_{t-1}} u_{t-1}(\tilde{X}_{-1}^B) \\
&\quad \times \mathbb{E}_{\epsilon_t^m} \left[\Pr(h_t = \hat{h} | \tilde{X}_{-1}^A, \epsilon_t^m) \right] \Pr(\Omega_1^U(\tilde{X}_t)) \lambda_u^{ed} \\
&+ \sum_{h_{t-1}} \sum_{INS} \int g_{t-1}(\tilde{X}_{-1}^A, \theta, INS) \\
&\quad \times \mathbb{E}_{\epsilon_t^m} \left[\Pr(h_t = \hat{h} | \tilde{X}_{-1}^A, \epsilon_t^m) \right] \Pr(\Omega_2^U(\tilde{X}_t, \theta, INS)) dF(\theta, INS) (1 - \delta^{ed})(1 - \lambda_e^{ed}) \\
&+ \sum_{h_{t-1}} \sum_{INS} \int g_{t-1}(\tilde{X}_{-1}^A, \theta, INS) \\
&\quad \times \mathbb{E}_{\epsilon_t^m} \left[\Pr(h_t = \hat{h} | \tilde{X}_{-1}^A, \epsilon_t^m) \right] \Pr(\Omega_3^U(\tilde{X}_t, \theta, INS)) dF(\theta, INS) (1 - \delta^{ed}) \lambda_e^{ed} \\
&+ \sum_{h_{t-1}} \sum_{INS} \int g_{t-1}(\tilde{X}_{-1}^A, \theta, INS) \\
&\quad \times \mathbb{E}_{\epsilon_t^m} \left[\Pr(h_t = \hat{h} | \tilde{X}_{-1}^A, \epsilon_t^m) \right] d\theta \delta^{ed} (1 - \lambda_e^{ed}) \\
&+ \sum_{h_{t-1}} \sum_{INS} \int g_{t-1}(\tilde{X}_{-1}^A, \theta, INS) \\
&\quad \times \mathbb{E}_{\epsilon_t^m} \left[\Pr(h_t = \hat{h} | \tilde{X}_{-1}^A, \epsilon_t^m) \right] \Pr(\Omega_1^U(\tilde{X}_t)) d\theta \delta^{ed} \lambda_e^{ed} \tag{2.8}
\end{aligned}$$

where \tilde{X}_{-1}^A and \tilde{X}_{-1}^B are defined as above, and $\Pr(\Omega_1^U(\tilde{X}_t))$ is the probability that the unemployed with characteristics \tilde{X} decides to turn down the offer:

$$\Pr(\Omega_1^U(\tilde{X}_t)) = \int \Pr(V_0^t(\tilde{X}_t) + \epsilon_t^n > V_1^t(\tilde{X}_t, \theta, INS)) dF^{ed}(\theta, INS).$$

$\Pr(\Omega_2^U(\tilde{X}_t, \theta, INS))$ is the probability that an employed individual with characteristics \tilde{X}_t who has a job with compensation package (θ, INS) decides to quit and become non-employed:

$$\Pr(\Omega_2^U(\tilde{X}_t, \theta, INS)) = \Pr(V_0^t(\tilde{X}_t) + \epsilon_t^n > V_1^t(\tilde{X}_t, \theta, INS)).$$

Finally, $\Pr(\Omega_3^U(\tilde{X}_t, \theta, INS))$ is the probability that an employed individual with charac-

teristics \tilde{X}_t who has a job with compensation package (θ, INS) receives a job offer from another firm but decides to quit and become non-employed:

$$\begin{aligned} & \Pr(\Omega_3^U(\tilde{X}_t, \theta, INS)) \\ &= \int \mathbf{1} \left(V_0^t(\tilde{X}_t) + \epsilon_t^n = \max \left\{ \begin{array}{l} V_0^t(\tilde{X}_t) + \epsilon_t^n, \\ V_1^t(\tilde{X}_t, \theta, INS), \\ V_1^t(\tilde{X}_t, \theta', INS') \end{array} \right\} \right) dF^{ed}(\theta', INS'). \end{aligned}$$

Again, one can interpret (2.8) very intuitively. The first two terms of the right-hand side of (2.8) are the inflow from the currently non-employed workers with state variable \tilde{X}_{-1}^B . They stay non-employed and have state variable \tilde{X}_t in the following period if they are not offered health insurance, or if they receive a job offer but turn it down. The third and the fourth terms are the inflow from currently employed workers with state variable \tilde{X}_{-1}^A who choose to quit and become non-employed: the third term is the case where they do not receive a job offer from another firm; the fourth term is the case where they receive a job offer from another firm but turn it down. The fifth and sixth terms are the inflow from the currently employed who are hit by job destruction shock. The fifth term depicts those individuals who are forced to be unemployed and the sixth term depicts individuals who also receive job offers from other firms, but choose to quit and become non-employed.

Finally, the steady state condition requires that the sum of all measures of workers is equal to M , that is

$$\sum_t \sum_{\tilde{X}_t} u_t(\tilde{X}_t) + \sum_t \sum_{\tilde{X}_t} \sum_{INS} \int g_t(\tilde{X}_t, \theta, INS) d\theta = M. \quad (2.9)$$

Therefore, by using (2.7), (2.8) and (2.9), one can characterize the steady state worker distribution. Although the expression looks rather complicated, it can be derived fairly easily by using forward induction. That is, we can analytically calculate the whole dis-

tribution once we know the period 1 distribution g_1 and u_1 , the worker's value function, health transition function, and the offer distribution $F^{ed}(\theta, INS)$.

From the steady state employment measure $g_t(\tilde{X}_t, \theta, INS)$, one can define the terms related to firm size, by following the same spirit as Burdett and Mortensen (1998) and Postel-Vinay and Robin (2002). Specifically, the density of employees with age t and characteristics \tilde{X}_t for firms offering compensation package (θ, INS) is given by

$$l_t(\tilde{X}_t, \theta, INS) = \frac{g_t(\tilde{X}_t, \theta, INS)}{f^{ed}(\theta, INS)}$$

where $f^{ed}(\theta, INS)$ is density of firms offering compensation package (θ, INS) .

Firm's optimization problem

Next, I formalize the firm's problem. Firms choose wage offers and health insurance offerings to maximize steady state profit flow. I assume that the firm draws a shock, ϵ^{ESHI} , in each period, which is specific to its choice of whether to offer health insurance. The shock is additively separable from the steady state profit flow, but affects the payoff of health insurance provisions. I incorporate the choice-specific shock to smooth the insurance provision decisions of the employers. This problem can be formulated as

$$\Pi(p, \epsilon^{HI}) = \max \{ \Pi_1(p) + \epsilon^{ESHI}, \Pi_0(p) \}$$

where Π_{INS} is the conditional profit under the health insurance offering status $INS \in \{0, 1\}$. It is a solution of the following problems:

$$\Pi_1(p) = \max_{\theta_1^{ed}} \sum_t \sum_{\tilde{X}} \pi_1(\tilde{X}, t, p, \theta_1^{ed}) l_t(\tilde{X}, \theta_1^{ed}, 1) - \xi_{ESHI} \quad (2.10)$$

$$\Pi_0(p) = \max_{\theta_0^{ed}} \sum_t \sum_{\tilde{X}} \pi_0(\tilde{X}, t, p, \theta_0^{ed}) l_t(\tilde{X}, \theta_0^{ed}, 0). \quad (2.11)$$

where $\pi_1(\tilde{X}, t, p, \theta_1^{ed})$ is the flow net profit of firms with productivity p offering skill price θ_1^{ed} and health insurance when hiring a worker with characteristics (\tilde{X}, t)

$$\pi_1(\tilde{X}, t, p, \theta_1^{ed}) = \exp(e_w^*(ed, type, E_t)) \left(\exp(p) \exp(e_h^*(h_t)) - \theta_1^{ed} \right) - E[m_*^{\tilde{X}, t} | INS(p) = 1]. \quad (2.12)$$

The second term of the right-hand side is the expected medical expenditure of the employed worker having characteristics X and age t at firms with productivity p offering health insurance. $\pi_0(\tilde{X}, t, p, \theta_1^{ed})$ is the flow of net profit of firms with productivity p offering skill price θ_1^{ed} but not offering health insurance by hiring a worker with characteristics (\tilde{X}, t) :

$$\pi_0(\tilde{X}, t, p, \theta_0^{ed}) = \exp(e_w^*(ed, type, E_t)) \left(\exp(p) \exp(e_h^*(h_t)) - \theta_0^{ed} \right). \quad (2.13)$$

I assume that ϵ^{ESHI} follows an i.i.d. Type-I extremum value distribution with scale parameter σ_f . As a result, the fraction of firms with productivity p offering health insurance is characterized by

$$\Delta(p) = \frac{\exp\left(\frac{\Pi_1(p)}{\sigma_f}\right)}{\exp\left(\frac{\Pi_0(p)}{\sigma_f}\right) + \exp\left(\frac{\Pi_1(p)}{\sigma_f}\right)}. \quad (2.14)$$

Steady state equilibrium

Now, I am in a position to define an equilibrium.

Definition 1. *A steady state equilibrium consists of workers' value functions $\{V_0^t, V_1^t\}$ and corresponding policy functions, the steady state measure of workers with characteristics of $\tilde{X} = (ed, type, E, h)$, $g_t(\tilde{X}, \theta, INS)$ and $u_t(\tilde{X})$, the firms' compensation packages consisting of skill price $\{\theta_1^{ed}(p), \theta_0^{ed}(p)\}$ and insurance offer $\{\Delta(p)\}$ for all p , and offer distribution $F^{ed}(\theta, INS)$ such that*

1. *Given offer distribution $F^{ed}(\theta, INS)$, the value functions at age t $\{V_0^t, V_1^t\}$ and corresponding policy functions solve (2.4), (2.5), and (2.6).*

2. *Given worker's optimization behavior described by $\{V_0^t, V_1^t\}$ and corresponding policy functions and offer distribution $F^{ed}(\theta, x)$, $g_t(\tilde{X}, \theta, INS)$ and $u_t(\tilde{X})$ must satisfy*

(2.7), (2.8) and (2.9).

3. Given $F^{ed}(\theta, INS)$ and the steady state employee sizes implied by $g_t(\tilde{X}, \theta, INS)$ and $u_t(\tilde{X})$, a firm with productivity p chooses to offer health insurance with probability $\Delta(p)$, where $\Delta(p)$ is given by (2.14). Moreover, conditional on insurance choice x , the firm offers skill price $\{\theta_{INS}^{ed}(p)\}$ that solves (2.10) and (2.11) respectively for $INS \in \{0, 1\}$.

4. The postulated distributions of offered compensation packages are consistent with the firms' optimization behavior $\{\theta_1^{ed}(p), \theta_0^{ed}(p), \Delta(p)\}$. Specifically, $F^{ed}(\theta, INS)$ must satisfy

$$F^{ed}(\theta, 1) = \int_0^\infty \mathbf{1}(\theta_1^{ed}(p) < \theta^{ed}) \Delta(p) d\Gamma(p), \quad (2.15)$$

$$F^{ed}(\theta, 0) = \int_0^\infty \mathbf{1}(\theta_0^{ed}(p) < \theta^{ed}) [1 - \Delta(p)] d\Gamma(p). \quad (2.16)$$

for each education group $ed \in \{NC, C\}$.

Characterization of equilibrium

I can characterize the firm's optimal skill price $\theta_{INS}^{ed}(p)$ by extending an approach by Bontemps, Robin, and Van den Berg (1999) and Bontemps, Robin, and Van den Berg (2000). By applying an envelope condition to (2.10), we obtain that

$$\Pi_1'(p) = \sum_t \sum_{\tilde{X}} \left(e_{\tilde{X}}(p) - \frac{\partial E[m_*^{\tilde{X}, t} | HI = 1]}{\partial p} \right) l_t(\tilde{X}, \theta_1^{ed}(p), 1).$$

Integrating over $[\underline{p}_1, p]$, where \underline{p}_1 represents the lowest productivity firms that hire a positive number of workers, we obtain that

$$\Pi_1(p) = \Pi_1(\underline{p}_1) + \int_{\underline{p}_1}^p \sum_t \sum_{\tilde{X}} \left(e_{\tilde{X}}(p') - \frac{\partial E[m_*^{\tilde{X}, t} | INS = 1]}{\partial p'} \right) l_t(\tilde{X}, \theta_1(p'), 1) dp'.$$

By equating this with (2.10), one can characterize $\theta_1^{ed}(p)$ and $\theta_0^{ed}(p)$ as follows:

Proposition 2. For $p > \underline{p}$, $\theta_1^{ed}(p)$ and $\theta_0^{ed}(p)$ satisfy

$$\theta_1^{ed}(p) = \frac{\sum_t \sum_{\tilde{X}} \left(e_{\tilde{X}}(p) - E[m_*^{\tilde{X},t}] \right) l_t \left(\tilde{X}, \theta_1^{ed}(p), 1 \right) - \Pi_1(\underline{p}_1)}{\sum_t \sum_{\tilde{X}} \exp(e_w^*(ed, type, E_t)) l_t \left(\tilde{X}, \theta_1^{ed}(p), 1 \right)} - \frac{\int_{\underline{p}_1}^p \sum_t \sum_{\tilde{X}} \left(e_{\tilde{X}}(p') - \frac{\partial E[m_*^{\tilde{X},t} | INS=1]}{\partial p'} \right) l_t \left(\tilde{X}, \theta_1^{ed}(p'), 1 \right) dp'}{\sum_t \sum_{\tilde{X}} \exp(g_w^*(ed, type, E_t)) l_t \left(\tilde{X}, \theta_1^{ed}(p), 1 \right)} \quad (2.17)$$

$$\theta_0^{ed}(p) = \frac{\left(\sum_t \sum_{\tilde{X}} e_{\tilde{X}}(p) l_t \left(\tilde{X}, \theta_1^{ed}(p), 0 \right) - \Pi_0(\underline{p}_0) - \int_{\underline{p}_0}^p \sum_t \sum_{\tilde{X}} e_{\tilde{X}}(p') l_t \left(\tilde{X}, \theta_0^{ed}(p'), 0 \right) dp' \right)}{\sum_t \sum_{\tilde{X}} \exp(e_w^*(ed, type, E_t)) l_t \left(\tilde{X}, \theta_1^{ed}(p), 0 \right)}. \quad (2.18)$$

This form can be utilized when we numerically solve the equilibrium of the model. $\theta_{INS}^{ed}(\underline{p}_1)$ and $\theta_{INS}^{ed}(\underline{p}_0)$ must be solved by maximizing (2.10) and (2.11) without relying on (2.17) and (2.18).

Due to the complexity of the model, I cannot solve the equilibrium analytically. I instead solve the equilibrium numerically using the algorithm described in Appendix B.1. The complexity of the model also prevents me from establishing a proof of the existence and uniqueness of the equilibrium, but, by conducting extensive numerical simulations, I always find a unique equilibrium based on this algorithm.

2.3 Data Set

In this section, I describe my data set and its sample selection. The model describes rich individual-level dynamics over the life cycle regarding labor market outcome, health status, health insurance coverage, and medical expenditure. Moreover, it describes how firms with different productivity decide wage and health insurance provisions which determine the size of the firm. Therefore, an ideal data set for estimating the model is to use single employee-employer matched data which contains this information. However,

such a data set is not available in the U.S. Instead, I combine three separate data sets for the estimation: (1) 2004 Survey of Income and Program Participation; (2) 2004-2007 Medical Expenditure Panel Survey; and (3) Kaiser Family 2004-2007 Employer Health Benefit Survey. I choose the data period 2004-2007 because estimating the model using the data after 2008 is not ideal for the following two reasons. First, the Great Recession has generated dramatic changes in the labor market. Second, possibly due to the policy announcement effect, there was a sharp jump in the health insurance offering rate in 2010, which disappears after 2011. These short-term dramatic changes are difficult to capture with my model. Instead, I choose 2004-2007, when the economic environment was relatively stable.

2.3.1 Survey of Income and Program Participation

I obtain an individual-level labor market outcome and associated health status, health insurance coverage status, and demographic information from the 2004 Panel of Survey of Income Program Participation (hereafter, SIPP 2004). The SIPP 2004 interviews individuals every four months for up to twelve times, so that an individual may be interviewed over a four-year period. It consists of two parts: (1) the core module, and (2) the topical module. The core module, which is based on the interviews from each wave, contains detailed monthly information regarding individuals' demographic characteristics and labor force activity, including earnings, number of weeks worked, average hours worked, and employment status, as well as whether the individual changed jobs during any month in the survey period. In addition, at each interview date the SIPP 2004 gather a variety of health insurance variables. It specifies the source of insurance so we know whether it is ESHI, private individual insurance, or Medicaid, and we also know whether it is obtained through the individual's own name or the spouse's name. The topical module, which is based on annual interviews (i.e., at interview waves 3, 6, 9 and 12), contains yearly information about the worker and his/her family members' health status and out-of-pocket

medical expenditure.²² To link individual health status with individual labor market outcomes and health insurance coverage status information, I match the core module with the topical module.

I construct the estimation sample as follows. The total sample size after matching the topical module and the core module is 131,532. I restrict my sample to men (dropping 71,283 female individuals) whose age are between 25-59 (dropping an additional 33,652 individuals). In addition, I only keep individuals who are not in school, are not self-employed, do not work in the public sector, do not engage in military service and do not participate in any government welfare program (dropping an additional 11,433 individuals). I also limit our sample to individuals who are either uninsured or covered by ESHI in their own name (dropping an additional 1,949 individuals). Finally, I exclude individuals receiving Social Security income (dropping an additional 95 individuals). The sample size for the estimation is a total of 11,797 individuals.

2.3.2 Medical Expenditure Panel Survey (MEPS)

The SIPP data set allows me to capture the dynamics of health insurance coverage driven by the labor market mobility, one of the main drivers determining individual insurance status under the pre-ACA health insurance system. However, a problem with using SIPP data for my estimation is the lack of information about total medical expenditure. To obtain the information, I use the Medical Expenditure Panel Survey (hereafter, MEPS) 2004-2007. MEPS is a set of large-scale annual rotating panel surveys. I use its Household Component (HC), which surveys households in two consecutive years, collecting detailed information for each person in the household on demographic characteristics, health conditions, health status, use of medical services, charges and source of payments, access to care, satisfaction with care, health insurance coverage, income, and employment. To

²²In both SIPP and MEPS, I use the self-reported health status to construct whether the individual is healthy or unhealthy. The self-reported health status has five categories. I categorize “Excellent”, “Very Good” and “Good” as Healthy (*H*) and “Fair” and “Poor” as Unhealthy (*U*).

construct the estimation sample, I use the same criteria as SIPP 2004. The sample size for the estimation is a total of 17,536 individuals.

2.3.3 Kaiser Family Employer Health Benefits Survey

Finally, I obtain firm-side information about the health insurance offering status and associated firm characteristics from the Kaiser Family 2004-2007 Employer Health Benefit Survey.²³ It is an annual survey of the nation's private and public firms having three or more workers. It contains information about firms' characteristics (such as industry and size) and categorical information about employees' demographics (such as age and annual wage), as well as information about health insurance (such as whether the employer offers health insurance, the type of plan offered, employees' eligibility and enrollment, and whether the employer, the employee, or both contribute to the purchase of insurance.) I restrict the sample to firms which belong to the private sector and have at least three employees. The estimation sample size is 18,593.²⁴

2.3.4 Descriptive Statistics

Table 2.1 reports the descriptive statistics of key variables in SIPP 2004. It is shown that the employed who receive health insurance receive a higher wage than the employed who do not. Moreover, they are slightly healthier than individuals who do not have health insurance.

In Table 2.2, I report the descriptive statistics for the individuals in MEPS 2004-2007. Note that the proportion of healthy workers is similar to that reported in SIPP. The proportion of individuals who are insured is somewhat under-reported relative to SIPP.

I provide the descriptive statistics for firm-side data based on Kaiser 2004-2007 in

²³Note that other studies investigating ESHI often use Robert Wood Johnson Foundation Employer Health Insurance Survey (e.g., Cebul, Rebitzer, Taylor, and Votruba (2011) and Aizawa and Fang (2013)). The unit of observation is the establishment. However, the survey has not conducted after 1997.

²⁴The data also shows whether firms consist of a single establishment or not. In my data selection, 90.15% of firms consist of a single establishment.

Variable	Mean	Std. Dev.
Fraction of workers who are college graduates	0.4521	0.4977
Average worker age	40.5462	8.8822
Fraction of insured among employed	0.8409	0.3658
Average 4-month wage for employed, in \$10,000	1.9383	2.0832
Average 4-month wage for insured, in \$10,000	2.1445	2.1896
Average 4-month wage for uninsured, in \$10,000	0.8483	0.7230
Employment rate	0.9111	0.2847
Fraction of healthy workers	0.9338	0.2487
Fraction of healthy workers among insured	0.9501	0.2178
Fraction of healthy workers among uninsured	0.8771	0.3284

Table 2.1: Summary Statistics: SIPP 2004.

Variable	Mean	Std. Dev.
Average worker age	41.0386	9.7948
Fraction of healthy workers	0.9116	0.2839
Fraction of insured among employed workers	0.7244	0.4468
Annual medical expenditure, in \$10,000	0.2478	0.9583

Table 2.2: Summary Statistics for MEPS 2004-2007.

Table 2.3. In general, firms that offer health insurance tend to have many employees. Moreover, the composition of employees systematically differs by firm’s health insurance offering status. First, firms offering health insurance tend to employ a larger share of high income employees. Moreover, firms offering health insurance also consist of a larger share of older workers.

2.4 Estimation

2.4.1 Identification

In this section, I first discuss identification of several key parameters of the model. Due to the complexity of the model, the argument is mainly heuristic. The key parameters related to health and health insurance are risk aversion, γ_{type} , the disutility from bad health, η_h^{type} , productivity loss due to bad health, α_h , a scale parameter for firms offering ESHI, σ_f , and the fixed cost of offering ESHI, ξ_{ESHI} . These parameters all affect firm’s health insurance coverage and it may be unclear how each parameter is separately identified.

Variable Name	Mean	Std. Dev.
Average firm Size	20.8198	53.8675
... for those that offer health insurance	29.6444	67.7277
... for those that do not offer health insurance	7.8903	12.3044
Health insurance coverage rate	0.5943	0.4911
... for those with less than 50 workers	0.5686	0.4954
... for those with 50 or more workers	0.9338	0.2486
Average frac. of employees more than age 26	0.8364	0.2093
... for those that offer health insurance	0.8636	0.1519
... for those that do not offer health insurance	0.7991	0.2650
Average frac. of employees annual salaries \leq \$21,000	0.2108	0.3107
... for those that offer health insurance	0.1244	0.2328
... for those that do not offer health insurance	0.3288	0.3621
Average frac. of employees annual salaries \geq \$50,000	0.2345	0.2833
... for those that offer health insurance	0.2739	0.2935
... for those that do not offer health insurance	0.1807	0.2601

Table 2.3: Summary Statistics: Kaiser 2004-2007.

Note: Standard deviations are in parentheses.

While it is true that higher $\gamma_{type}, \eta_h^{type}$, and lower α_h increase firm's coverage rate, γ_{type} and η_h^{type} affect worker-side moments such as life cycle patterns of coverage rate and associated job-to-job transitions and medical expenditure. Specifically, if γ_{type} and η_h^{type} are larger, we would expect to observe more frequent transitions of workers from jobs without health insurance to jobs with health insurance, even if the transition involves a reduction in wages. Moreover, the magnitude of the wage cut a worker is willing to tolerate in order to switch from a job without health insurance to a job with health insurance increases with γ_{type} and η_h^{type} . Then, η_h^{type} is disciplined by fitting medical expenditure: because medical expenditure is a choice variable in the model and the benefit of medical expenditure is to improve future health status, I can identify the disutility from bad health through the variation of the fraction of positive medical expenditure by individual health insurance status conditional on individual characteristics, e.g., age. Finally, heterogeneity of γ_{type} and η_h^{type} by unobserved type can be identified by differential patterns of the above moments by education status, as the type distribution differs across education.

The remaining parameters are then identified in the following way. The scale parameter σ_f affects the relationship between the probability of offering health insurance and

firm productivity (and thus firm size). The parameter α_h has an additional effect on the differences in wages offered by firms, depending on whether they offer health insurance. This therefore affects the wage difference between the insured and the uninsured. Finally, the administrative cost ξ_{ESHI} is identified from the the probability (in the level) of small firms offering health insurance.

Finally, I discuss additional identification assumptions related to labor market parameters. As in standard labor market equilibrium models, it is impossible to identify intercepts in individual skill function described by (2.2) from skill prices. Therefore, I assume that the mean of log normal distribution for firm productivity distribution μ_p is set to be 0. Furthermore, I assume that the intercept for type 2 individual dummy α_1^2 is equal to 0.

2.4.2 Estimation Strategy

The solution of the model serves as an input to the estimation procedure. Estimation is by the method of simulated moments (MSM). Specifically, a weighted average distance between sample moments and simulated moments is minimized with respect to the model's parameters. The weights are the inverses of the estimated variances of the moments. The procedure requires a choice of moments.

The following is a list of moments used in the estimation. Each moment on the worker-side is a conditional moment by education and age cohort.

1. Labor market status and its dynamics
 - (a) employment rate
 - (b) transition rate from non-employment to employment
 - (c) job-to-job (JJ) transition rate
 - (d) transition rate from employment to non-employment

2. Wage and health insurance

- (a) the fraction of the uninsured among employed
- (b) wage change through JJ conditional on before-after health insurance coverage status
- (c) the distribution of wage conditional on health insurance coverage status
- (d) wage change of job stayers conditional on health insurance coverage status
- (e) the distribution of wage among previously unemployed workers, conditional on health insurance coverage status

3. Health and medical expenditure

- (a) health status conditional on employment and health insurance status
- (b) annual health transition conditional on health
- (c) annual health transition conditional on health and health insurance status
- (d) annual medical expenditure conditional on health
- (e) annual medical expenditure conditional on health and health insurance
- (f) the fraction of zero medical expenditure conditional on health and health insurance

4. Firm characteristics

- (a) the fraction of firms with less than 50 workers
- (b) health insurance coverage rate by whether firm size is less than 50 workers
- (c) mean firm size conditional on health insurance offering coverage²⁵

²⁵One could incorporate moments regarding within-firm worker distribution. However, the available demographic information (e.g., the fraction of employees whose age is less than 26, the fraction of employees earning more than \$15 per hour, etc.) is very coarse. Therefore, I decide not to include them as a target. However, the model prediction of these moments (specifically conditional moments) is qualitatively consistent. The prediction is available from the author upon request.

The estimation is done in the standard nested fixed point algorithm and described as follows:

1. Guess a vector of parameters q .
2. Given the parameters q , solve the equilibrium of the model and then simulate the data.
3. Evaluate the objective function using the simulated moments:

$$\min_{\{q\}} G(q)' \Omega G(q),$$

where $G(q)$ is the vector of the value of each moment j , $G_j(q)$. Each moment $G_j(q)$ is constructed as

$$G_j(q) = \tilde{G}_j - \mu_j(q)$$

where \tilde{G}_j is the sample moment of j and $\mu_j(q)$ is the simulated moment.

4. Repeat steps 1-3 and find q to minimize the objective value.

I compute asymptotic standard errors following Gourieroux, Monfort, and Renault (1993). Certain parameters of the model are calibrated without using the model. In this study, I do not try to estimate β but set $\beta = 0.99$ so that the annual interest rate is about 3%.²⁶ Moreover, the population growth rate n is estimated using the SIPP 2004 sample by running a regression of cohort size on age. Estimates are $n = 1.0005$ per four-month period. Finally, the after-tax income schedule is specified by following Kaplan (2012) who approximates the U.S. income tax code as $T(w) = \tau_0 + \tau_1 \frac{w^{(1+\tau_2)}}{1+\tau_2}$. Kaplan (2012) estimates the parameters as $\tau_0 = 0.0056$, $\tau_1 = 0.6377$ and $\tau_2 = -0.1362$.

²⁶ It is known from Flinn and Heckman (1982) that it is difficult to separately identify the discount factor β from the flow unemployed income b in standard search models.

2.5 Estimation Result

The parameter estimates are summarized in the Tables 2.4-2.8. Table 2.4 reports the parameter estimates for individual preferences. Table 2.4 shows that the CARA coefficient is very heterogenous across different types. In my empirical analysis, I consider the case where type 1 individuals are more risk averse.²⁷ If we convert this into relative risk aversion, we get 2.99 for the very high income group of type 1 workers, while it is 1.67 for type 2 workers. These estimates are much lower than the standard estimates in consumption/saving literature, but consistent with the literature of labor supply elasticity.²⁸ On the other hand, the disutility from bad health is relatively homogenous across different types. I also find that the disutility of working is increasing over ages. Moreover, the coefficient on the interaction between disutility from working and bad health is positive, implying that old unhealthy individuals suffer higher disutility from working relative to young unhealthy individuals.

Parameter	Estimates	Std. Error
γ_1	1.2973	(0.038)
γ_2	0.7259	(0.033)
η_u^1	0.0360	(0.001)
η_u^2	0.0380	(0.001)
η_p	0.0005	(0.00001)
η_{hp}	0.0223	(0.00003)
b	0.0891	(0.001)
v_U^T	-4.0711	(2.003)

Table 2.4: Parameter Estimate for Individual Preference

Next, Table 2.5 shows the parameter estimates for latent medical expenditure shocks and health transition processes. I find that the distribution of latent medical expenditure shocks differs substantially between healthy and unhealthy individuals. For example, al-

²⁷Note that I do not impose any restrictions in terms of the correlation of risk type with other characteristics such as labor market skills, health transition, and disutility from bad health. These correlations are estimated to fit the data.

²⁸For example, French and Jones (2011) estimate the CRRA coefficient as greater than 8 in his life-cycle model of labor supply and saving for older workers. On the other hand, Chetty (2006) shows that the upper bound of the relative risk aversion coefficient for the average worker should be 2 if one estimates it from labor supply behavior, under the assumption that both college graduates and non-college graduates share a common parameter value.

though the constant term for healthy individuals in health shock process is much smaller than the constant term for unhealthy individuals, the age coefficient in the health shock process for healthy individuals is much larger than the age coefficient for unhealthy individuals. Moreover, the health transition process differs substantially between type 1 and type 2 workers. I find that type 1 workers are more likely to transition to being healthy in the next period relative to type 2 workers.

Parameter	Estimates	Std. Error
ω_1^H	-2.8652	(0.004)
ω_2^H	0.0097	(0.0001)
ω_3^H	1.7001E-07	(1.0E-08)
σ_H	0.5613	(0.003)
κ^H	0.0684	(0.0004)
ω_1^U	-1.9286	(0.008)
ω_2^U	0.00001	(0.00001)
ω_3^U	0.0000	(1.0E-08)
σ_U	1.7218	(0.005)
κ^U	0.0000	(0.0005)
ϕ_{1H}^H	3.2557	(0.004)
ϕ_{1U}^U	0.4002	(0.001)
ϕ_{2U}	-0.3895	(0.001)
ϕ_{3U}	0.000015	(0.000002)
ϕ_{4U}	0.004	(0.001)
ϕ_{5U}	0.1530	(0.002)
ϕ_{6U}	-0.00016	(0.00002)
ϕ_{7U}^U	0.0025	(0.0013)

Table 2.5: Parameter Estimate for Latent Medical Expenditure Shocks and Health Transition Process

Table 2.6 shows the parameter estimates for the individual labor market environment. An important finding is that type 1 workers, who are more risk averse and are more likely to be healthy, are more productive than type 2 workers. Another important finding from Table 2.6 is that the productivity loss of being unhealthy, α_h , is -0.337, which means that unhealthy workers produce 71.39% of the output of healthy workers.

Table 2.7 shows the parameter estimates for the distribution of workers. It shows that the fraction of type 1 workers among the college graduates is larger than that among the non-college graduates. Therefore, by combining results from Tables 2.4, 2.5, and 2.6,

Parameter	Estimates	Std. Error
α_1^C	-0.8008	(0.002)
α_1^{NC}	-1.1200	(0.002)
α_1^1	0.3668	(0.002)
α_4^C	0.00083	(0.0003)
α_4^{NC}	3.91E-05	(1.5E-05)
α_4^1	2.16E-06	(5.5E-06)
α_5^C	1.0E-09	(1.16E-08)
α_5^{NC}	-1.5E-06	(1.90E-07)
α_5^1	1.0E-08	(2.1E-07)
α_h	-0.3370	(0.004)
σ_n	0.0126	(0.0003)
λ_u^C	0.4340	(0.001)
λ_e^C	0.1360	(0.001)
δ^C	0.0101	(0.0001)
λ_u^{NC}	0.2048	(0.001)
λ_e^{NC}	0.0874	(0.001)
δ^{NC}	0.0244	(0.0001)
σ_w	0.0100	(0.0002)

Table 2.6: Parameter Estimate for Individual Labor Market Activities

parameter estimates show that the college graduates are on average more risk averse, more productive, and more healthy than non-college graduates in the model. Moreover, the total measure of workers (relative to the measure of firms) is estimated to be 22.2050 to fit the mean of firm size observed in the data.

Parameter	Estimates	Std. Error
The frac. of type 1 workers among the college	0.4045	(0.004)
The frac. of type 1 workers among the non-college	0.2715	(0.003)
M	22.2050	(1.26)

Table 2.7: Parameter Estimate for the Distribution of Workers

Finally, Table 2.8 shows parameter estimates for firm-side characteristics. Note that the scale parameter of firm productivity distribution is normalized to 0 and therefore only the shape parameter is estimated. The key finding here is that the fixed cost of offering ESHI is estimated to be \$552 per four-month, which is equivalent to \$1,656 annually.

Parameter	Estimates	Std. Error
σ_p	0.8517	(0.002)
ξ_{ESHI}	0.0552	(0.001)
σ_f	0.3587	(0.0008)

Table 2.8: Parameter Estimate for Firm-side Characteristics

2.5.1 Model Fit

In this section, I first provide tables about model fits to show that the model can fit the most salient features of the data.

Table 2.9 shows the model fit for the average wage conditional on health insurance status, age, and education group. In the data, there is a strong positive correlation between wage and age for college graduates for individuals who have health insurance. On the other hand, wage-age slopes in other groups are rather flat. The model is able to fit these moments well.

	College graduate				Non-college graduate			
	with <i>HI</i>		w/o <i>HI</i>		with <i>HI</i>		w/o <i>HI</i>	
Age	Data	Model	Data	Model	Data	Model	Data	Model
25-31	1.8974	1.8695	0.8753	0.8680	1.2624	1.4014	0.7500	0.7772
32-38	2.5425	2.3713	0.9752	1.0004	1.4231	1.4741	0.8915	0.8012
39-45	2.7911	2.7291	0.9603	0.9815	1.6000	1.5107	0.8607	0.8468
46-52	2.9782	3.0157	1.0895	0.9501	1.6467	1.5203	0.8252	0.8461
53-59	2.8987	2.9818	0.9935	0.9065	1.5719	1.5723	0.7103	0.6754

Table 2.9: Model fit: mean 4-month wage conditional on coverage, age and education.

Notes: (a) the unit is \$ 10,000. (b) with HI indicates employed individuals with ESHI; w/o HI indicates employed individuals are uninsured.

Next, Table 2.10 shows the pattern of health insurance coverage status among the employed workers over ages. The data shows a positive correlation between coverage rate and age, regardless of education group. The model captures the insurance gain over the life cycle quite well. The model somewhat underpredicts the coverage rate among college graduates in the age group 25-31, while it overpredicts the coverage rate among non-college graduates in the age group 25-31. Table 2.11 shows the model fit for the employment rate across each of the age groups. The model can quantitatively explain the age profile of the employment rate for both education groups.

Age	College graduate		Non-college graduate	
	Data	Model	Data	Model
25-31	0.9045	0.8516	0.6557	0.7494
32-38	0.9494	0.9417	0.7621	0.8005
39-45	0.9516	0.9653	0.8235	0.8478
46-52	0.9527	0.9766	0.8843	0.9028
53-59	0.9607	0.9786	0.8911	0.9185

Table 2.10: Model fit: coverage rate among the employed conditional on age and education.

Age	College graduate		Non-college graduate	
	Data	Model	Data	Model
25-31	0.9788	0.9692	0.9175	0.9282
32-38	0.9787	0.9795	0.9380	0.9278
39-45	0.9693	0.9809	0.9262	0.9268
46-52	0.9624	0.9757	0.9207	0.9129
53-59	0.9504	0.9678	0.9129	0.9094

Table 2.11: Model fit: employment rate conditional on age and education.

Next, Table 2.12 shows the model fit for health status. It reports the fraction of healthy individuals across age groups and education status. One striking pattern observed in the data is that the difference between the fraction of college graduates and non-college graduates that are healthy increases over time. The model quantitatively accounts for this pattern well.

Age	College graduate		Non-college graduate	
	Data	Model	Data	Model
25-31	0.9889	0.9613	0.9585	0.9278
32-38	0.9784	0.9523	0.9397	0.8968
39-45	0.9590	0.9362	0.9229	0.8692
46-52	0.9564	0.9185	0.8820	0.8857
53-59	0.9182	0.8576	0.8223	0.7822

Table 2.12: Model fit: health status conditional on age and education.

Note: the number in the table is the fraction of individuals who are healthy.

Table 2.13 reports the model fit for the pattern of health status conditional on insurance and employment status. It is shown that the employed workers with ESHI are the most healthy; the employed who do not have health insurance are less healthy and the non-employed are the least healthy. The model somewhat over-predicts the fraction of

the healthy among the non-employed, but it accounts for these qualitative patterns.

	Emp. with <i>HI</i>		Emp. w/o <i>HI</i>		Non-employed	
	Data	Model	Data	Model	Data	Model
College graduate	0.9354	0.9507	0.8796	0.8953	0.6803	0.7851
Non-College graduate	0.9009	0.9222	0.8365	0.8632	0.6458	0.7311

Table 2.13: Model fit: mean health status conditional on insurance, employment and education.

Table 2.14 shows the model fit for the health transition rate. The data demonstrates increasing persistence of transitions from unhealthy to unhealthy. Moreover, there is a stark difference in health transitions between college graduates and non-college graduates. The table shows that the model is able to capture these patterns.

Table 2.15 shows how well the model fits medical expenditure patterns conditional on health status. One interesting pattern in the data is that the variation of medical expenditure due to health status is much larger for older individuals than for younger. While the model tends to produce a steeper relationship between age and medical expenditure relative to the data, it captures the overall pattern of the data reasonably well.

Finally, Table 2.16 shows the model fit of firm-side moments. It demonstrates that the model fits reasonably well with the data. Specifically, it fits remarkably well for coverage rate and firm size distributions.

Overall, the model does a good job of quantitatively explaining most salient features of health, health insurance, and labor market outcomes. While the model contains various mechanisms generating these outcomes, one of the most important mechanisms is as follows. First of all, the model estimates predict that high productivity firms tend to offer health insurance. In the model, firms want to attract more productive workers. High productivity firms are more likely to attract such workers because they can offer higher compensation. These more productive workers are typically experienced workers, who tend to be older and thus have a higher demand for health insurance. College graduates are another type of productive workers. From my estimates, college graduates are more risk-averse and have a higher demand for health insurance than less educated

Age	College graduate						Non-college graduate					
	Healthy→Healthy		Unhealthy→Unhealthy		Healthy→Healthy		Healthy→Healthy		Unhealthy→Unhealthy		Unhealthy→Unhealthy	
	Data	Model	Data	Model	Data	Model	Data	Model	Data	Model	Data	Model
25-31	0.9831	0.9803	0.2000	0.1987	0.9711	0.9666	0.1539	0.1539	0.2446	0.2446	0.1539	0.2446
32-38	0.9913	0.9778	0.1667	0.1919	0.9638	0.9546	0.3056	0.3056	0.3194	0.3194	0.3056	0.3194
39-45	0.9716	0.9697	0.4231	0.2123	0.9514	0.9413	0.3929	0.3929	0.4058	0.4058	0.3929	0.4058
46-52	0.9781	0.9617	0.1667	0.2149	0.9332	0.9445	0.4729	0.4729	0.3910	0.3910	0.4729	0.3910
53-59	0.9444	0.9515	0.3750	0.2839	0.8669	0.9234	0.5625	0.5625	0.5297	0.5297	0.5625	0.5297

Table 2.14: Model fit: Annual health transition

	College graduate				Non-college graduate			
	healthy		unhealthy		healthy		unhealthy	
Age	Data	Model	Data	Model	Data	Model	Data	Model
25-39	0.1177	0.0875	1.0769	0.8367	0.0688	0.0596	0.5168	0.5931
40-49	0.2032	0.1861	0.9329	1.0988	0.1242	0.0899	0.6964	0.7269
50-59	0.3857	0.3571	1.2109	1.0727	0.2164	0.2219	0.5638	0.5455

Table 2.15: Model fit: mean annual medical expenditure conditional on age and education and health.

Note: The unit is \$ 10,000.

Variable Name	Data	Model
Average establishment Size	20.8198	20.9779
... for those that offer health insurance	29.6444	31.6429
... for those that do not offer health insurance	7.8903	6.1163
The frac. of firms having less than 50 workers	0.9295	0.8858
Health insurance Coverage Rate	0.5943	0.5957
... for those having less than 50 workers	0.5686	0.5486
... for those having 50 or more workers	0.9338	0.9612

Table 2.16: model fit for firm-side moment

workers. The consequence of sorting of workers with high health insurance demand leads high-productivity firms to offer ESHI. This then leads to a positive correlation between firm size and the rate of firms offering ESHI, as high productivity firms attract more workers and become larger. Furthermore, the model also explains why the coverage rate increases over age. By moving from low to high productivity firms through job-to-job transitions over their life cycles, workers gain health insurance. Overall, this mechanism simultaneously explains positive correlations among workers' age, wage, education status, and health insurance coverage and correlations among firm size and ESHI offering rate.²⁹

²⁹In addition to this mechanism, the model has several additional mechanisms that contribute toward generating a higher rate of firms offering ESHI by large firms. These include tax-deductibility of ESHI and an improvement of health, and therefore labor productivity, through more usage of health care (conditional on health shocks). Given my estimates, these mechanisms quantitatively play a minor role relative to the mechanism described in the main text. The quantitative comparison is available on request.

2.6 Counterfactual Experiments: Evaluating the ACA and the components of HIX

In this section, I use the estimated model to evaluate the impact of ACA and each component of HIX and then study the optimal design of HIX. To do so, I first consider a stylized version of the ACA as a benchmark counterfactual environment, and then evaluate the role of HIX in such a context. I consider a stylized version of the ACA which incorporates its five main components: first, all individuals are required to have health insurance or pay a penalty (individual mandate); second, all firms with more than 50 workers are required to offer health insurance or pay a penalty (employer mandate); third, HIX are established where individuals can purchase health insurance at a modified community rated premium; fourth, the individuals purchasing health insurance from HIX can obtain income-based subsidies; fifth, Medicaid is expanded.

The introduction of HIX requires a substantial departure from the pre-ACA model because the premium in HIX will be endogenously determined. As a result, I will first describe how I extend and analyze the pre-ACA model to incorporate HIX.

2.6.1 The Model for evaluating the ACA

I provide a brief explanation of the main changes in the economic environment, as well as the definition of equilibrium, for the model used to evaluate the ACA and components of HIX.

The Main change in individuals' environment

I assume that individuals who are not offered health insurance by their employers and those who are non-employed can purchase individual insurance from HIX. I also assume that the insurance product offered from HIX is the same as that offered by the employers, in that it also fully insures medical expenditure risk. Moreover, as in the pre-ACA econ-

omy, health insurance premia for ESHI are pre-paid by firms before an individual accepts a job offer from the firm. Therefore, there is no incentive for the workers not to take the offer. Thus in the extended model, an individual's insurance status INS is defined as:

$$INS = \begin{cases} 0 & \text{if uninsured} \\ 1 & \text{if insured through ESHI} \\ 2 & \text{if insured through HIX.} \end{cases}$$

I also incorporate the premium subsidies to the individuals and penalties if uninsured into the model. Let $S(y, R^{HIX}(t))$ denote income-based subsidies to an individual with income y who purchases health insurance from HIX at premium $R^{HIX}(t)$; note that the subsidy amount does not explicitly depend on individual age t under the specification of the ACA. Similarly, let $IM(y)$ denote the penalty to individuals who remain uninsured, which is merely function of income under the ACA.

$$\begin{aligned} C_t &= \tau_w(w_t)P_t + (1 - P_t)b - OOP^{HI}(x_t m_t) \\ &\quad - \mathbf{1}(x = 0)IM(y) - \mathbf{1}(x = 2)(R^{HIX}(t) - S(y, R^{HIX}(t))) \end{aligned} \quad (2.19)$$

where $y = w_t$ if employed ($P_t = 1$) at firms offering (θ, INS) and $y = b$ otherwise ($P_t = 0$).

Modeling individual decisions of health insurance purchases from HIX requires me to modify the timing in a period specified in Section 2.2.1. I assume that the decision is made at the end of the period: after making a working decision in the labor market, an employed individual not offered health insurance can decide whether to purchase health insurance from HIX.

Of course, the introduction of HIX into the individual decision problem makes the expression of individual value functions and steady state worker distributions rather complicated. However the derivation itself is a straightforward extension of the pre-ACA version. Therefore, I introduce value functions and steady state worker distribution in

Appendix (B.2).

The main change in firms' environment

Because of the employer mandate, firms pay a penalty if they do not offer health insurance. This changes the determination of profits obtained by firms not offering ESHI. The flow profit $\Pi_0(p)$ is specified as

$$\begin{aligned} \Pi_0(p) = & \max_{\theta_0^{ed}} \sum_t \sum_{\tilde{X}} \pi_0(\tilde{X}, t, p, \theta_0^{ed}) l_t(\tilde{X}, \theta_0^{ed}, 0) \\ & - EM \left(\sum_t \sum_{\tilde{X}} (l_t(\theta_0^{ed}, 0) + l_t(\theta_0^{ed}, 2)) \right) \end{aligned}$$

where $\pi_0(\tilde{X}, t, p, \theta_0^{ed})$ is defined in the pre-ACA economy, which is expressed in (2.13), and $EM(l)$ is the tax penalty amount, which depends on firm size $\sum_{\tilde{X}} (l(\theta_0^{ed}, 0) + l(\theta_0^{ed}, 2))$. The flow profit of firms offering ESHI, $\Pi_1(p)$, as well as the decision of whether to offer ESHI, $\Delta(p)$, are determined as before, and are expressed in (2.10) and (2.14) respectively.

HIX

I assume that HIX is a competitive insurance market, as in Hackmann et al. (2013) and Handel et al. (2013).³⁰ The premium in HIX is regulated as a modified community rating. It can vary based on individual age. In equilibrium, it must satisfy

³⁰Ericson and Starc (2012), in their analysis of the HIX in Massachusetts (MA), instead assume that HIX is an imperfect insurance market. As I will explain, one of the main differences between the MA reform and the ACA is that the MA reform does not have medical loss ratio regulation while the ACA has. This limits insurance companies from charging a higher markup for insurance premia. While allowing imperfect competition in HIX allows us to study the optimal choice of medical loss ratio, such an assumption substantially complicates the whole analysis and requires the actual data on national samples of individuals purchasing health insurance from HIX, which is not available at present. Therefore, I leave this issue for future work.

$$\begin{aligned}
& \sum_t \sum_{\tilde{X}} \int g_t(\tilde{X}, \theta, 2) d\theta R^{HIX}(t) \\
&= (1 + \xi_{HIX}) \sum_t \sum_{\tilde{X}} \int E[m_*^{\tilde{X}} | INS_t = 2] g_t(\tilde{X}, \theta, 2) d\theta \quad (2.20)
\end{aligned}$$

where the left-hand side is the total premium paid by individuals purchasing health insurance from HIX, and the right-hand side is the total expected medical expenditure by those participants multiplied by the loading factor ξ_{HIX} .

Note that the health insurance premium $R^{HIX}(t)$ is subject to government regulation such that it cannot depend on individual health status. The variation of premia due to age is partially limited due to the age-based rating regulation. Specifically, the regulation determines the maximum allowable premium ratio between the oldest and the youngest ω_{AGE} :

$$\omega_{AGE} R^{HIX}(1) \geq R^{HIX}(T).$$

The definition of equilibrium

The definition of steady state equilibrium is a straight forward extension of the equilibrium under the pre-ACA economy. Specifically, it now includes premiums in HIX $\{R^{HIX}(t)\}$ as equilibrium objects. Insurance premia must satisfy a break even condition defined in (2.20).

Parameterization of policy components in the ACA

In order to evaluate the impact of ACA and the components of HIX, I need to parameterize each component. It requires me to address several issues regarding how to introduce the specifics of the ACA provisions into my model, such as the penalty associated with the individual mandate and employer mandate, the premium subsidies, age based rating regulation, and Medicaid. First, the final picture of the reform regarding Medicaid ex-

pansion is still unclear. As of the end of August 2013, 25 states plan to expand Medicaid as the ACA requires, while the others plan not to.³¹ Because the model does not describe heterogeneity in terms of states, I can only consider extreme cases: (a) all states expand Medicaid; (b) no states expand Medicaid. In this analysis, I first show the result for case (a), and then discuss how results change in case (b). Second, I need to decide on the magnitude of the loading factor ξ_{HIX} that appeared in (2.20), which is applicable in HIX. I calibrate ξ_{HIX} based on the ACA requirement that all insurance sold in the exchange must satisfy the newly-imposed regulation by the ACA that the medical loss ratio must be at least 80%.³² This implies that $\xi_{HIX} = 0.25$.³³ Third, the amount of penalties and subsidies are defined as an annual level, while my model uses four-month as a model period. I simply divide all monetary units by 3 to obtain the applicable number for a four-month period. Given these assumptions, I specify the details of the components of the ACA, which is in Appendix B.3.

2.6.2 The Main Results

The aggregate impact of the ACA

Table 2.17 shows the main results from simulating the ACA and a variety of its combinations. The results for the ACA are reported in Column (2). It reports the outcomes of several important aggregate variables determined in the model. Panel A reports the main outcomes regarding the firm side. First, I find that the ACA substantially increases firms' coverage rate from 59.6% to 62.7%. The increase in firm's coverage rate comes not only from firms with more than 50 workers which are subject to the penalty of employer

³¹On June 28, 2012, the U.S. Supreme Court's ruled unconstitutional the law's provision that, if a State does not comply with the ACA's new coverage requirements, it will lose not only the federal funding for those requirements, but all of its federal Medicaid fund. This ruling allows states to opt out of ACA's Medicaid expansion, leaving each state's decision to participate in the hands of the nation's governors and state leaders.

³²The medical loss ratio is the ratio of the total claim costs that the insurance company incurs to total insurance premium collected from participants.

³³The medical loss ratio implied by (2.20) is simply $1/(1 + \xi)$, thus an 80% medical loss ratio corresponds to $\xi = 0.25$. The ACA requires that $\xi \leq 0.25$.

	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Effects on the Firm Side						
Frac. of firms offering ESHI	0.5957	0.6268	0.5394	0.6768	0.4093	0.5993
...if firm size is less than 50	0.5486	0.5797	0.4814	0.6359	0.3382	0.5487
...if firm size is 50 or more	0.9612	0.9999	0.9999	0.9999	0.9998	0.9999
Average labor productivity	2.4883	2.4736	2.4725	2.4862	2.4664	2.4730
Output per capita	2.2483	2.2439	2.2427	2.2476	2.2430	2.2436
Panel B: Effects on the Worker Side						
Uninsured rate	0.2359	0.0788	0.1656	0.0934	0.0200	0.0842
The frac. of ind. with ESHI	0.7641	0.7687	0.7057	0.7992	0.5886	0.7686
Non-employment rate	0.0958	0.0918	0.0918	0.0941	0.0908	0.0921
Average wage	1.7859	1.7717	1.7756	1.7778	1.7802	1.7667
The fraction of employed in the top 50% productivity rank firms	0.7412	0.7332	0.7331	0.7402	0.7301	0.7335
Medical expenditure	0.0724	0.0847	0.0823	0.0838	0.0861	0.0844
The frac. of healthy workers	0.9501	0.9589	0.9572	0.9581	0.9598	0.9587
Panel C: Effects on Government Revenue:						
Revenue from Income Tax	0.6853	0.6814	0.6832	0.6726	0.6860	0.6792
Subsides to HIX & Medicaid	-	0.0139	0.0150	0.0115	0.0226	0.0166
Revenue from penalties	-	0.0016	1.63e-07	0.0034	0.0003	0.0020
Total Revenue	0.6853	0.6690	0.6682	0.6726	0.6639	0.6655

Table 2.17: Counterfactual Policy Experiments: Evaluation of the ACA and its Variations.

Notes: (a) Column (1) reports the statistics generated under the pre-ACA economy. Note that all the statistics are calculated by including individuals less than age 25. The main pattern is unchanged even if I exclude them in this table. (b) Column (2) reports the statistics generated under the ACA. (c) Column (3) reports the statistics generated under the ACA without individual mandate. (d) Column (4) reports the statistics generated under the ACA without premium subsidies in the HIX. (e) Column (5) reports the statistics generated under the ACA with additional premium subsidies which are set that individuals who are eligible for premium subsidies in HIX obtain the full premium subsidies, subsidies which are equal to premia. (f) Column (6) reports the statistics generated under the ACA without maximum premium ratio regulation between the oldest and youngest. out premium subsidies in the HIX.

mandate, but also from firms with less than 50 workers, which are not. Secondly, and more interestingly, I find that the ACA reduces labor productivity and output per worker. The aggregate labor productivity, defined as the output per employed worker, decreases by 0.6%, while output per capita decreases by 0.2%, which are substantial.

Panel B reports the main outcomes regarding the worker side. It shows that the uninsured rate in the estimated sample decreases from 23.6% to 7.8%, which is a substantial reduction. Half of the reduction is due to the Medicaid expansion, which covers 9.2% of non-employed workers. Moreover, 10.2% of workers, who are employed workers at firms not offering ESHI, have health insurance from HIX. The remaining increase is explained by the fact that more workers are covered by ESHI. It also shows that the non-employment rate decreases by 0.4 percentage point. Furthermore, the fraction of healthy workers increases by 1 percentage point.

Combining the results from Panel A and B, it is somewhat surprising to observe the decline of aggregate labor productivity and output. It is important to stress that the impact of the ACA on aggregate labor productivity and output are theoretically ambiguous. There are several mechanisms counteracting the decline of these variables. First, the fraction of healthy workers increases as more workers are insured and utilize more health care when they are hit by health shocks. Because the unhealthy individuals produce 30% less than the healthy individuals, this should increase aggregate labor productivity. Second, the non-employment rate decreases relative to the pre-ACA economy, which is a consequence of the increase in the fraction of healthy individuals, as the unhealthy individuals face higher disutility of working.³⁴ This channel has a positive effect on the output per capita. However, the main reason for observing the decrease in aggregate labor productivity and output is that more workers are allocated to low productivity firms, reported in Panel B. The fraction of employed workers in the top 50% of firms in productivity

³⁴Note that an increase in the employment rate is not a trivial result. Indeed, given the expansion of Medicaid which provides free coverage to non-employed workers, the ACA gives a strong disincentive to work, as argued by Mulligan (2013b) and Mulligan (2013a). However, I find that such a disincentive effect is smaller than the effect from improved health which leads to more labor force participation.

ranking decreases by 1.11%.

Panel C reports the main outcome regarding the government budget. It shows that government revenues decrease by 3%. The source of the declines is an expansion of subsidized coverage in HIX and an expansion of Medicaid.

The heterogeneous impact of the ACA

While Table 2.17 is mainly about aggregate outcome, it is also useful to know whether the ACA has different effects on individuals with different characteristics.

Age Group	Pre-ACA	ACA	
Age Group	(1)	(2)	(3)
25-35	0.18	0.17	0.12
35-45	0.12	0.12	0.07
45-55	0.09	0.10	0.004

Table 2.18: Counterfactual Policy Experiments: the Impact of the ACA by Age Groups. Notes: Column (1) is the measure of employed workers who are not offered ESHI under the pre-ACA; Column (2) is the measure of employed workers who are not offered ESHI under the ACA; Column (3) is the measure of individuals who are uninsured and employed.

Table 2.18 shows the uninsured rate by age cohorts. Interestingly, the third column in Table 2.18 shows that individuals older than age 45 achieve almost universal coverage. Most remaining uninsured are younger individuals. The third column in Table 2.18 shows that 10% of individuals between age 45 and 55 are working for firms not offering ESHI, while it is 17% for age 25-35. Therefore, this age difference is caused by the low participation to HIX by younger individuals who are working for firms not offering ESHI. This pattern is explained by the following two reasons. First, the maximum premium ratio between the youngest and oldest is binding in my counterfactual. Therefore, younger individuals need to pay higher premia relative to their expected medical expenditures. Second, the distribution of medical expenditure among the young is more skewed relative to the old in my estimates. As a result, HIX for the young pool suffers from more of an adverse selection problem. These effects lead to a lower proportion of the young participating in HIX. Another interesting finding is that the fraction of older workers in

firms not offering ESHI is higher in the ACA than the fraction in the pre-ACA economy. Because firms not offering ESHI tend to be low productivity firms, a higher proportion of older workers are allocated to low productivity firms. Such a pattern is documented in Figure 2.1, which shows that the fraction of older individuals in more productive firms is lower in the ACA than the fraction in the pre-ACA economy. Moreover, these older workers are typically less skilled, as seen from Figure 2.2.

Pre-ACA	ACA	
(1)	(2)	(3)
0.94	0.96	0.99

Table 2.19: Counterfactual Policy Experiments: the Impact of the ACA by Health Status. Notes: Column (1) is the fraction of healthy workers among those who are working in firms not offering ESHI under the pre-ACA; Column (2) is the fraction of healthy workers among those who are working in firms not offering ESHI under the ACA; Column (3) is the frac. healthy uninsured is the fraction of healthy workers among those who choose to stay uninsured under the ACA.

Table 2.19 shows the change in the distribution of health status among the uninsured. As shown in the table, the fraction of healthy workers who are working at firms not offering ESHI is 0.94 in the pre-ACA economy. The ACA increases it to 0.96. Moreover, those who remain uninsured are healthy individuals. Therefore, almost all unhealthy individuals are covered by health insurance.

The impact of the components of HIX

To understand the main driving forces and mechanisms generating the impact of the ACA, I show the importance of each component of HIX in the ACA, which is summarized in Table 2.17.³⁵ First, Column (3) shows the simulation result under the ACA without individual mandate. In this case, the uninsured rate more than doubles. This increase in the uninsured rate is explained by the decline in both the participation rate in HIX and the rate of firms offering ESHI. This indicates the importance of the individual mandate in this economy for achieving a lower uninsured rate. The increase in the uninsured rate

³⁵An additional set of the results showing each component of the ACA, in addition to the HIX, is available on request.

comes from a lower proportion of firms offering ESHI and lower take-up in HIX. Mainly due to an increase in the uninsured rate without individual mandates, the fraction of unhealthy individuals increases, contributing to the decline of labor productivity and output. However, the magnitude of the decline is quite small.

Column (4) shows the simulation result under the ACA without premium subsidies. Relative to the case under the ACA without individual mandate, an increase in the uninsured rate from the ACA is relatively small: it is 9.34% while it is 16.56% under the ACA without individual mandate. One reason for the increase in the uninsured rate is the lower participation in HIX, which is the effect we observed in the case of the removal of the individual mandate. However, this small response is due to the fact that the fraction of firms offering ESHI in the absence of premium subsidies is 5 percentage points higher relative to the ACA with the premium subsidies. This is due to the fact that premium subsidies decrease the ESHI offering rate. Therefore, while both an individual mandate and premium subsidies increase participation in HIX, each of the components of the ACA has a qualitatively different effect on the rate of firms offering ESHI.

An interesting finding from Column (4) is that the decline in labor productivity and the output per capita in the ACA, reported in Column (2) is accounted for by the premium subsidies. Column (2) also shows that premium subsidies account for the most of the decline in the fraction of employed workers in firms ranked in the top 50% productivity ranking.

One can see a similar pattern from Column (5). Here, I consider the case where individuals whose income is less than 400% obtain full premium subsidies so that they can obtain free health insurance from HIX. Both output and labor productivity decline, while the fraction of healthy workers and employment rate increase.

The main reason why more workers are allocated to low productivity firms is that firms cannot allow individual workers to choose between ESHI and HIX as their source of coverage. Specifically, if firms offer ESHI, their workers lose the chance to buy subsidized

health insurance from HIX. This will be a more important problem for older and low skilled workers, who obtain health insurance at lower costs from the HIX relative to high skilled workers. Therefore, more older and low skilled workers show their preference for staying at low productivity firms not offering ESHI by turning down offers from high productivity firms offering ESHI, unless these high productivity firms offer wages which are high enough. These effects are confirmed in Figures 2.1 and 2.2. On the other hand, high skilled workers, who gain less from purchasing health insurance from HIX prefer to work at jobs offering ESHI. Such heterogeneous preferences can arise from not only skill heterogeneities, but also other characteristics such as age. These preference heterogeneities then lead to an inefficient sorting of workers across different firms.

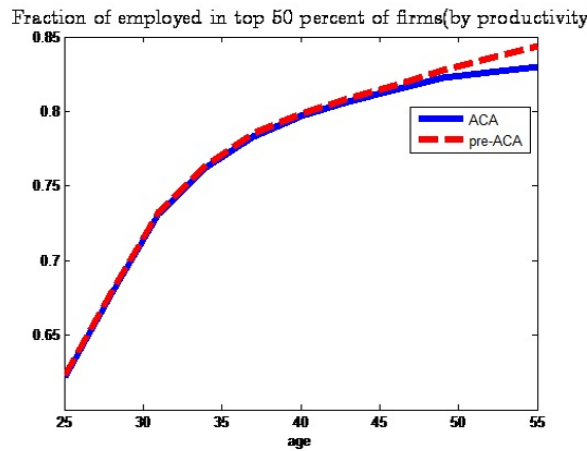


Figure 2.1: The fraction of employed workers in top 50% of firms (by productivity)

This result is very different from the standard perception that HIX can lead to more efficient sorting of workers across firms. Indeed, there has been a large literature investigating how the pre-ACA, employer based health insurance system leads to misallocation of workers by hampering reallocation of workers from low to high productivity firms, which are labeled as job-lock and job-push problems.³⁶ However, the existing evidence about the quantitative importance of these channels is mixed. Indeed, Dey and Flinn (2005), who estimate a search, matching, and bargaining model of wage and health in-

³⁶See Currie and Madrian (1999) and Gruber (2000) for an excellent survey of this literature.

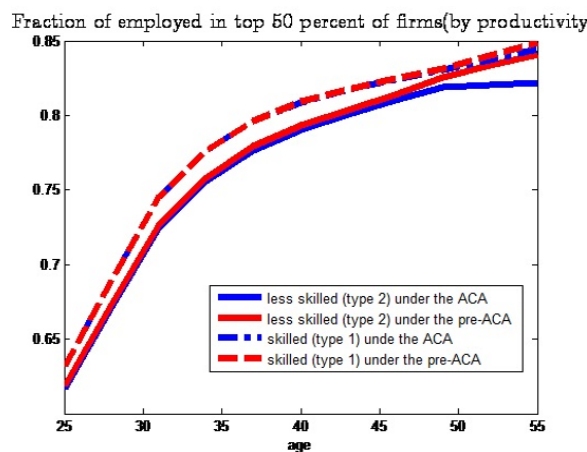


Figure 2.2: The fraction of employed individuals in top 50% of firms (by productivity): by individual type

insurance provisions, find that the existing employer based health insurance system leads to a negligible amount of inefficient job mobility decisions (i.e., decisions to turn down job offers from more productive jobs.) My study, rather, points out that the particular design of HIX can lead to less efficient allocation of workers across firms.

In terms of the quantitative significance of this result, it is comparable to other studies investigating the impact of institution on labor productivity. For example, Gourio and Roys (2012) find that size dependent regulations on firms in France lead to a decline in labor productivity of 0.27%. By using a different model, Garicano, LeLarge, and Van Reenen (2013) evaluate the impact of the same institution in France and find that it reduces labor productivity by 0.02% if wages are endogenously adjusted while it reduces labor productivity by 4% if wages are not adjusted.

Finally, column (6) shows the simulation result under the ACA without age based rating regulation. Here, I assume that each age group consists of different risk pools and that premia for each group is determined as a competitive price. Interestingly, the uninsured rate increases slightly, which reflects the fact that given the individual mandate and premium subsidies, perfect age based rating causes more adverse selection in each group. More interestingly, the premium subsidies to HIX substantially increase relative to the original ACA.

The partial equilibrium analysis: the role of equilibrium labor market interactions

To understand the role of equilibrium labor market interactions in Table 2.17, I conduct counterfactual experiments where the distribution of offers of compensation packages is assumed to be exogenous and the same as in the pre-ACA economy. As I find in Table 2.17, the rate of firm offering ESHI varies substantially depending on the design of HIX. The key question is how important is this channel to understanding the impact on the uninsured rate or other labor market outcomes. Note that this analysis is not directly comparable to Ericson and Starc (2012), Handel et al. (2013) and Hackmann, Kolstad, and Kowalski (2013). In these models, individuals can choose either to stay uninsured or to purchase health insurance from HIX. Therefore, their models do not account for the possibility that individuals change labor supply and job mobility decisions to take advantage of the benefit of purchasing health insurance from HIX. In this analysis, I still allow the individual-level responses in the labor market.

The result is reported in Table 2.20. In Column (2), I show the results under the case where all the components of the ACA are implemented. Compared with Column (2) in Table 2.17, where the offer distribution is endogenously determined, the labor productivity decline is much smaller. A similar result is obtained in Column (4) and (5), where the amount of premium subsidies is changed from the ACA.

To understand this result, it is important to recognize how the offer distribution is adjusted. As I explain in the last section, low skilled workers prefer to stay with low productivity firms if they have premium subsidies from HIX. This makes low productivity firms that do not offer ESHI more attractive relative to more productive firms offering ESHI, as these low productivity firms can hire and retain workers for longer. As a result, low productivity firms are less likely to offer ESHI, which increases the likelihood that these workers work at low productivity firms. Consequently, this channel becomes an additional mechanism that lower aggregate labor productivity.

	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Effects on the Firm Side						
Frac. of firms offering ESHI	0.5957	0.5957	0.5957	0.5957	0.5957	0.5957
...if firm size is less than 50	0.5486	0.5486	0.5486	0.5486	0.5486	0.5486
...if firm size is 50 or more	0.9612	0.9612	0.9612	0.9612	0.9612	0.9612
Average labor productivity	2.4883	2.4777	2.4771	2.4820	2.4742	2.4738
Output per capita	2.2483	2.2461	2.2474	2.2477	2.2452	2.2448
Panel B: Effects on the Worker Side						
Uninsured rate	0.2359	0.0665	0.1114	0.0972	0.0200	0.0977
The frac. of ind. with ESHI	0.7641	0.7794	0.7748	0.7834	0.5886	0.7776
Non-employment rate	0.0958	0.0924	0.0925	0.0942	0.0908	0.0921
Average wage	1.7859	1.7817	1.7821	1.7854	1.7798	1.7794
Medical expenditure	0.0724	0.0850	0.0838	0.0836	0.0864	0.0839
The frac. of the healthy workers	0.9501	0.9591	0.9581	0.9579	0.9599	0.9583
Panel C: Effects on Government Revenue:						
Revenue from Income Tax	0.6853	0.6854	0.6856	0.6857	0.6851	0.6848
Subsides to HIX & Medicaid	-	0.0138	0.0132	0.0117	0.0166	0.0162
Revenue from penalties	-	0.0013	1.63e-05	0.0019	1.57e-04	0.0019
Total Revenue	0.6853	0.6729	0.6724	0.6757	0.6687	0.6686

Table 2.20: Counterfactual Policy Experiments: Partial Equilibrium.

Notes: (a) Results between Column (1)-(6) are obtained under the case where the offer distribution of compensation package is exogenous and fixed as the distribution under the pre-ACA economy. (b) Column (1) reports the statistics generated under the pre-ACA economy. Note that all the statistics are calculated by including individuals less than age 25. The main pattern is unchanged even if I exclude them in this table. (c) Column (2) reports the statistics generated under the ACA. (d) Column (3) reports the statistics generated under the ACA without individual mandate. (e) Column (4) reports the statistics generated under the ACA without premium subsidies in the HIX. (f) Column (5) reports the statistics generated under the ACA with additional premium subsidies which are set that individuals who are eligible for premium subsidies in HIX obtain the full premium subsidies, subsidies which are equal to premia. (g) Column (6) reports the statistics generated under the ACA without maximum premium ratio regulation between the oldest and youngest. out premium subsidies in the HIX.

Another interesting pattern involves the uninsured rate. In Column (2) and (3), the reported uninsured rate is much lower relative to the one in Table 2.17. The main difference is the rate of firms offering ESHI. In Table 2.17, firms' ESHI offering rate is much higher, particularly for large firms. This causes workers to have less incentive to obtain health insurance and utilize health care when they are young, as they expect to be covered in the future.

Column (6) is the case where the premium is perfectly rated on the basis of age. Interestingly, in this case we find much more decline in output and labor productivity relative to the case where the offer distribution is endogenously adjusted. Here, young workers, who are more likely to experience job-to-job quits, face much lower premia relative to older workers in HIX. These workers have more incentive to stay at firms not offering ESHI. In the case where offer distribution of compensation packages is endogenously adjusted, this effect is partially muted because firms increase wage offer to take into account premium differences in HIX.

Overall, these results indicate that the simulated outcomes substantially differ between partial and general equilibrium settings. This indicates that an explicit modeling of labor market interactions is crucial to understand the welfare impact of HIX.

2.7 Normative Analysis: Optimal Design of Health Insurance Exchanges

In this section, I study whether the government can increase the welfare by altering the major design components of HIX. The previous section has shown mechanisms whereby each component of HIX leads to changes in the uninsured rate and labor market outcomes. I find that the current design of HIX system decreases the aggregate labor productivity by allocating more workers to relatively low productivity firms. This is mainly due to the premium subsidies, which contributes to lower uninsured rate in HIX. Moreover, I

also find that an adverse selection problem is much more severe for young individuals under the current structure of HIX. Given these findings, the natural question is whether there is a room for improving the welfare by altering the HIX system. To answer the question, I study the optimal design problem of HIX system. First, I specify policy tools for designing HIX. Next, I define the government problem of designing HIX. Then, I show the main results obtained from the optimal design problem.

2.7.1 The Government Problem

Policy instruments for HIX

In my welfare analysis, I consider the following three policy instruments for HIX: (1) age-based pricing regulation ω_{AGE} , (2) premium subsidies S , and (3) tax penalties on the uninsured IM ³⁷. I choose these policy instruments because these policies are regarded as the major instruments affecting the uninsured rate in HIX. Moreover, in public economics, these policy instruments are also regarded as the major instruments to correct welfare loss from adverse selection in insurance markets.³⁸ The key novelty of my analysis is to take into account equilibrium labor market interactions. To define the government problem, I specify their policy constraints in the optimal design problem. I proceed with analysis in the following three cases:

Case 1. S and IM follow the same functional form as the ACA. For example, tax penalties are specified as $(\omega_0^{im}, \omega_1^{im})$ such that $IM(y) = \max\{\omega_0^{im} \times y, \omega_1^{im}\}$. In the ACA, $\omega_0^{im} = 0.025$ and $\omega_1^{im} = \$695$. I parameterize S in the same way:

$$S(y, R^{HIX}(t)) = \begin{cases} \max\left\{R^{HIX}(t) - \left[\omega_0^s + \omega_1^s \frac{3y}{\omega_2^s}\right] y, 0\right\} & \text{if } y < \frac{\omega_2^s}{3} \\ 0, & \text{otherwise,} \end{cases}$$

In the ACA, $\omega_0^s = 0.035$, $\omega_1^s = 0.060$, and ω_2^s is the income at 400% federal poverty

³⁷Note that the specification under the ACA is explained in Appendix B.3.

³⁸See Einav, Finkelstein, and Levin (2010b) and Einav and Finkelstein (2011).

level $FPL400\%$.

Case 2. S and IM take a more flexible functional form. Specifically, they are specified so that

$$S(y, R^{HIX}(t)) = \frac{\exp(\omega_a^s + \omega_b^s y_t + \omega_c^s y_t^2)}{1 + \exp(\omega_a^s + \omega_b^s y_t + \omega_c^s y_t^2)} R^{HIX}(t)$$

$$IM(y, R^{HIX}(t)) = \frac{\exp(\omega_a^{im} + \omega_b^{im} y_t + \omega_c^{im} y_t^2)}{1 + \exp(\omega_a^{im} + \omega_b^{im} y_t + \omega_c^{im} y_t^2)} R^{HIX}(t)$$

Case 3. S and IM are explicitly age-dependent:

$$S(y, t, R^{HIX}(t)) = \frac{\exp\left(\begin{array}{l} \omega_a^s + \omega_b^s y_t + \omega_c^s y_t^2 \\ + \omega_d^s t + \omega_e^s t^2 \end{array}\right)}{1 + \exp\left(\begin{array}{l} \omega_a^s + \omega_b^s y_t + \omega_c^s y_t^2 \\ + \omega_d^s t + \omega_e^s t^2 \end{array}\right)} R^{HIX}(t)$$

$$IM(y, t, R^{HIX}(t)) = \frac{\exp\left(\begin{array}{l} \omega_a^{im} + \omega_b^{im} y_t + \omega_c^{im} y_t^2 \\ + \omega_d^{im} t + \omega_e^{im} t^2 \end{array}\right)}{1 + \exp\left(\begin{array}{l} \omega_a^{im} + \omega_b^{im} y_t + \omega_c^{im} y_t^2 \\ + \omega_d^{im} t + \omega_e^{im} t^2 \end{array}\right)} R^{HIX}(t)$$

I study the possibility that the government can condition premium subsidies and tax penalties on the uninsured by individual age for the following reasons. First, in HIX, insurance premia are, to some extent, pooled across individuals with different ages due to the age based pricing regulation. Second, I find in Section 2.6 that there is a sharp difference in terms of the impact of the ACA across different age groups. Third, it is known from recent optimal taxation and social insurance literature that age dependent policies are welfare improving (see, for example, Akerlof (1978); Michelacci and Ruffo (2011); Weinzierl (2011); Farhi and Werning (2013)). Motivated by these reasons, I examine the welfare impact of age dependent premium subsidies and tax penalties, in addition to age

based rating regulation.

Optimal design of HIX

The government chooses a combination of three policy instruments for HIX, $T_{HIX} = \{\omega_{AGE}, S, IM\}$, to maximize the social welfare subject to the revenue constraint. There are several ways to specify the social welfare function. By following the standard approach used in the social insurance and optimal taxation literature, I assume that the government is utilitarian. Moreover, I define the social welfare function as the ex-ante lifetime utility of newborn individuals:

$$SW(T_{HIX}) = \sum_{\tilde{X}_0} V_0(\tilde{X}_0) g_0(\tilde{X}_0, 0, 0)$$

where $V_0(\tilde{X})$ is lifetime utility of newborn individuals. The revenue constraint is given by

$$RV_{tax}(T_{HIX}) + RV_p(T_{HIX}) - EXP_{sub}(T_{HIX}) \geq R.$$

The first term in the left hand side is the revenue from the income tax in the equilibrium under the policy parameters T_{HIX} , which is given as

$$RV_{tax}(T_{HIX}) = \sum_t \sum_{\tilde{X}} \sum_{INS} \int T_I(\tilde{X}, \theta, INS) g_t(\tilde{X}, \theta, INS) d\theta,$$

and $RV_p(T_{HIX})$ is the revenue from tax penalties imposed on the uninsured and on large firms not offering ESHI, given as

$$\begin{aligned}
RV_p(T_{HIX}) = & \sum_t \sum_{\tilde{X}} \int IM(y_t, R^{HIX}(t), t) g_t(\tilde{X}, \theta, 0) d\theta \\
& + \sum_t \sum_{\tilde{X}} IM(b, R^{HIX}(t), t) u_t(\tilde{X}, 0) \\
& + \int_p EM \left(\sum_t \sum_{\tilde{X}} l_t(\tilde{X}, \theta_0^{ed}, 0) + l_t(\tilde{X}, \theta_0^{ed}, 2) \right) \\
& \times (1 - \Delta(p)) d\Gamma(p)
\end{aligned}$$

where the first line is the tax penalty on the uninsured who are employed, the second line is the penalty on the uninsured among non-employed, and the third term is the penalty on the large firms not offering ESHI. $EXP_{sub}(T_{HIX})$ is government subsidies for health insurance, which consist of the expenditure on the premium subsidies to HIX and Medicaid:

$$\begin{aligned}
EXP_{sub}(T_{HIX}) = & \sum_t \sum_{\tilde{X}} \int S_y(y_t, R^{HIX}(t), t, \tilde{X}) g_t(\tilde{X}, \theta, 2) d\theta \\
& - \sum_t \sum_{\tilde{X}} S_u(R^{HIX}(t), t, \tilde{X}) u_t(\tilde{X}, 1)
\end{aligned}$$

Finally, I specify that R is the government revenue obtained at the ACA, $T_{HIX} = T_{HIX}^{ACA}$. Therefore, the solution of this government problem tells us the maximum welfare gain achieved given the policy constraints relative to the ACA. In order to solve the planner's problem numerically, I use KNITRO, which is a solver for nonlinear optimization allowing nonlinear inequality constraints.³⁹ Moreover, I fix revenue constraint R as the revenue obtained if T_{HIX} takes the parameter values adopted by the ACA.

³⁹KNITRO is a derivative based optimization toolbox, and thus requires smoothness of the objective function. To guarantee the smoothness, I add preference shock for insurance purchasing decisions for HIX. I specify that it follows a Type-I extreme value distribution, where the scale parameter is set very small: 0.005.

2.7.2 Main Results

The general equilibrium analysis (with endogenous offer distribution of compensation)

First, I investigate the optimal structure of HIX in a general equilibrium setting, i.e., the setting where the offer distribution of compensation package is endogenously determined by firm's optimization problem. I begin by reporting the optimal policy parameters for Case 1, which is the case where the government's policy set is restricted within the functional form implemented under the ACA.

Policy Parameter	ACA	optimal
Panel A: policy for age based rating		
the MPR between the youngest and the oldest (ω_{AGE})	3.00	4.044
Panel B: premium subsidies to HIX		
constant term (ω_0^s)	0.035	0.005
the coeff. for income (ω_1^s)	0.060	0.052
the maximum 4-month income eligible for subsidies	400% FPL	242% FPL
Panel C: individual mandates (tax penalty to the uninsured)		
the minimum amount of annual tax penalty (ω_0^{im})	\$695	\$40.7
the percentage of tax penalty as a function of income (ω_1^{im})	2.5%	2.99%

Table 2.21: Optimal Policy Parameters under Case 1

Table 2.21 shows the optimal structure of HIX under Case 1 and its comparison from the ACA. The major differences are as follows: (1) the maximum premium ratio between the oldest and the youngest ω_{AGE} is larger; (2) the premium subsidies becomes more progressive: available up to 242% of FPL, more generous subsidies to low income; (3) individual mandates are set $\max\{0.03 \times y, \$29.5\}$, which indicates a larger tax penalty for high income individuals and a lower penalty to low income individuals. However, the welfare gain is modest. I measure the welfare gain by the annual lump-sum transfer to individuals to have the same utility under the ACA environment as the one in the optimal. For Case 1, it is merely \$45.

Next, I examine the optimal design under Case 2, which allows more nonlinearity in terms of the choice of premium subsidies and tax penalties. The result is shown in Table

2.22. I find that ω_{AGE} is even larger and the premium subsidies are still progressive. To calculate the welfare gain, I use the same measure as used in Case 1. The welfare gain is equivalent to an annual lump-sum transfer to each individual of \$111. Overall, I find that progressive subsidies and more transfers from the old to the young are welfare improving.

Optimal Policy Parameter	
Panel A: policy for age based rating	
The maximum premium ratio between the youngest and the oldest (ω_{AGE})	4.70
Panel B: premium subsidies to HIX	
Constant term (ω_a^s)	0.05
The coeff. for income (ω_b^s)	-0.50
The coeff. for income squared (ω_c^s)	0.002
Panel C: individual mandates (tax penalty to the uninsured)	
Constant term (ω_a^{im})	-3.63
The coeff. for income (ω_b^{im})	0.16
The coeff. for income squared (ω_c^{im})	0.01

Table 2.22: Optimal Policy Parameters under Case 2

Now, I consider Case 3, which allows age dependent premium subsidies and individual mandates. The optimal policy parameter is reported in Table 2.23. The general feature of the optimal policy relative to the ACA scheme is: (a) it allows a larger maximum premium ratio between the youngest and the oldest; (b) the ratio of premium subsidies to the premium is substantially decreasing in income and age; (c) the ratio of tax penalties over premium is very slightly increasing in income and age. Specifically, an important pattern is that premium subsidies are decreasing in age, while they are set independently in the ACA. The shape of optimal subsidies is summarized in Figures 2.3 and 2.4. As is clear from Figure 2.3, individuals with income at 300% of the Federal Poverty Level (FPL 300%) have very small subsidies over age, while those at FPL 133% receive large premium subsidies when they are young, decreasing gradually over time. Moreover, from Figure 2.4, one can see that premium subsidies become almost zero around a 4-month income of \$15,000, which is close to FPL 400% regardless of age. This pattern is similar to the ACA.

Optimal Policy Parameter	
Panel A: policy for age based rating	
The MPR between the youngest and the oldest (ω_{AGE})	3.74
Panel B: premium subsidies to HIX	
Constant term (ω_a^s)	3.95
The coeff. for income (ω_b^s)	-0.46
The coeff. for income squared (ω_c^s)	0.06
The coeff. for age ($t = \dots, 132$) (ω_d^s)	-0.02
The coeff. for age squared (ω_e^s)	-0.02
Panel C: individual mandates (tax penalty to the uninsured)	
Constant term (ω_a^{im})	-3.60
The coeff. for income (ω_b^{im})	0.005
The coeff. for income squared (ω_c^{im})	0.000
The coeff. for age ($t = \dots, 132$) (ω_d^{im})	0.003
The coeff. for age squared (ω_e^{im})	0.000

Table 2.23: Optimal Policy Parameters under Case 3: Allowance of Age-Dependent Subsidies and Individual Mandate

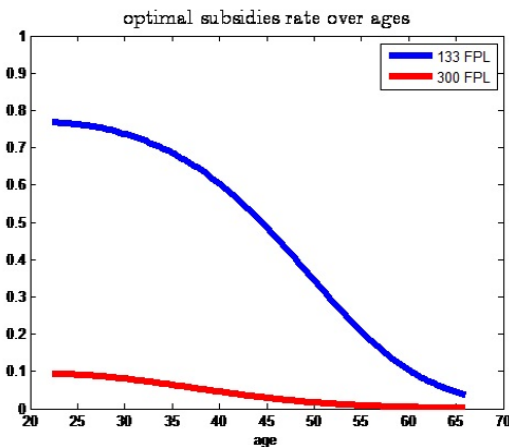


Figure 2.3: Optimal subsidies rate over ages

Column (2) in Table 2.24 reports the outcome under optimal HIX. First, it shows the substantial increase in output and labor productivity: output per worker increases by 0.1% and aggregate labor productivity increases by 0.6%. The increase in output is achieved despite the fact that the nonemployment rate increases. The uninsured rate is 7.53% which is somewhat lower relative to the ACA, while the magnitude of the changes is quite small compared with the pre-ACA economy where the uninsured rate is 23.6%.

In order to measure the welfare gain relative to the ACA, I assume that the government provides a lump-sum transfer to all individuals in the ACA economy. I find that the ACA

	(1)	(2)	(3)	(4)	(5)
Panel A: Effects on the Firm Side					
Frac. of firms offering ESHI	0.6268	0.6073	0.5672	0.6201	0.6079
Average labor productivity	2.4777	2.4875	2.4701	2.4814	2.4870
Output per capita	2.2461	2.2472	2.2464	2.2492	2.2471
Panel B: Effects on the Worker Side					
Uninsured rate	0.0788	0.0753	0.1159	0.0357	0.0907
Non-employment rate	0.0918	0.0941	0.0910	0.0925	0.0935
Average wage	1.7717	1.7840	1.7743	1.7799	1.7836

Table 2.24: Counterfactual Policy Experiments: Optimal Design of HIX under case 3 and its Variations.

Notes: (a) Column (1) reports the main aggregate outcomes under the ACA. (b) Column (2) reports the main aggregate outcomes under the optimal structure of the HIX. (c) Column (3) reports the main aggregate outcomes under the optimal structure of the HIX but the premium subsidies are replaced by the premium subsidies which are implemented under the ACA. (d) Column (4) reports the main aggregate outcomes under the optimal structure of the HIX but individual mandates (tax penalties on the uninsured) are replaced by the individual mandates which are implemented under the ACA. (e) Column (5) reports the main aggregate outcomes under the optimal structure of the HIX but an age based rating regulation (maximum premium ratio between the oldest and youngest) is replaced by the regulation which is implemented under the ACA.

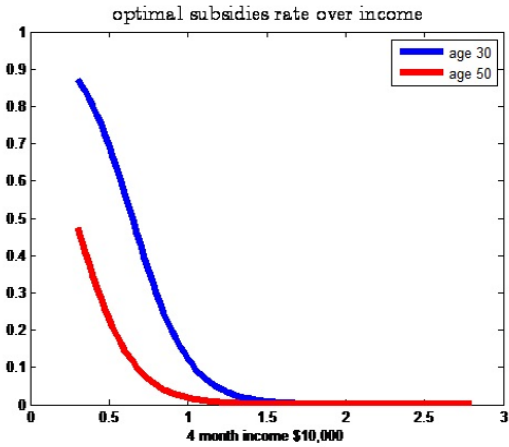


Figure 2.4: Optimal subsidies rate over income

will achieve the same level of utility as the optimal HIX if the government provide the transfer which is \$195 annually. The transfer corresponds to 7.6% of medical expenditure. This eventually contributes to an increase in government expenditure of 1.1%, which is substantial.

Columns (3) to (5) report what happens when each component of optimal HIX is replaced by the component of HIX implemented under the ACA. Column (3) reports the major outcomes if the premium subsidies scheme is replaced by the scheme implemented under the ACA. I find that the uninsured rate is much higher while labor productivity is lower. The rise in the uninsured rate is mainly explained by the fact that more young and healthy individuals are uninsured. Column (3) reports the major outcomes if the individual mandate scheme is replaced by the scheme implemented under the ACA. I find that the uninsured rate is much smaller, as it induces more participation into HIX. While labor productivity is still higher than the ACA, it is lower relative to the optimal HIX reported in Column (2). Finally, Column (5) reports the major outcomes if the age based rating regulation is replaced by the regulation under the ACA. I find that the uninsured rate is higher while most labor market outcomes are similar to those reported in Column (2). These results demonstrate that all of the components contribute to the changes in outcomes, but redesigning premium subsidies has a crucial effect on labor productivity.

To understand the mechanism generating the welfare gain, it is important to recognize that under the optimal structure of HIX, relatively old workers gain less benefit than young workers from purchasing health insurance from HIX, because the optimal structure sets higher maximum premium ratio (MPR) and also allows premium subsidies to decrease with age. By giving more subsidies to the young, the optimal structure can ameliorate the adverse selection problems explained in Section (2.6.2). Because young workers are given more premium subsidies, they have more incentive to participate in HIX. Moreover, because the gain from participating in HIX is smaller for relatively old workers, this structure gives old workers more incentives to work at firms offering ESHI. Because high productivity firms are more likely to offer ESHI, this optimal structure can give more incentives to workers to move from low to high productivity firms over the life cycle. This intuition is confirmed by Figure 2.5, which shows that more older individuals are allocated on more productive firms relative to the outcome under the ACA.

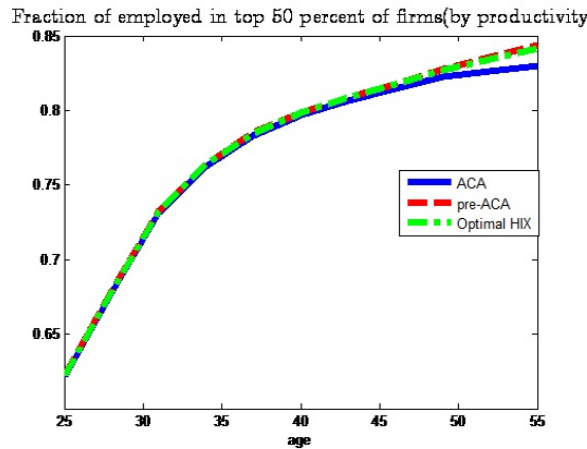


Figure 2.5: The fraction of employed workers in top 50% of firms (by productivity)

While these economic forces give the benefit of redistribution from old to young workers, its cost is to make old workers worse off, in particular for those that lose their job offering ESHI due to exogenous job destruction shocks and start working at firms not offering ESHI. Indeed, Table 2.25 shows that the measure of the uninsured among the age group 45-55 increases due to the decrease in premium subsidies. The optimal structure

of HIX is therefore determined to take into account such a cost as well as the benefit described above.

Age Group	<u>ACA</u> (1)	<u>Optimal HIX</u> (2)
Age Group	Uninsured who are employed	Uninsured who are employed
25-35	0.12	0.07
35-45	0.07	0.07
45-55	0.004	0.04

Table 2.25: Comparison of the uninsured rate by different age groups under the ACA and under the optimal structure of HIX in Case 3.

Note: Uninsured who are employed is the measure of individuals who are uninsured and employed.

Finally, an interesting feature of the optimal structure is while premium subsidies are substantially decreasing in age, tax penalties are not. This difference reflects the differential effect of subsidies and tax penalties, which arise in an economy where both HIX and ESHI are modeled. Both larger subsidies and higher tax penalties give individuals more incentive to obtain health insurance. However, while larger subsidies give individuals incentive to purchase health insurance from HIX rather than to obtain health insurance from employers, higher penalties do not directly give such incentive to individuals. Therefore, in order to give incentive the old individuals to obtain health insurance from employers, premium subsidies are set to decrease in age. This differential impact of the premium subsidies and tax penalties has not pointed out in the existing works that study HIX designs, as they do not consider ESHI. Therefore, this finding indicates the importance of modeling ESHI and the labor market to understand the optimal design of HIX.

The partial equilibrium analysis

Next, I investigate how the optimal structure differs if one ignores general equilibrium effects of labor markets. To this end, I assume that the offer distribution of compensation package is fixed and the same as the one in the pre-ACA economy. Then, I solve the worker's problem taking into account the equilibrium determination of health insurance premia in HIX.

The optimal policy parameters under Case 3 are summarized in Table 2.26. The key features of the optimal policies relative to the optimal HIX under the general equilibrium case are as follows. First, average premium subsidies are, in general, larger, particularly for low income individuals. This reflects the fact that the offer distribution is not adjusted, and therefore does not contribute a reduction of labor productivity. Second, the tax penalty on the uninsured is higher, particularly for high income individuals. Because tax penalties do not increase the ESHI offering rate, higher penalties are needed to improve the pool of HIX. Finally, the maximum premium ratio (MPR) between the oldest and the youngest is slightly higher than the MPR under the general equilibrium case.

Optimal Policy Parameter	
Panel A: policy for age based rating	
The maximum premium ratio between the youngest and the oldest (ω_{AGE})	3.91
Panel B: premium subsidies to HIX	
Constant term (ω_a^s)	4.63
The coeff. for income (ω_b^s)	-0.31
The coeff. for income squared (ω_c^s)	0.22
The coeff. for age ($t = \dots, 132$) (ω_d^s)	-0.01
The coeff. for age squared (ω_e^s)	-0.001
Panel C: individual mandates (tax penalty to the uninsured)	
Constant term (ω_a^{im})	-3.62
The coeff. for income (ω_b^{im})	0.16
The coeff. for income squared (ω_c^{im})	0.01
The coeff. for age ($t = \dots, 132$) (ω_d^{im})	0.00
The coeff. for age squared (ω_e^{im})	0.00

Table 2.26: Optimal Policy Parameters under Case 3 and Exogenous Offer Distribution

To understand how the implications differ if we use the optimal policies obtained under the partial equilibrium setting, I report the comparison of outcomes simulated under the general equilibrium setting between optimal HIX policies under general equilibrium (in Column (2)) and optimal HIX policies under partial equilibrium (in Column (3)) in Table 2.27. First, by providing more subsidies, aggregate labor productivity is lower under optimal HIX policies under partial equilibrium (in Column (3)) than the productivity under general equilibrium (in Column (2)). Moreover, due to a larger expenditure on premium subsidies, the fraction of firms offering ESHI decreases and the total government expen-

diture is slightly higher under the partial equilibrium. These results show the importance of taking into account equilibrium effects when designing the optimal HIX.

	Optimal HIX: GE	Optimal HIX: partial
	(1)	(2)
Frac. of firms offering ESHI	0.6073	0.5646
Average labor productivity	2.4875	2.4775
Uninsured rate	0.0753	0.0853
Non-employment rate	0.0941	0.0938

Table 2.27: Counterfactual Policy Experiments: Comparison: Outcomes in general equilibrium under (1) optimal policies obtained with endogenous offer distribution of compensation packages and (2) optimal policies obtained with exogenous offer distribution of compensation packages.

2.7.3 Discussion: the role of Medicaid expansion

So far, all the exercises are conducted under the assumption that Medicaid is expanded as projected. However, as of August 2013, 25 states do not plan to expand Medicaid. In this section, I investigate the interaction between Medicaid expansion and HIX. Medicaid expansion affects HIX as it affects the risk pools in HIX. Specifically, if Medicaid is not expanded, individuals who would have been eligible for Medicaid will be absorbed by HIX. Because these individuals are less healthy relative to the average population in the economy, their absorption into HIX may increase health insurance premium in HIX, worsening the adverse selection problem.⁴⁰ Therefore, the question is how much effect Medicaid expansion will have on risk pooling in HIX. At this stage, it is not clear how state governments that decide not expand Medicaid will choose premium subsidies for those who are eligible for Medicaid but ineligible for federal premium subsidies. However, for simplicity, I assume that individuals who were eligible for Medicaid obtain full premium subsidies from HIX. That is, those individuals will be guaranteed to be insured through HIX. I call this situation "no-Medicaid expansion" in the remainder of the study.

⁴⁰The link between Medicaid and the risk pools in individual insurance markets is investigated by Clemens (2013).

Therefore, the impact of no-Medicaid expansion is realized purely through its impact on risk pools in HIX.

First, I evaluate the impact of the ACA with no-Medicaid expansion. I find that the uninsured rate is 8.3%, which is higher than the rate under the ACA with Medicaid expansion, which was 7.8%. The main reason is that the risk pool in HIX is worse due to the entry of pre-Medicaid pools who are less healthy. This negative externality is large enough that more employed workers choose to be uninsured. Next, I examine the optimal structure of HIX, given the ACA budget constraint. To make the comparison transparent, I assume that the pre-Medicaid eligible population still obtain health insurance for free from HIX. I find that the optimal structure of HIX is qualitatively similar to the case with Medicaid expansion, including the age dependence of premium subsidies. More interestingly, I find that the welfare gain from the optimal design of HIX is rather small: it amounts to \$80, much smaller than the welfare gain with Medicaid expansion, which amounts to \$195. These results highlight the importance of interactions between Medicaid and HIX.

2.8 Conclusion

In this study, I evaluate the current HIX system and investigate its optimal design, accounting for adverse selection and equilibrium labor market interactions. I first develop and estimate a life cycle equilibrium labor market search model integrated with the pre-ACA health insurance market. Various forms of individual heterogeneity are incorporated to understand the welfare consequences of HIX. Through counterfactual experiments, I find that the ACA substantially reduces the uninsured rate. However, the ACA also decreases aggregate labor productivity by allocating more workers to less productive firms.

Next, I examine the optimal design of HIX by choosing the values of three major design components – individual mandates, premium subsidies and age-based rating regulations. I found that the optimal combination of these components makes it less beneficial for older

workers to purchase health insurance from HIX, relative to young workers. Implementing the optimal structure leads to a substantial welfare gain relative to HIX implemented under the ACA, while achieving higher labor productivity and a slightly lower uninsured rate. In this structure, the adverse selection problem in HIX among the young is reduced. Moreover, it gives older workers an incentive to work at firms offering ESHI. This increases the allocation of workers from low to high productivity firms, raising aggregate labor productivity.

Finally, I assess the role of equilibrium labor market interactions by assuming that the distribution of offers of compensation package is fixed and evaluating the impact of each component of HIX as well as the optimal design of HIX. I find that both the impact of HIX components and their optimal design are qualitatively and quantitatively very different from the equilibrium with endogenous compensation packages, indicating the importance of modeling equilibrium labor market interactions to evaluate HIX design.

There are a number of dimensions in which my stylized model could be extended to capture other important features. One of the most important extensions is to allow multiple insurance products offered in HIX to understand the importance of adverse selection across insurance plans within HIX. Such an extension allows us to study the optimal regulation of insurance contracts, such as the choice of minimum creditable coverage. However, there are a number of difficulties. First, it requires careful choices about the notion of equilibrium, as certain types of equilibria (e.g., Nash equilibrium) may not exist in such environment. Handel et al. (2013) make an important contribution in this regard. Second, an additional important challenge is to model the situation where firms offer a menu of ESHI contracts to workers, and to obtain data about the set of insurance plans that workers face and firms offer. At this stage, it is very difficult to obtain such data. Third, it also requires a more sophisticated modeling of individual health care utilization decisions. While each of these issues is very challenging and requires a substantial departure from this framework, these extensions are exciting opportunities for future work.

Chapter 3

Advertising Competition and Risk Selection in Health Insurance Markets: Evidence from Medicare Advantage

This chapter is co-authored with You Suk Kim.

3.1 Introduction

Medicare provides health insurance for the majority of elderly Americans. Although traditional fee-for-service Medicare is public insurance provided by the government, many Medicare beneficiaries opt out of traditional Medicare to receive coverage from Medicare Advantage (MA) plans offered by private insurance companies. A main factor that differentiates MA plans from traditional Medicare is the provision of additional services at the cost of a restricted provider network. In 2011, about 25% of Medicare beneficiaries enrolled in MA. An MA plan receives a capitation payment from the government for its enrollee and then bears the health care costs incurred by the enrollee. The capitation

payment accounts for most of the plans' revenues, even though MA plans often charge a premium.

A potential problem of MA is that private insurers have incentives to selectively enroll low-cost, healthy individuals (or “risk-select”) due to an imperfect risk adjustment of capitation payments. Table 3.1 illustrates the presence of strong incentives for risk selection by private insurers, and the incentives are observed not only in Los Angeles but also in other regions throughout the nation. Given that regulations prohibit an MA plan from charging different premiums to individuals with different health risks, the opportunity to increase profits by enrolling healthier individuals provides insurers incentives to risk-select. Moreover, there is regional variation in the amounts of over-payment for the healthy, which creates incentives for MA plans to risk-select more intensively in regions with these higher over-payments. Indeed, previous research on MA finds that MA enrollees are healthier than traditional Medicare enrollees.¹ Although preference heterogeneity between healthy and unhealthy individuals for MA plans can partly account for the observed selection patterns, incentives for risk selection, as illustrated by Table 3.1, are strong.

	Self-reported Health Status				
	Excellent	Very Good	Good	Fair	Poor
Monthly Capitation Payment (\$)	601.0	619.5	646.6	708.0	796.3
Monthly Health Expenditure (\$)	266.0	347.8	575.4	923.7	2029.4
Monthly Over-payment (\$)	335.0	271.3	71.2	-215.7	-1233.1

Table 3.1: Capitation Payment and Health Expenditure by Health Status in Los Angeles County

Note: Over-payment = Capitation payment - Health Expenditure.

Source: Medicare Current Beneficiary Survey 2000–2003

The main goal of this study is to empirically study incentives for private insurers to use advertising as a means of risk selection and the impacts of advertising on the MA market. Previous work on risk selection views advertising as one of the central tools of

¹For examples, see Langwell and Hadley (1989); Mello et al. (2003); Batata (2004).

risk selection (Van de Ven and Ellis 2000; Brown et al. 2012). MA plans might target advertising to healthy beneficiaries, for example, through its content (Neuman et al. 1998; Mehrotra et al. 2006). Moreover, advertising can be targeted to regions having greater over-payments for the healthy. When private insurers can risk-select with advertising, the effects will not be limited to MA enrollees and insurers, but the government's budget will also be affected through over-payments for the healthy. Despite the potential importance of advertising in MA, however, there is no existing quantitative analysis on the effects of advertising on risk selection or its effects on health insurance markets in general.

In order to understand the role of advertising, we develop and estimate an equilibrium model of the MA market, which incorporates strategic advertising by insurers. On the demand side of the model, consumers make a discrete choice to enroll with one of the available MA insurers or to select traditional Medicare. We assume advertising affects a beneficiary's indirect utilities, thus capturing persuasive, prestige and signaling effects of advertising. We capture the effect of advertising on risk selection with its heterogeneous effects on demand, depending on an individual's health status. Customer preferences for a plan also depend on its other characteristics such as premiums and coverage benefits. On the supply side, insurers simultaneously choose premiums and levels of advertising to maximize profits. A firm's revenue from an enrollee equals the sum of the premium and capitation payment for the enrollee, while its cost of insuring an enrollee depends on plan characteristics and the enrollee's health risk. Thus the optimal pricing and advertising of a plan takes into account the effects of these choices on the plan's composition of health risks.

Our empirical analysis relies on data from a variety of sources. First, we use data on advertising expenditures by health insurers in the 100 largest local advertising markets for the period 2000–2003 from AdSpender Database of Kantar Media, a leading market research firm.² Second, we use data on individual MA insurer choices, together with

²A local advertising market consists of a major city and its surrounding counties, and the 100 largest markets cover more than 80% of the total U.S. population.

information on the respondents' demographic and health statuses. Third, we use data sets published by the Center for Medicare and Medicaid Services, which have information on the number of enrollees and plan benefit characteristics for each plan in each county in each year and capitation payments in each county in each year. The data show the potential importance of advertising in relation to risk selection: There is a large variation in advertising expenditures across local markets, and advertising efforts by insurance companies are concentrated in markets with higher margins from enrolling healthier individuals. Within a market, moreover, healthier individuals are more likely to enroll with MA insurers that use more advertising.

We estimate the demand and supply side of the model in two steps, using generalized method of moments. For estimation of the demand model in the first step, we allow for time-invariant plan fixed effects and use instrumental variables to account for the endogeneity of premiums and advertising stemming from (time-varying) unobserved plan heterogeneity. In the second step, the supply model is estimated using the estimated demand model and optimality conditions for observed pricing and advertising choices by insurers. In the supply model, we account for the possibility that insurers choose zero advertising, which is frequently observed in the data. Parameter estimates show that advertising has a positive effect on overall demand, but a much larger effect on healthier consumers.

With the estimated model, we investigate the effects of advertising on the MA market and evaluate the effects of a policy that adjusts capitation payments based on an individual's health risks. In order to investigate the effect of advertising on the MA market, we simulate the model in an environment in which advertising is banned. The ban decreases overall MA enrollment by 4% and enrollment for MA plans with above-average advertising spending by 9%. Despite the lower demand without advertising, we find that insurers lower their premiums by very little, which results from the fact that MA enrollees become less healthy on average without advertising, raising the MA insurers' cost. The absence

of advertising decreases the difference in the expected health expenditures of enrollees in traditional Medicare and MA by 15%, which reduces the average excess capitation payment per MA enrollee by 4%. This finding implies that risk selection with advertising accounts for 15% of the selection of healthier individuals into MA.

We also investigate the effects of a policy that reduces the incentive for risk selection. We consider a perfectly risk-adjusted capitation payment so that the difference between an enrollee's capitation payment and expected health expenditure is the same for any individual. We find that the risk adjustment policy has large effects on the equilibrium. Monthly premiums increase from \$30.1 to \$51.1; advertising expenditures decrease by 30%; and overall MA enrollment rates decrease by 9%. Because the risk adjustment policy reduces capitation payments for healthy enrollees, insurers compensate for the decrease in revenues by increasing premiums. Moreover, insurers reduce advertising because insuring the healthy is now less profitable. These findings highlight a strong link between risk selection and advertising.

This study contributes to a large body of literature empirically investigating adverse selection and risk selection in insurance markets. Previous research finds that an individual's heterogeneous characteristics, such as risk, risk preference, income, and cognitive ability, are important determinants of selection patterns in insurance markets.³ More recently, researchers empirically investigated the possibility that the insurer affects consumer selection in different health insurance market settings. Bauhoff (2012) studies risk selection in a German health insurance market by looking at how insurers respond differently to insurance applications from regions having different profitabilities. Brown et al. (2012) provide descriptive evidence that insurers engage in risk selection in MA, using the introduction of sophisticated risk adjustment of capitation payments to MA plans. Kuziemko et al. (2013) study risk selection among private Medicaid managed-care insurers in Texas and provide evidence that the insurers risk-select more profitable individuals.

³For examples, see Chiappori and Salanie (2000) for automobile insurance, Finkelstein and McGarry (2006) for long-term care insurance, and Fang et al. (2008) for Medicare supplement insurance.

Although the occurrences of risk selection are well documented in the related works, there is still little research on its channels. This study adds to this literature by investigating the role of advertising on risk selection.

Our focus on an insurance company's behavior in insurance markets is related to a new and growing body of literature studying demand and competition in insurance markets. For example, Lustig (2011) studies adverse selection and imperfect competition in MA with an equilibrium model that endogenizes a firm's choice of premium and plan generosity by creating an index of generosity. Starc (2012) investigates the impact of adverse selection on an insurer's pricing and consumer welfare in an imperfectly competitive market (Medicare supplement insurance).⁴ This study adds to this literature by examining how advertising, which is a less explored and less regulated channel relative to competition on pricing and coverage, affects risk selection and competition.

Lastly, this study is also related to the literature on advertising. Many empirical papers in the literature study the channels through which advertising influences consumer demand—i.e., whether advertising gives information about a product or affects utility from the product.⁵ More recently, researchers have studied the effects of advertising in an equilibrium framework for different markets. Goeree (2008) studies advertising in the personal computer market in the U.S., and Gordon and Hartmann (2013) study advertising in a presidential election in the U.S. A paper that is closely related to ours is Hastings et al. (2013), who also study advertising in a privatized government program (the privatized social security market in Mexico). An important difference between this study and the related works on advertising is that advertising in MA affects not only consumers and insurers but also the government. If MA insurers can risk-select with advertising, the enrollment decisions made by healthy individuals will directly affect government expenditures because the government over-pays for the insurance of these individuals.

⁴For other works in this literature, see Bajari et al. (2011); Bundorf et al. (2012b); Carlin and Town (2007); Cohen and Einav (2007); Dafny and Dranove (2008); Einav et al. (2010a,c); Nosal (2012); Town and Liu (2003).

⁵For examples, see Akerberg (2001, 2003); Ching and Ishihara (2012); Clark et al. (2009).

The rest of this chapter is organized as follows. Section 2 describes Medicare Advantage in greater detail. Section 3 describes the data and presents results from the preliminary analysis. Section 4 outlines the model while Section 5 discusses estimation and identification of the model. Section 6 provides estimates of the model, and Section 7 describes results from counterfactual analyses. Section 8 concludes.

3.2 Background on Medicare Advantage

Medicare is a federal health insurance program for the elderly (people aged 65 and older) and for younger people with disabilities in the United States. Before the introduction of Medicare Part D in 2006, which provides prescription drug coverage, Medicare had three Parts: A, B, and C. Part A is free and provides coverage for inpatient care. Part B provides insurance for outpatient care. Part C is the Medicare Advantage program, previously known as Medicare + Choice until it was renamed in 2003.⁶

The traditional fee-for-service Medicare is comprised of Parts A and B, which reimburse costs of medical care utilized by a beneficiary who is covered by Parts A and B. As an alternative to traditional Medicare, a Medicare beneficiary also has the option to receive coverage from an MA plan run by a qualified private insurer. Insurers wishing to enroll Medicare beneficiaries sign contracts with the Center for Medicare and Medicaid Services (CMS) describing what coverage they will provide, and at what costs. The companies that participate in the MA program are usually health maintenance organizations (HMOs) or preferred provider organizations (PPOs), many of which have a large presence in individual or group health insurance markets, such as Blue Cross Blue Shield, Kaiser Permanente, United Healthcare, etc. They contract with the Center for Medicare and Medicaid Services on a county-year basis and compete for beneficiaries in each county where they operate.

⁶Although we will focus on the period 2000–2003 for our analysis, we will refer to Medicare private plans as Medicare Advantage plans instead of Medicare + Choice plans.

The main attraction of MA plans for a consumer is that they usually offer more comprehensive coverage and provide benefits that are not available in traditional Medicare. For example, many MA plans offer hearing, vision, and dental benefits which are not covered by Parts A or B. Before the introduction of Part D, prescription drug coverage was available in MA plans, but not in traditional Medicare. Although a beneficiary in traditional Medicare is able to purchase Medicare supplement insurance (known as Medigap) for more comprehensive coverage than basic Medicare Parts A and B, the Medigap option is priced more expensively than a usual MA plan, many of which require no premium. Therefore, MA is a relatively cheaper option for beneficiaries who want more comprehensive coverage than traditional Medicare offers. In return for greater benefits, however, MA plans usually have restrictions on provider networks. Moreover, MA enrollees often need a referral to receive care from specialists. In contrast, an individual in traditional Medicare can see any provider that accepts Medicare payments.

Previous works on MA find that healthier individuals are systematically more likely to enroll in a MA plan.⁷ The selection pattern may result from preference heterogeneity between healthy and unhealthy individuals for MA plans. For example, unhealthy individuals may dislike certain aspects of MA plans such as restricted provider networks and referral requirements. However, it is also possible that insurers' risk-selection reinforces the direction of consumer selection. Indeed, incentives for MA plans to risk-select are strong. By regulation, MA insurers must charge the same premium for individuals with different health statuses in a county. More importantly, capitation payments from the government do not fully account for variation in health expenditures across individuals. Until the year 2000, the CMS paid capitation payments equal to 95% of the expected costs of treating a beneficiary within traditional Medicare, and adjustments to payments were made based only on an enrollee's age, gender, welfare status, institutional status, and location. However, these adjustments, based solely on demographic information, were

⁷For example, see Langwell and Hadley (1989); Mello et al. (2003); Batata (2004).

found to account for only about 1% of an enrollee’s expected health costs (Pope et al. 2004). During the period of 2000–2003, which is a focus of this study, the CMS made 10% of capitation payments depend on inpatient claims data using the PIP-DCG risk adjustment model, but the fraction of variations in expected health costs by the newer system remained around 1.5% (Brown et al. 2012).⁸

3.3 Data and Preliminary Analysis

3.3.1 Data

This study combines data from multiple sources. We use the Medicare Current Beneficiary Survey (MCBS) for the years 2000–2003 for individual-level information on MA enrollment and demographic characteristics, including health status. Our data on advertising by health insurers in local advertising markets for the years 2000–2003 were retrieved from the AdSpender Database of Kantar Media, a leading market research firm. Market share data for the years 2000–2003 are taken from the CMS State-County-Plan (SCP) files, and insurers’ plan characteristics are taken from the Medicare Compare databases for the years 2000–2003.⁹

The reason we study MA for the years 2000–2003 is because the MCBS does not provide information on an individual’s choice of MA insurer from 2006 onward. We also avoid using data right before 2006 because Medicare Part D was introduced in that year, changing many aspects of the MA market.

⁸From the year 2004, a more sophisticated risk adjustment model is implemented. However, Brown et al. (2012) find that MA insurers were still able to selectively enroll more profitable individuals because even the new model did not perfectly account for variation in health expenditures across individuals. The reason that we focus on the period 2000–2003 is discussed later when we introduce our data.

⁹We thank Kathleen Nosal for sharing Medicare Compare data with us.

Individual-level Data

The MCBS is a survey of a nationally representative sample of Medicare beneficiaries. This dataset provides information on a beneficiary's demographic information such as age, income, education, and location, as well as an extensive set of variables on an individual's health status: self-reported health status, difficulties in activities of daily living (ADL), difficulties in instrumental activities of daily living (IADL), and a history of diseases such as cancers, heart diseases, diabetes, etc. An important feature of this dataset is that it is linked to administrative data in Medicare, which provides information on an individual's MA insurer choice, the amount of the capitation payment paid for an MA enrollee in the sample, and the amount of Medicare claims costs for individuals in traditional Medicare.

For our analysis, we only use observations who are at least 65 years old. This means that we exclude the sample of individuals under 65 who are on Medicare solely due to disability. Although these individuals can purchase MA plans, we exclude them because the main factor that affected capitation payments for the years 2000–2003 was age and because we want to have samples of individuals who are more or less similar in terms of their capitation payments. Because beneficiaries younger than 65 years old represent a small fraction of MA enrollment (7%), we do not view this exclusion as a serious problem.

Health status An important variable from this dataset is an individual's health status. A health status can be measured in many different ways, and there are plenty of variables in the MCBS that are related to health status. Because it is very difficult to include all possible measures separately in the empirical analysis, we construct a one-dimensional continuous measure of health status. Our measure of an individual's health status is expected claims costs if an individual were to be insured by Medicare Parts A and B. To construct this measure of health status, we use information on an extensive set of observed health statuses and the realized amount of Medicare Parts A and B claims for each individual who remained in traditional Medicare. Because information on Medicare

claims is available only for individuals in traditional Medicare, we have to impute expected claims costs for MA enrollees using their observed health statuses. Thus, we first estimate equations that relate Medicare claims costs to an extensive list of health characteristics using beneficiaries enrolled in traditional Medicare. Then we calculate expected claims costs not only for traditional Medicare enrollees, but also for MA enrollees.¹⁰ A detailed discussion on constructing the health status variable is in the Appendix.

Capitation Payment For our analysis, we need to know how much an MA plan would receive when enrolling a Medicare beneficiary with certain characteristics. Unfortunately, the MCBS does not provide such information. Instead, it contains information on how much an MA plan received for a Medicare beneficiary enrolled in MA. In order to calculate a capitation payment amount for an enrollee, we exploit the fact that capitation payments were mostly based on the simple demographic factors for the years 2000–2003, as described in the previous section.¹¹ First, we regress an actual capitation payment for an MA enrollee in the MCBS on the enrollee’s demographic characteristics that are used in the calculation of actual capitation payments. With coefficient estimates from the regression, we calculate a capitation payment for any Medicare beneficiary. Because capitation payments depend only on exogenous demographic characteristics, selection bias is not a concern here even though the regression is run with data on MA enrollees only. The coefficient estimates in the regression are reported in Table 3.2. The results show that the variables included in the regression explain a large part of variation in capitation payments, with R-squared of 0.822. The estimates are used to calculate a capitation payment amount for all Medicare beneficiaries including those who chose traditional Medicare.

¹⁰An implicit assumption here is that traditional Medicare and MA enrollees do not differ in unobserved health status. Given the extensive list of variables on health status used in imputation, however, it is reasonable to assume that we can capture most of the meaningful differences in health status.

¹¹As explained in the previous section on MA, 10% of capitation payments depended on inpatient claims data for the years 2000–2003. For this version of this study, we ignore the dependence of capitation payments on the data, which only accounted for 0.5% of health expenditures Brown et al. (2012). Given the small role of inpatient data in the calculation of capitation payments, we do not view the omission as a serious problem.

Dep. Var: Capitation Payment Paid for an MA Enrollee		
Variables	Coefficient	Std. Err.
Female	219.3**	(88.20)
Male with Age 65–69	-165.1**	(81.18)
Female with Age 65–69	-286.2***	(34.91)
Male with Age 70–74	-67.81	(80.72)
Female with Age 65–69	-209.1***	(33.11)
Male with Age 75–79	36.74	(80.45)
Female with Age 75–79	-140.0***	(31.20)
Male with Age 80–84	97.85	(80.00)
Female with Age 80–84	-88.06***	(29.93)
Male with Age 85–89	135.2*	(80.32)
Female with Age 85–89	-24.38	(29.44)
Male with Age 90–94	117.7	(83.38)
Female with Age 90–94	-51.31*	(31.05)
Age	-17.34***	(1.671)
Living in a Nursing Home	-414.0	(287.0)
Medicaid Eligible	-97.88*	(56.98)
Avg. Capitation	-1.485***	(0.180)
Avg. Capitation × Female	-0.272***	(0.0341)
Avg. Capitation × Nursing Home	1.664***	(0.502)
Avg. Capitation × Medicaid	0.624***	(0.0885)
Avg. Capitation × Age	0.0336***	(0.00243)
Observations	8,020	
R-squared	0.822	

Table 3.2: Capitation Payments and Demographic Characteristics

Note: Avg. Capitation means the average capitation payment in county of residence for each individual, which is extracted from the Medicare State-County-Plan databases 2000–2003. The regression was run only with the sample of MA enrollees.

Source: Medicare Current Beneficiary Survey 2000–2003; Medicare State-County-Plan databases 2000–2003.

Advertising Data

AdSpender contains information on the annual advertising expenditures and quantities of health insurers in different media such as TV, newspaper, and radio in the 100 largest local advertising markets in the U.S. A local advertising market consists of a major city and its surrounding counties, and its size is comparable to that of a Metropolitan Statistical Area (MSA).¹² Advertising quantity is defined as the number of times an advertisement appeared in a medium in a given period, and this information is only available for TV and newspapers. AdSpender categorizes advertising across product types whenever specific product information can be detected in an advertisement, which allows us to isolate advertising expenditures for an insurer's MA plan in some instances. For example, some expenditures are reported in detail (e.g. Humana Gold plan, which is an MA plan offered by Humana Insurance Company), while others are reported more generally (e.g., Blue Cross Blue Shield health insurance in general). An advertisement falls into the latter category when it does not mention product names, or when it is for an insurer itself (not for its specific products).

In constructing a measure of advertising levels for MA plans, we excluded advertising expenditures specific to insurance products that are not MA plans. Whenever information on a product is available in the data, for example, we can tell whether the product was sold in individual or group markets for individuals not on Medicare. In the end, we use advertising expenditures for MA plans and general advertising expenditures. Because the latter is likely to be meant not only for the Medicare population but also for the non-Medicare population, we make adjustments for expenditures for general advertisements, while we do not make any changes to advertising expenditures for MA plans. To be more precise, we denote ad_{jmt}^{ma} and ad_{jmt}^g as a firm j 's MA-specific and general advertising expenditures in a local advertising market m in year t , respectively. Our final measure of

¹²In the advertising industry, this local market is usually referred to as a Designated Media Market, which is defined by Nielsen company.

advertising expenditures for firm j 's MA plans in market m in year t , ad_{jmt} , is that:

$$ad_{jmt} = ad_{jmt}^{ma} + \psi_{mt}ad_{jmt}^g$$

where $\psi_{mt} \in [0, 1]$ is a number we use to adjust ad_{jmt}^g . An important issue here is the choice of ψ_{mt} . For example, if $\psi_{mt} = 1$, the total advertising spending for MA will simply be the sum of the two kinds of advertising expenditures, which may overstate “true” MA advertising spending. For our analysis, we use ψ_{mt} equal to the fraction of the population that is at least 65 years old in each advertising market. Although the choice of ψ_{mt} is not likely to lead to a perfect measure of advertising expenditures for MA, the choice of ψ_{mt} will be a reasonable proxy for the relative importance of MA business for a firm operating in a local advertising market.¹³

In our analysis, we do not distinguish between an insurer's advertising expenditures in different media.¹⁴ Instead, we use an insurer's total advertising expenditure in a local advertising market by summing the insurer's advertising expenditures across all media in the market. In analysis, we also use an insurer's total advertising quantity. Because information on advertising quantity is available only for TV and newspaper advertising, and because a unit of TV advertising is very different from a unit of newspaper advertising, we measure an insurer's advertising quantity in terms of TV-advertising-equivalent quantity. We construct this variable by dividing an insurer's total advertising expenditures in a local advertising market by the average cost of a unit of TV advertising in the market.

¹³We plan to conduct robustness checks for the choice of ψ_{mt} .

¹⁴We make this choice for two reasons. The first reason is that advertising in different media does not have very distinctive effects on demand in our preliminary analysis. The second reason is that because we endogenize advertising choices in the model, and because we simulate advertising equilibrium in our counterfactual analysis, we did not want to add multiple advertising variables for which we would need to find new equilibria.

Plan-level Data

The Medicare Compare Database is released each year to inform Medicare beneficiaries which private insurers are operating in their county, what plans they offer, and what benefits and costs are associated with each plan. We take a variety of plan benefit characteristics from the data such as premiums, dental coverage, vision coverage, brand and generic prescription drug coverage, and the copayments associated with prescription drugs, primary care doctor visits and specialist visits, emergency room visits, skilled nursing facility stays, and inpatient hospital stays. In addition to information about plan benefits, the data also provide information from report cards on MA plan quality.¹⁵ We use four measures of plan quality: ease of getting referral to specialists, overall rating of health plan, overall rating of health care received, and how well doctors communicate.

The CMS State-County-Plan (SCP) files provide the number of Medicare beneficiaries, number of enrollees of each MA insurer, and average capitation payments in each county-year. A problem with this dataset is that although many insurers offer multiple plans in the same county, the aggregate enrollment information is at the insurer-county-year level, not at the plan-insurer-county-year level. One way to deal with this issue is by taking the average of characteristics of plans offered by an insurer as representative characteristics of the insurer; and another approach is to take the base plan of each MA insurer as a representative plan because the base plan is usually the most popular.¹⁶ ¹⁷ For the current version of this study, we take the first approach, and, as a result, each MA insurer will

¹⁵Dafny and Dranove (2008) find that the report cards on MA plan quality had an impact on demand for MA plans.

¹⁶Previous research on MA also faced the same issue and had to deal with the issue in one of these ways. For examples, see Hall (2007); Nosal (2012).

¹⁷Another approach taken previously by Lustig (2011) is to use the individual-level data, MCBS. This dataset contains beneficiaries' answers to questions about characteristics of MA plans they chose such as premium paid, whether it provides vision, hearing, prescription drug coverage, etc. Using this information, Lustig (2011) was able to match plans chosen by individuals in the MCBS with a specific plan. In the current version of this study, we do not take this approach for two reasons. First, information on an individual's choice of a specific plan is not the most important information for us given our focus on an individual's choice of an MA insurer. Second, the approach requires extensive data work because we have to compare the characteristics of an individual's plan to the characteristics of each plan offered by an insurer to match an individual with a specific plan. However, we plan on conducting robustness checks with this approach later when revising this study.

Mean Utility	Interaction with Health Status
Generic drug	Drug coverage (Generic + Brand)
Brand drug	Inpatient copay ≤ 5 Days
Unlimited Drug Coverage	Nursing Home copay ≤ 20 Days
Dental	Emergency care copay
Routine Eye Exam	Primary care physician copay
Glasses	Specialist copay
Hearing Aids	ease of getting referral to specialists
Hearing Exam	overall rating of health plan
Nursing Home Copay ≤ 20 Days	Dummy for Secure Horizon
Nursing Home Copay ≤ 100 Days	Dummy for United Healthcare
Emergency Care Copay	Dummy for Kaiser Permanente
Emergency Care Coinsurance	Dummy for Blue Cross Blue Shield
ER Worldwide Coverage	Dummy for Aetna
Primary Physician Copay	Dummy for Humana
Primary Physician Coinsurance	Dummy for Health Net
Specialist Copay	
Specialist Coinsurance	
Inpatient Copay ≤ 5 Days	
Inpatient Copay ≤ 90 Days	
Inpatient Coinsurance	
ease of getting referral to specialists	
overall rating of health plan	
overall rating of health care received	
doctors communicate well	
# plans offered by a Firm-county-year	

Table 3.3: Plan Characteristics Included in Analysis

Note: Dummies for different brands are implicitly included in insurer-county fixed effects in the mean utility.

	(1)	(2)	(3)
No. of county-year	800	1264	385
Population (age ≥ 65)/ county	15,589	34,220	113,047
% MA enrollment	0.07	0.21	0.32
Monthly Capitation Payments / enrollee	\$520.3	\$534.2	\$613.3
Monthly Medicare costs/ enrollee	\$458.4	\$459.1	\$566.3
No. of Firms	1.30	2.63	5.19
No. of Firms with Advertising	0	1.88	4.43
Premium for Firms w/ ad < mean	\$57.8	\$42.7	\$29.9
Premium for Firms w/ ad > mean	n/a	\$39.0	\$47.4
Avg Market Shares for Firms w/ ad < mean	0.05	0.08	0.06
Avg Market Shares for Firms w/ ad > mean	n/a	0.12	0.11
Total Number of Insurer-county-year	893	2648	1440

Table 3.4: Summary Statistics at County Level

Note: Column (1) is about counties belonging to a local advertising market having no advertising spending; Column (2) is about counties belonging to a local advertising market where total advertising spending is below \$250,000; and Column (3) is about counties belonging to a local advertising market where total advertising spending is at least \$250,000. Source: AdSpender 2000–2003; CMS state-county-plan files 2000–2003.

have only one representative plan available in each county in analysis.

3.3.2 Preliminary Analysis

In this section, we provide summary statistics from the data and descriptive evidence on how advertising relates to risk selection. Table 3.4 displays characteristics of counties depending on total advertising spending in a local advertising market to which a county belongs. Although there are plenty of counties having no advertising spending, these counties are small in population. There is also a strong correlation between advertising and other county-level characteristics. Counties with larger advertising expenditures tend to have a larger fraction of Medicare beneficiaries in MA, higher capitation payments, higher health care costs in terms of traditional Medicare reimbursement rates, and more MA insurers.

Table 3.5 shows the presence of strong incentives for risk selection in MA. A common

Self-reported Health Status		Market Categories		
		(1)	(2)	(3)
Excellent or Very Good	Capitation (\$)	435.1	450.4	520.0
	Health Expenditures (\$)	213.2	225.2	257.9
	Over-payments (\$)	221.9	225.2	262.1
Good	Capitation (\$)	440.0	464.3	536.4
	Health Expenditures (\$)	394.9	385.4	444.7
	Over-payments (\$)	45.1	78.9	91.7
Fair or Poor	Capitation (\$)	454.6	470.5	549.4
	Health Expenditures (\$)	721.3	736.1	912.7
	Over-payments (\$)	-266.7	-265.7	-363.3
Number of Observations		2729	7594	7729

Table 3.5: Incentives for Risk Selection

Note: Column (1) is about counties belonging to a local advertising market without any advertising spending; Column (2) is about counties belonging to a local advertising market where total advertising spending is below \$250,000; and Column (3) is about counties belonging to a local advertising market where total advertising spending is at least as large as \$250,000.

Source: Medicare Current Beneficiary Survey 2000–2003; AdSpender 2000–2003

pattern observed in this table is that monthly capitation payments do not account for the large variation in health expenditures across individuals having different health statuses. MA insurers are paid capitation payments greater than necessary to cover the health expenditures of relatively healthy individuals whereas capitation payments for relatively unhealthy individuals are not sufficient to cover their health expenditures. As a result, MA insurers would have very strong incentives to selectively enroll healthier individuals in any county. Moreover, there is regional variation in incentives for risk selection. In counties belonging to local advertising markets with relatively large advertising spending, enrolling healthy individuals is more profitable, and enrolling unhealthy individuals results in a larger loss.¹⁸

In order to investigate incentives for risk selection and their regional variation more

¹⁸The regional variation results from the fact that the average and variance of health expenditures are positively correlated. In a region where health care is more expensive, the average health expenditure is higher. At the same time, the variance of health expenditures across individuals is also greater in the region because it is usually the health expenditures of unhealthy individuals that increase disproportionately more in a more expensive region.

precisely, we run the following regression with the individual-level data:

$$Overpayment_i = \beta_1 rh_i + \beta_2 rh_i \times cap_{ct} + \beta_3 cap_{ct} + X_i \gamma + \epsilon_i$$

$Overpayment_i$ is the difference between individual i 's capitation payment and health status (measured in terms of expected traditional Medicare claims costs), which are calculated with the individual-level data; rh_i is individual i 's relative health status, which is defined as a ratio of individual i 's health status to the average Medicare claims cost in county c where individual i resides in year t ; cap_{ct} is the average capitation payment in county c in year t ; and X_i is a vector of other controls that determine the capitation payment for individual i such as age, Medicaid status, and institutional status. Regression results are presented in Table 3.6. Because the minimum county-level average capitation payment is larger than 200, $\hat{\beta}_1 + \hat{\beta}_2 cap_{ct} < 0$ in any county in any year. This means that more over-payments will be made in regions having healthier individuals (lower rh_i). Moreover, $\hat{\beta}_2 rh_i + \hat{\beta}_3 > 0$ for $rh_i < 0.97$, and the median and mean of rh_i are 0.6 and 0.89, respectively. This means that over-payments for relatively healthy individuals are greater in regions with higher average capitation payments. These results are summarized in Figure 3.1, which is based on an individual of age 75 who is not eligible for Medicaid and not living in a nursing home. The plots show that MA plans can increase profit by enrolling healthier individuals and that risk selection is more profitable in regions with higher average capitation payments.

Now given that insurers have more incentives for risk selection in regions with higher capitation payments, we investigate how an insurer's advertising in a local advertising market is related to regional variation in capitation payments with the following regressions:

$$ad_{jmt} = \beta cap_{mt} + X_{mt} \gamma + \delta_j + \epsilon_{jmt}$$

$$ad_{jmt} = \beta cap_{mt} + X_{mt} \gamma + \xi_{jm} + \epsilon_{jmt}$$

Dependent Variable = Expected Over-payment		
VARIABLES	Coefficient	Std. Err.
Relative Health	69.37***	(15.31)
Relative Health \times Avg Capitation Paymentl	-1.012***	(0.0278)
Avg Capitation Payment at the county-level	0.990***	(0.0214)
Observations	31,756	
R-squared	0.977	

Table 3.6: Relationship between Health Status and Over-payment by Location
 Note: Other Controls are age, age-squared, age-cubed, Medicaid status, and whether one lives in a nursing home.

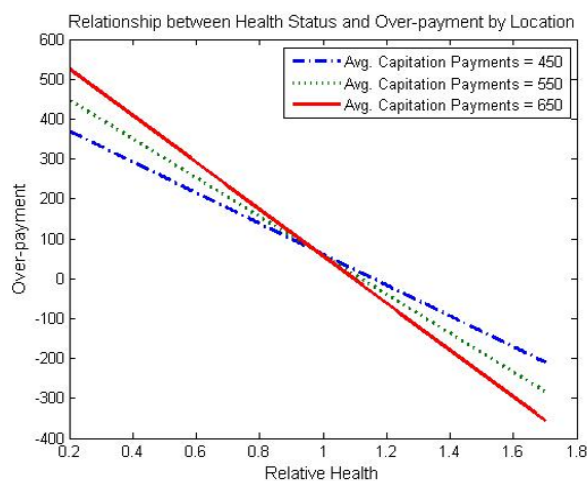


Figure 3.1: Relationship between Health Status and Over-payment by Location
 Note: Relative health = $\frac{\text{Expected Health Expenditure}}{\text{County-level Medicare Costs}}$; median of relative health = 0.6; mean of relative health = 0.89. These plots were generated based on a regression of an amount of over-payment on relative health, average capitation payment in each county, interaction between relative health and average capitation payment, as well as other control variables that determine a capitation payment. The regression results are reported in Table 3.6. The plots were generated for an individual of age 75 that is not eligible for Medicaid and not living in a nursing home. The plots show that over-payments are greater for healthier enrollees and that over-payments for the healthy are greater in regions with higher average capitation payments.

The two regressions are different only with respect to fixed effects. In the first specification, δ_j denotes insurer fixed effects which are invariant over local advertising markets (m). In the second specification, ξ_{jm} denotes insurer-advertising market fixed effects. ad_{jmt} is either an advertising quantity or expenditure by insurer j in local advertising market m in year t , depending on specification.¹⁹ cap_{mt} is the weighted average of cap_{ct} (the average capitation payment in county c in year t) across counties in local advertising market m , with the population of each county as a weight. X_{mt} is a vector of other control variables such as the population of market m , local TV advertising cost and number of competing insurers in an advertising market. The results are reported in Table 3.7. For any specification, the results indicate that more advertising is done in local advertising markets with higher average capitation payments, where over-payments for healthy enrollees are greater. That is, MA insurers' amounts of advertising respond to regional variation in the profitability of risk selection.

Lastly, Table 3.8 shows that individuals with different health statuses are likely to be enrolled with different insurers. MA plans in general tend to have healthier Medicare beneficiaries than traditional Medicare, which is consistent with previous findings on selection into MA. Among MA insurers, moreover, firms with more advertising tend to have healthier enrollees.

3.4 Model

As discussed in a previous section, MA insurers contract with CMS for each county (c) in each year (t). As a result, consumers in different counties face different choice sets. However, each advertising decision is typically made on the basis of a local advertising market (m), which contains several counties. Thus we assume individuals in different c but in the same m are exposed to the same advertising level by the same firm. If county

¹⁹After all, we run four different regressions. Each equation is estimated with each of the two dependent variables.

Dependent Variable	(1)	(2)	(3)	(4)
	Ad Qty		Ad Expenditure	
Avg. Capitation	1.257*** (0.247)	1.032* (0.590)	0.568*** (0.128)	0.693** (0.294)
Population (65 j)	7.94e-05** (4.03e-05)	3.94e-05 (0.000109)	0.000139*** (2.09e-05)	2.48e-05 (5.50e-05)
Local TV Ad Cost	-68.71** (34.16)	-94.20* (48.08)	17.75 (17.71)	-13.37 (23.98)
No. of Competitors	13.60* (7.244)	-20.59* (11.58)	2.291 (3.755)	-9.311 (5.788)
Fixed Effect	Ins.	Ins. - market	Ins.	Ins. - market
Year FE	Yes	Yes	Yes	Yes
R-squared	0.094	0.039	0.185	0.060
Observations	1,035	1,035	1,035	1,035

Table 3.7: Relationship between Advertising and Capitation Payments

Note: The dependent variable in specification (1) and (2) is advertising quantity by an MA insurer in a local advertising market in a year, and that in specification (3) and (4) is advertising spending in a local advertising market in a year. Specification (1) and (3) have market-invariant insurer fixed effects, whereas specification (2) and (4) allow for market-specific insurer fixed effects. Average capitation payments in a local advertising market is the average across average capitation payments in each county belong to the advertising market, weighted by population of each county. The variable, number of competitors, is constructed in a similar way by taking the average across counties with a county population as a weight.

County Category	Insurer Category		
	(1)	(2)	(3)
Counties with total ad spend = 0	0.936	0.930	N/A
Counties with total ad spend $\in (0, \$250K]$	0.918	0.811	0.701
Counties with total ad spend $> \$250K$	0.952	0.767	0.726
Counties with avg capitation $< \$500$	0.919	0.799	0.726
Counties with avg capitation $\in [\$500, \$600]$	0.918	0.818	0.722
Counties with avg capitation $> \$600$	0.989	0.742	0.722
Overall	0.934	0.798	0.722

Table 3.8: Health Status and Insurer Choice by Medicare Beneficiaries

Note: Colum (1) is about Traditional Medicare; Column (2) is about MA insurers with the total advertising spending being below \$150,000; (3) Column (2) is about MA insurers with the total advertising spending being more than \$150,000.

The reported number in each cell is the average relative health status of enrollees in each insurer and market category.

Source: Medicare Current Beneficiary Survey 2000–2003.

c is included in ad market m , we denote $c \in m$.

3.4.1 Demand

Consider a consumer i , living in a county $c (\in m)$ in year t . Consumer i chooses to enroll with one of the available MA insurers in each c and t or in traditional Medicare. We assume that consumer i , living in a county c in year t , obtains indirect utility u_{ijct} from insurer j as follows:

$$u_{ijct} = g(ad_{jmt}, rh_i; \phi) + p_{jct}\alpha_i + x_{jct}\beta_i + \overline{\xi_{jc}} + \Delta\xi_{jct} + \epsilon_{ijct}$$

where

$$\begin{aligned} g(ad_{jmt}, rh_i; \phi) &= (\phi_0 + \phi_1 \log(rh_i)) \times \log(1 + \phi_2 ad_{jmt}); \\ \alpha_i &= \alpha_0 + \alpha_1 \log(rh_i); \\ \beta_i &= \beta_0 + \beta_1 \log(rh_i). \end{aligned}$$

Each insurer has observable characteristics (ad_{jmt} , p_{jct} , and x_{jct}), insurer-county fixed effect ($\overline{\xi_{jc}}$), and an unobservable characteristic ($\Delta\xi_{jct}$). First, ad_{jmt} denotes insurer j 's advertising quantity in advertising market m in year t . The effect of advertising on indirect utility u_{ijct} is captured by $g(ad_{jmt}, rh_i; \phi)$, which depends on individual i 's relative health status (rh_i).²⁰ Parameter ϕ_0 reflects the effects of advertising that are independent of an individual's health status. The effects of advertising on risk selection are captured by its heterogeneous effects on individuals with different rh_i (ϕ_1). We assume that the effects of advertising diminish in its quantity by assuming that ad_{jt} enters $g(\cdot)$ in logarithm. Parameter ϕ_2 determines the curvature of function $g(\cdot)$.

With this specification of u_{ijct} , we assume that advertising affects indirect utility

²⁰ rh_i is defined as a ratio of individual i 's health status (in terms of expected Medicare claims cost) to the average Medicare expenditure in county c where individual i resides in year t . This definition of rh_i is used in the previous section for preliminary analyses.

from an insurer, which is consistent with the persuasive, prestige and signaling effects of advertising. The persuasive and prestige effects of advertising would directly affect utility from an insurer, for example, by creating a certain positive image associated with the insurer (Stigler and Becker 1977; Becker and Murphy 1993). Indeed, many advertisements for MA show images of seniors living healthy lives: engaging in physically demanding activities like running and golfing (Neuman et al. 1998; Mehrotra et al. 2006). These advertisements may create a positive image associated with an insurer and lead to a higher utility level from a plan of that insurer. The signaling effects of advertising will affect demand for an insurer through expected utility by giving a signal about the (unobservable) quality of the insurer (Nelson 1974; Milgrom and Roberts 1986). Because indirect utility u_{ijct} is supposed to capture expected utility from an insurer, $g(ad_{jmt}, rh_i; \phi)$ will contain both effects of advertising. Another possible effect of advertising we do not exactly model is the provision of information about the existence of a product, which is likely to affect an individual's consideration set. If advertising in MA indeed has such effects, they will be captured as an increase in $g(ad_{jmt}, rh_i; \phi)$ because we do not model the effects of advertising on an individual's consideration set.²¹

p_{jct} denotes the premium of plan jct which a consumer pays in addition to the Medicare Part B premium.²² The effect of p_{jct} on utility is also potentially heterogeneous depending on an individual's health status. This is captured by parameter α_1 . x_{jct} describes plan jct 's characteristics other than ad_{jmt} and p_{jct} . For example, x_{jct} includes copayments for a variety of medical services such as inpatient care and outpatient doctor visits and variables

²¹Effects of advertising on a consumer's consideration set would be especially important in an environment where the number of available insurers is so large that consumers cannot easily know about available options. In the MA market, however, the number of available insurers is limited for many individuals. About 40% of Medicare beneficiaries have at most two insurers available in their county of residence; and about 70% of Medicare beneficiaries have at most four insurers available in their county of residence. Thus, although the informative effects of advertising can be still important in the MA market, the effects are not likely to be as important as in markets with a large number of available products.

²²When enrolling in a MA plan, an individual must pay the Medicare Part B premium as well as the premium charged by the plan. Here I did not include Medicare Part B premium in p_{jct} because almost all Medicare beneficiaries, who remain in traditional Medicare, enroll in Medicare Part B and pay the Medicare Part B premium.

describing drug coverage, vision coverage, dental coverage, etc. x_{jct} also includes quality measures of insurers taken from report cards on MA plan quality. The quality measures included in x_{jct} are ease of getting a referral, overall rating of health care received through a MA plan, and how well doctors in a MA plan communicate.²³ With these quality measures, we can control for an insurer’s characteristics that would be usually considered unobserved. The effects of x_{jct} are potentially heterogeneous with parameter β_1 capturing the differential effects of x_{jct} on individuals having different health statuses.²⁴

$\overline{\xi_{jc}}$ denotes insurer-county fixed effects that capture time-invariant unobserved characteristics of insurer j in county c such as size and quality of the insurer’s networks in a region. An individual’s utility also depends on aspects of an insurer that are unobserved by researchers but observed by consumers and insurers. $\Delta\xi_{jct}$ is a time-specific deviation from $\overline{\xi_{jc}}$. $\Delta\xi_{jct}$ captures time-varying unobserved characteristics and/or shocks to demand for this insurer. We assume that $\Delta\xi_{jct}$ is known by consumers and insurers when they make decisions. Lastly, ϵ_{ijct} is idiosyncratic preference shock, which we assume is drawn from Type I extreme value distribution and i.i.d across individuals, insurers, counties and years.

In the model, the outside option is to enroll in traditional Medicare, from which a consumer receives utility of u_{i0ct} :

$$u_{i0ct} = z_i\lambda + \epsilon_{i0ct}.$$

z_i is a vector of an individual’s characteristics including relative health status (rh_i), age, Medicaid status, and whether the individual receives insurance benefits from an (former) employer. These individual characteristics in u_{i0ct} will control for the possibility of different values of the outside option relative to MA, depending on individual characteristics.

²³A detailed list of the variables used in analysis is reported in the Appendix.

²⁴In order to reduce the number of parameters to be estimated, we do not interact every variable in x_{jct} with health status. We select which variables to interact with health status based on the results of the preliminary analysis. A complete list of variables interacted with health status is reported in the Appendix.

For example, Medicaid-eligible individuals will receive more comprehensive coverage in traditional Medicare without having to pay an additional premium. Those who receive insurance benefits from employers will also have a different value of the outside option compared to individuals only with basic Medicare Parts A and B coverage. Moreover, many Medicare beneficiaries in traditional Medicare purchase Medicare supplement insurance (so-called Medigap). Medigap is used in conjunction with traditional Medicare and covers out-of-pocket expenditure risks of individuals in traditional Medicare.²⁵ Because we do not allow for an additional choice of purchasing Medigap in the model, the utility from the possibility of purchasing Medigap is included in u_{ioct} . Previous research on Medigap finds that selection into Medigap depends on an individual's characteristics such as health status (Fang et al. 2008). Then coefficient λ will also capture heterogeneous preference for Medigap depending on z_i . Moreover, it is possible that individuals have different preferences for MA. For example, unhealthier individuals may dislike common aspects of MA plans such as restricted provider networks and referral requirements for specialized treatments. In this case, parameter λ will also capture heterogeneous preferences for MA.

With the functional-form assumption on ϵ_{ijct} , we can analytically calculate the probability for an individual i with characteristics z to enroll plan jct . By defining $u_{jct}(z_i) \equiv u_{ijct} - \epsilon_{ijct}$, we can write the choice probability for plan jct as follows:

$$q_{jct}(z) = \frac{\exp(u_{jct}(z))}{\exp(u_{0ct}(z)) + \sum_{k \in J_{ct}} \exp(u_{kct}(z))} \quad (3.1)$$

Then aggregate market share for a firm jct is

$$Q_{jct} = \int_z q_{jct}(z) dF_{ct}(z) \quad (3.2)$$

where $F_{ct}(z)$ is the distribution of individual characteristics z in county c and year t .

²⁵About 25% of Medicare beneficiaries purchase a Medigap plan.

3.4.2 Supply

We assume that insurers play a simultaneous game in choosing optimal pricing and advertising in each advertising market. In the model, a pricing decision is made for each county (c) in each year (t), and an advertising decision is made for each advertising market (m) in each year (t).

When insuring an individual with health status h (a nominal health expenditure, not relative health rh) with plan characteristics x_{jct} and market characteristics w_{ct} , insurer jct expects to incur a marginal cost $c_{jct}(h)$ as follows:

$$c_{jct}(h) = x_{jct}\gamma_1 + w_{ct}\gamma_2 + h\gamma_3 + \psi_j + \eta_{jct}. \quad (3.3)$$

x_{jct} is a vector of plan characteristics which are included in the utility specification of a consumer such as drug coverage, copayment amounts for a variety of services, etc. w_{ct} includes county characteristics that can potentially influence the cost of providing insurance, including the number of hospitals, skilled nursing facilities and physicians in a county. For example, insurers may be able to negotiate lower payments with providers in markets having a large number of physicians and hospitals (Ho 2009). Importantly, the marginal cost of insuring a consumer depends on the consumer's health status h , and this aspect of $c_{jct}(h)$ creates incentives for risk selection. ψ_j is a firm fixed effect that capture different administrative costs and different ways of delivering health care at the firm level (e.g., Aetna, Blue Cross Blue Shield, Secure Horizon, etc.). Lastly, η_{jct} is a firm-county-year-specific shock to marginal costs that is constant across individuals with different h . We assume that η_{jct} is observed by all insurers making pricing and advertising decisions in a market.

Insurer j 's profit from a county c in year t , excluding advertising costs, is given by:

$$\pi_{jct} = M_{ct} \int_z (p_{jct} + cap_{ct}(z) - c_{jct}(h)) q_{jct}(z) dF_{ct}(z).$$

M_{ct} is the population of those who are at least 65 years old in county c in year t , which is the market size; p_{jct} is the premium charged by insurer j in county c in year t ; $cap_{ct}(z)$ is a capitation payment that depends on county, year, age, gender, Medicaid status and institutional status; and $q_{jct}(z)$ is demand for insurer j by an individual having characteristics z in (3.1).

Because each insurer makes an advertising decision for each advertising market, we need to consider an insurer's profit in an advertising market in order to analyze its advertising choice. An insurer j 's profit in advertising market m and year t is:

$$\pi_{jmt} = \sum_{c \in m} \pi_{jct} - mc_{jmt} ad_{jmt}$$

where mc_{jmt} is constant marginal cost per unit of advertising. We assume that

$$mc_{jmt} = \exp(x_{jmt}^{ad} \gamma_{ad} + \zeta_{jmt}).$$

x_{jmt}^{ad} includes the costs of TV advertising in media market m in year t , year dummies, and dummy variables for large firms. We included eight dummy variables for each of the eight largest firms. These dummy variables will capture different resources constraints faced by different firms.²⁶ ²⁷ ζ_{jmt} is a shock to the marginal cost, which is also known by all insurers in a media market, but unobserved by researchers. We assume that $\zeta_{jmt} \sim N(0, \sigma_\zeta^2)$.²⁸

Nash equilibrium conditions for the game for insurers are that insurers' choices maximize their profits given choices made by other insurers. For an insurer's optimal pricing

²⁶Although marginal cost of advertising is assumed to be constant, some firms using large amounts of advertising may face different advertising costs due to volume discounts. The dummy variables can capture the different discounts received by different firms having potentially different advertising amounts.

²⁷Included insurers are Secure Horizon, Blue Cross Blue Shield, Kaiser Permanente, United Healthcare, Aetna, Humana, Health Net, and Cigna. Although Secure Horizon is currently part of United Healthcare, they were separate companies during the period of 2000–2003.

²⁸The reason that we make a functional-form assumption for ζ will be discussed in the section for identification and estimation.

condition, we have the following condition for each p_{jct} :

$$\frac{\partial \pi_{jmt}}{\partial p_{jct}} = \frac{\partial \pi_{jct}}{\partial p_{jct}} = 0. \quad (3.4)$$

An insurer's optimal advertising conditions are:

$$\frac{\partial \pi_{jmt}}{\partial ad_{jmt}} = \sum_{c \in m} \frac{\partial \pi_{jct}}{\partial ad_{jmt}} - mc_{jmt} \begin{cases} = 0 & \text{for } ad_{jmt} > 0 \\ \leq 0 & \text{for } ad_{jmt} = 0 \end{cases}. \quad (3.5)$$

For the optimal advertising condition, we explicitly allow for the possibility of the corner solution, which is no advertising.²⁹ Because about 35% of insurer (j)-market (m)-year (t) combinations do not advertise at all, we have to explicitly allow for the possibility that insurers choose the corner solution. Condition (3.5) states that when an insurer spends a positive amount of advertising spending, the optimal quantity of advertising maximizes its profit, and that when an insurer does not advertise, its profit gain from a small quantity of advertising should not be greater than its cost.

3.5 Identification and Estimation

For the discussion of identification and estimation of the model, we define θ as a vector that contains all parameters in the model such that $\theta = (\theta^d, \theta^s)$. θ^d and θ^s are vectors of parameters that enter the demand and supply side, respectively.

²⁹Although premiums can be zero, we assume that even zero premium satisfies the pricing first order condition with equality. This assumption is made mainly for computational convenience when solving the model in counterfactual analysis. If an insurer chooses zero premium due to the constraint of nonnegative premium, it is possible that we overestimate the marginal cost of providing insurance for insurers having zero premium.

3.5.1 Demand

Mean Utility In utility u_{ijct} , there are two kinds of parameters: θ_1^d and θ_2^d . We define θ_1^d to be parameters that enter ‘mean utility’ δ_{jct} , which is a part of u_{ijct} that does not depend on individual characteristics. Precisely,

$$\delta_{jct} = \phi_0 \log(1 + \phi_2 ad_{jmt}) + \alpha_0 p_{jct} + x_{jct} \beta_0 + \overline{\xi_{jc}} + \Delta \xi_{jct}. \quad (3.6)$$

θ_2^d is defined as parameters for interaction terms between insurer characteristics and individual characteristics. We let ϕ_2 , which determines diminishing returns of advertising effects, be a part of θ_2^d . Berry et al. (1995) show that given a value for θ_2^d , there is a unique $\delta_{jct}^*(\theta_2^d)$ that solves for the system of equations given by the aggregate market share equation (3.2). Then parameter θ_1^d is estimated using equation 3.6. A well-known problem regarding identification of θ_1^d is that the unobserved characteristic ($\Delta \xi_{jct}$) and two endogenous variables in the model (p_{jct} and ad_{jmt}) are correlated, because $\Delta \xi_{jct}$ is assumed to be known by consumers and insurers when they make decisions. This problem is a typical endogeneity problem, and then a simple ordinary least squared regression of $\delta_{jct}^*(\theta_2^d)$ on the observed variables in (3.6) will result in inconsistent estimates of θ_1^d .

Although the endogeneity problem causes challenges in identification, fixed effects $\overline{\xi_{jc}}$ in δ_{jct} would control for a significant part of the unobserved heterogeneity of insurers. Important characteristics that are not included in x_{jct} are an insurer’s network size and quality in a local market. For example, Kaiser Permanente, which is one of the largest insurers in California, has a more extensive network in California than in other regions. As long as such characteristics do not vary much over the time period considered in this study, they will be controlled for by $\overline{\xi_{jc}}$. Moreover, x_{jct} includes an insurer’s quality measures from report cards on MA plan quality, such as ease of getting a referral, overall rating of an insurer, overall rating of health care received, and how well an insurer’s physicians communicates with patients. By including these characteristics, we will be able to control

for characteristics that would usually be considered unobservable.

However, it is still possible that x_{jct} cannot capture all relevant characteristics of an insurer that vary over time, which will result in the endogeneity problem. A typical approach to accounting for the endogeneity problem is to use instruments that are correlated with the endogenous variables, but not with the unobservable. We construct two sets of instruments. The first set of instruments are the averages of premiums and advertising of the same parent company in other advertising markets. The use of functions of endogenous variables in other counties as instruments is a strategy similar to Hausman (1996) and Nevo (2001). Town and Liu (2003) use similar instruments in estimating a model of demand for MA plans. The identifying assumption is that demand shock $\Delta\xi_{jct}$ is not correlated with shocks affecting the premiums of insurer j in other markets, such as demand and marginal cost shocks in the markets. A similar identifying assumption is made for advertising of the same firm in other markets. A premium in a county will be correlated with the average premiums of the same firm in other markets through, for example, common company-level components affecting premiums. The same argument also holds for advertising.

The second set of instruments are variables that affect a plan's premium and advertising choices, but do not affect utility directly. One such variable is the cost of a unit of TV advertising in a local advertising market, which affects an advertising decision, but does not affect utility directly. Other such variables are capitation payments in other counties in the same advertising market. Because capitation payments in other counties in the same advertising market will affect advertising in the advertising market, the payments in other counties can be valid instruments as long as they do not enter the utility of a consumer in a county.³⁰

³⁰The second instrument using capitation payments in other counties is similar to the instruments used in Nosal (2012), who studies demand for MA plans.

Resulting moment conditions employed in the estimation are:

$$E[\Delta\xi_{jct}|\Gamma] = 0. \tag{3.7}$$

Γ is a set of instruments that includes the aforementioned two sets of instruments as well as x_{jct} .

Preference Heterogeneity Important information for identification of parameters for preference heterogeneity θ_2^d is an individual's insurer choice from the MCBS (the individual-level data). Parameter θ_2^d will be identified by variation in the characteristics of insurers chosen by individuals having different characteristics. Identification of θ_2^d is aided by variation in insurer characteristics, not only across insurers within a region but also across regions. For example, advertising quantities vary across local advertising markets depending on how profitable risk selection is in the market, as illustrated in the previous section for preliminary analysis. Moreover, individuals in different regions will have different choice sets, and this variation in choice sets provides information on the substitution patterns of different individuals.

An important parameter in θ_2^d is the parameter that determines the heterogeneous effect of advertising depending on an individual's health status (ϕ_1), which captures the effect of advertising on risk selection. A potential concern in identifying ϕ_1 is that there may be insurer characteristics, not included in x_{jct} but correlated with ad_{jmt} , that have different effects on the demand of individuals having different health status. Given the available data, it is impossible to allow for insurer-county fixed effects $\overline{\xi_{jc}}$ that depend on an individual's health status and to control for them.³¹ In order to alleviate this concern, we interact many different variables in x_{jct} with health status, including not only usual characteristics such as drug coverage and copayments but also the quality measures from report cards on MA plans and dummy variables for each of the seven largest

³¹If there is information on an insurer's aggregate market share by different health statuses, it is possible to allow for $\overline{\xi_{jc}}$ that depends on health status.

insurers. The latter variables are highly correlated with ad_{jmt} , and their interactions with health status will limit the role of omitted insurer characteristics that can have differential effects on individuals having different health statuses. The quality measures will control for important aspects of insurers, with potential heterogeneous effect, that cannot be described by usual coverage characteristics. Moreover, an interaction between a dummy variable for a large insurer and health status will capture an aspect of the insurer that may have differential effects on individuals having different health statuses.

In order to construct micro-moments for an individual's insurer choice and combine them with aggregate moments (3.7), we use the score of the log-likelihood function for a choice by an individual observed in the MCBS, as in Imbens and Lancaster (1994). The likelihood function for an individual's choice is:

$$L = \prod_{i,j,c,t} q_{jct}(z_i)^{d_{ijct}}$$

where z_i is a vector of characteristics of individual i in the individual-level data; and d_{ijct} is an indicator variable that equals one when individual i chooses plan jct . Then our micro-moments are

$$\frac{\partial \log(L)}{\partial \theta_2^d} = 0. \tag{3.8}$$

3.5.2 Supply

Cost of Providing Insurance Estimation of parameters of the supply side relies on the optimality conditions for pricing and advertising, presented in (3.4) and (3.12). The first order condition for optimal pricing (3.4) is equivalent to the following condition:

$$\begin{aligned} \frac{Q_{jct} + \int_z (p_{jct} + cap_{ct}(z)) \frac{\partial q_{jct}(z)}{\partial p_{jct}} dF_{ct}(z)}{\frac{\partial Q_{jct}}{\partial p_{jct}}} &= \frac{\int_z c_{jct}(h) \frac{\partial q_{jct}(z)}{\partial p_{jct}} dF_{ct}(z)}{\frac{\partial Q_{jct}}{\partial p_{jct}}} \\ &= x_{jct} \gamma_1 + w_{ct} \gamma_2 + H(q_{jct}, F_{ct}) \gamma_3 + \psi_j + \eta_{jct} \end{aligned} \tag{3.9}$$

where $q_{jct}(z)$ and Q_{jct} are demand of an individual with characteristic z (which includes h) and aggregate demand for insurer j in county c in year t , respectively; and

$$H(q_{jct}, F_{ct}) \equiv \frac{\int_z h \frac{\partial q_{jct}(z)}{\partial p_{jct}} dF_{ct}(z)}{\frac{\partial Q_{jct}}{\partial p_{jct}}}.$$

An examination of (3.9) reveals that its left-hand side is a function of demand side parameters and data. Because demand side parameters can be identified with only the demand model and data, the left-hand side of (3.9) can be treated as known. Then optimality condition (3.9) leads to a linear estimating equation. Because we assume that an insurer's choice of x_{jct} is exogenous to the model, and because market characteristics w_{ct} are exogenous, we have the following moment conditions:

$$E[\eta_{jct}|x_{jct}] = 0 \text{ and } E[\eta_{jct}|w_{ct}] = 0. \quad (3.10)$$

These assumptions will identify parameters γ_1 and γ_2 .

However, we cannot have a similar condition for parameter γ_3 because $H(q_{jct}, F_{ct})$ is potentially endogenous to η_{jct} . Because an insurer's choice of p_{jct} will be directly dependent on η_{jct} in the model, and because p_{jct} will determine $q_{jct}(z)$, variable $H(q_{jct}, F_{ct})$ may be correlated with η_{jct} . This endogeneity problem necessitates an instrument that is correlated with $H(q_{jct}, F_{ct})$, but not with η_{jct} . In order to find an instrument for $H(q_{jct}, F_{ct})$, it is important to understand what $H(q_{jct}, F_{ct})$ means. By definition, $H(q_{jct}, F_{ct})$ measures the average health status of consumers switching from insurers jct to other insurers due to an increase in a premium of insurer jct . Because an individual's health status h is measured as expected claims cost for Medicare Parts A and B, an important determinant of $H(q_{jct}, F_{ct})$ is overall health care cost in county c in year t . As a result, $H(q_{jct}, F_{ct})$ must be highly correlated with county-level average Medicare claims cost FFS_{ct} , which exhibits large variation across counties. Since we control for market characteristics w_{ct} that may influence an insurer's marginal cost, it is likely that FFS_{ct} is uncorrelated with

η_{jct} , which leads to the identifying assumption for γ_3 such that

$$E[\eta_{jct}|FFS_{ct}] = 0. \quad (3.11)$$

Advertising Cost The optimality condition for an advertising quantity (3.5) identifies parameter γ_{ad} in advertising marginal cost mc_{jmt} . This condition is equivalent to the following condition:

$$\zeta_{jmt} \begin{cases} = \log \left(\sum_{c \in m} \frac{\partial \pi_{jct}}{\partial ad_{jmt}} \right) - x_{jmt}^{ad} \gamma_{ad} & \text{for } ad_{jmt} > 0 \\ \geq \log \left(\sum_{c \in m} \frac{\partial \pi_{jct}}{\partial ad_{jmt}} \right) - x_{jmt}^{ad} \gamma_{ad} & \text{for } ad_{jmt} = 0 \end{cases} \quad (3.12)$$

As is clear in (3.12), the optimality condition for insurers using zero advertising results in an inequality condition, which creates a challenge in estimation and identification. We deal with this problem by assuming a functional form for the distribution for advertising cost shock ζ_{jmt} such that $\zeta_{jmt} \sim N(0, \sigma_\zeta^2)$.^{32 33} In order to set up moment conditions, we use the score of the log-likelihood function for each insurer's observed advertising quantity choice, using the first order conditions (3.12). The likelihood function for the advertising choice is:

$$\Gamma = \prod_{j,m,t} f_\zeta(\zeta_{jmt}^*) \mathbf{1}^{[ad_{jmt}>0]} (1 - F_\zeta(\zeta_{jmt}^*)) \mathbf{1}^{[ad_{jmt}=0]}$$

where $\zeta_{jmt}^* = \log \left(\sum_{c \in m} \frac{\partial \pi_{jct}}{\partial ad_{jmt}} \right) - x_{jmt}^{ad} \gamma_{ad}$, and f_ζ and F_ζ are the pdf and cdf of ζ . Then the moment conditions for advertising cost are

$$\begin{aligned} \frac{\partial \log(\Gamma)}{\partial \gamma_{ad}} &= 0 \\ \frac{\partial \log(\Gamma)}{\partial \sigma_\zeta} &= 0. \end{aligned} \quad (3.13)$$

³²Goeree (2008) faces the same problem of rationalizing zero advertising by some firms in the personal computer market, and she also deals with this problem by making a functional-form assumption for the unobservable.

³³Note that a function-form assumption is not necessary for η when estimating the parameters in the marginal cost of providing insurance because there are no inequality optimality conditions for pricing.

An alternative approach, not taken in this study, is to set-identify γ_{ad} using the moment inequality method as in Pakes et al. (2011), which will result in an upper and lower bound for γ_{ad} . If the moment inequality method is used, it will be straightforward to calculate a lower bound by calculating an increase in profits (excluding advertising cost) when insurers increase a unit of advertising from the amount observed in the data. Marginal cost of advertising must be greater than the calculated increase in profits because the observed advertising quantity is assumed to maximize profits. A moment for a lower bound is calculated by averaging over each insurer's lower bounds for advertising cost.

A natural way to derive an upper bound of advertising cost is to calculate the decrease in profits when insurers decrease a unit of advertising from the observed advertising choice. However, deriving the upper bound is more challenging in this model because some insurers choose zero advertising and because an advertising quantity cannot be negative. As a result, we can calculate upper bounds only for insurers that choose positive advertising quantities. Because we can only average over insurers with positive advertising for a moment for the upper bound, we will have a selection problem. However, Pakes et al. (2011) show that if a researcher assumes that ζ comes from a symmetric distribution, it is still possible to derive an upper bound.

A tradeoff between the two approaches to dealing with the inequality first order conditions is that a functional-form assumption on ζ can lead to point-identification of parameters at the cost of a stronger assumption on unobservable ζ . However, the moment inequality method is not completely free of an assumption on ζ either. For this reason, we choose to make a functional-form assumption.³⁴

3.5.3 Estimation Algorithm

The demand and supply models are estimated separately in two steps. The estimation method we use is generalized method of moments. First, we estimate the demand model

³⁴For robustness checks, we plan to check how our results depend on different assumptions on ζ and to estimate the model with the moment inequality method.

using moments (3.7) and (3.8) with the nested fixed point algorithm as in Berry et al. (1995). We define $G_d(\theta^d)$ to be a vector of the moments for the demand side. Our criterion function is given by $\Psi_d(\theta^d) = G_d(\theta^d)'WG_d(\theta^d)$ where W is a weighting matrix. Our estimation routine searches for θ^d that minimizes $\Psi_d(\theta^d)$. Evaluation of $G_d(\theta^d)$ can be broken into the following steps for each choice of θ^d :

1. Given θ^d , we solve for mean utility $\delta^*(\theta^d) = \{\delta_{jct}^*(\theta^d)\}_{j,c,t}$ that satisfies the conditions for aggregate market shares (3.2), using the contraction mapping used in Berry et al. (1995).
2. With θ^d and $\delta^*(\theta^d)$, we calculate the demand $q_{jct}(z)$ of an individual with characteristic z using equation (3.1).
3. We evaluate $G_d(\theta^d)$ with $q_{jct}(z)$.

Once we estimate θ^d , the supply model is estimated using moments (3.10), (3.11), and (3.13).

3.6 Estimates

3.6.1 Utility

Table 3.9 displays estimates for the parameters of primary interest. The estimate of the parameter for the differential effects of advertising on utility is negative, which means that the effects of advertising are greater for healthier consumers because a healthier individual has lower rh_i . The total effect of advertising on an individual with relative health status rh_i is $\phi_0 + \phi_1 \log(rh_i)$. In the data, the median of $\log(rh_i)$ is -0.6, and the value of $\log(rh_i)$ is negative for a majority of individuals.³⁵ As a result, although the estimate for ϕ_0 is not large enough to be statistically significant, $\phi_0 + \phi_1 \log(rh_i)$ will be larger than ϕ_0 for many individuals with $\log(rh_i) < 0$. Moreover, less healthy individuals receive more

³⁵The distribution of rh has a long right-tail. The median of rh is 0.6, and the mean of rh is 0.9.

utility from the outside option than healthier individuals, according to the estimates for the parameters for relative health status in the utility for the outside option. In other words, healthier individuals are more likely to choose MA than less healthy individuals even without advertising. The estimates for price coefficients indicate that individuals receive negative utility from a higher premium, and that healthier individuals are less sensitive to premium although the estimate for α_1 is not statistically significant.

Variables	Estimates	Std. Err.
Ad effects (ϕ_0)	0.040	(0.063)
$\log(rh_i) \times$ Ad effects (ϕ_1)	-0.036**	(0.018)
Curvature of Ad effects (ϕ_2)	0.012**	(0.005)
$\log(rh_i) \times$ Outside option (part of λ)	0.233**	(0.097)
Premium (α_0)	-0.002***	(0.0005)
$\log(rh_i) \times$ Premium (α_1)	8.0e-4	(7.4e-4)

Table 3.9: Estimates for Key Parameters in Utility

Table 3.10 presents semi-elasticities of demand with respect to an increase of \$1,000 in advertising expenditures, which measures percentage change in demand for a \$1,000 increase in advertising expenditures.³⁶ An increase of \$1,000 in advertising expenditures by an insurer increases demand by 0.063% on average. Elasticities for different health statuses show that the effects of advertising are substantially different across individuals having different health statuses. The elasticity for an individual whose rh_i is lower than the 25th-percentile of the distribution of rh_i is more than four times greater than the elasticity for an individual, whose rh_i is more than the 75th-percentile of the distribution of rh_i . Semi-elasticity of demand with respect to a premium is -0.25, which means that a dollar increase in a premium decreases demand by 0.25%. Moreover, healthier individuals' price semi-elasticity is larger in its absolute value than that of less healthy individuals.

The estimates imply that although MA plans are preferred by healthy individuals in general, advertising reinforces the direction of selection into MA. As mentioned in a

³⁶We calculate semi-elasticity instead of elasticity because zero advertising is observed for about 35% of insurers. When an advertising expenditure is zero, elasticity becomes zero. For the same reason, we calculate semi-elasticity for premiums. MA insurers often charge a premium of zero.

Semi-Elasticities of Demand	Ad (\$1,000)	Price (\$1)
Overall Semi-elasticity	0.063%	-0.25%
Semi-elasticity for $h_i < 25\%$	0.092%	-0.28%
Semi-elasticity for $h_i > 25\%$ and $h_i < 50\%$	0.070%	-0.26%
Semi-elasticity for $h_i > 50\%$ and $h_i < 75\%$	0.050%	-0.24%
Semi-elasticity for $h_i > 75\%$	0.020%	-0.22%

Table 3.10: Elasticity of Demand with Respect to Advertising and Premiums

previous section, unhealthy individuals may dislike the HMO aspects of MA plans such as restricted provider networks and referral requirements for specialized medical treatment. These aspects will be especially inconvenient especially for unhealthy individuals, who expect to utilize medical care intensively. In addition to the heterogeneous preferences between healthy and unhealthy individuals for MA, advertising also attracts healthier individuals into MA.

There are several mechanisms to generate the estimated heterogeneous effects of advertising on demand. First, the estimates may reflect contents of advertising designed to be more appealing to healthy individuals, as claimed by Neuman et al. (1998) and Mehrotra et al. (2006). Alternatively, insurance companies may deliberately choose which media to advertise because individuals with different characteristics may be exposed to different media to different degrees. For example, more educated individuals are more likely to read a newspaper, and insurers may target these individuals with newspaper advertising because more educated individuals tend to be healthier.³⁷ Another possibility is that individuals with different health statuses respond differently to the same advertising. In order for an insurer's advertising to induce an individual to enroll with the insurer, the individual must be able to purchase a plan from the insurer. In fact, many Medicare beneficiaries have difficulties with activities related to purchasing a plan according to the individual-level data: About 10% of Medicare beneficiaries have difficulties in using the

³⁷An example of research that studies the effects of advertising in different media on individuals with different characteristics is Goeree (2008), who studies advertising in the U.S. personal computer market. We are unable to incorporate this detailed mechanism of risk selection into our analysis because of the lack of data that relate an individual's characteristics and media consumption patterns.

telephone; about 20% of them have difficulties in shopping for personal items; about 15% of them have difficulties in managing money; and about 50% of them do not use the Internet. Moreover, individuals with such characteristics are more likely to be unhealthy in the data. Then individuals without the difficulties who would be induced by advertising are likely to be relatively healthy.

Estimates for other parameters in utility are reported in Table 3.11 and 3.12. Many variables that enter mean utility are statistically significant. For example, consumers prefer insurers that offer generic and brand drug coverage and drug coverage without an annual coverage limit. However, many variables that interact with health status are not statistically significant. Exceptions are the coefficients for Medicaid status and whether an individual receives health insurance benefits from a (former) employer, which determine heterogeneous utility of the outside option. As expected, individuals on Medicaid are less likely to purchase a MA plan; and individuals with employer-sponsored benefits are also less likely to purchase MA. These estimates result from the fact that having either option usually increases the value of staying in traditional Medicare. Medicaid, combined with Medicare, provides more generous coverage than traditional Medicare, without an additional premium. Moreover, employer-sponsored benefits also provide a cheap option for supplemental coverage without MA plans.

The imprecise estimates for the parameters for most interaction terms imply that many plan characteristics do not have large impacts on the insurer choice of individuals with different health statuses. This may be because variation in the data that identifies the relevant parameters comes from observed insurer choices, not plan choices, by individuals with different health statuses. Even if individuals with different health statuses select into plans with different characteristics within an insurer, an observed insurer choice cannot provide information on such selection patterns unless the characteristics of overall plans of different insurers are very different.³⁸ However, parameters for the effects of

³⁸As a robustness check, we plan to consider the possibility that individuals make a choice at the plan-level, not at the insurer-level. See footnote 17 for details.

VARIABLES	Estimates	Std. Err.
Generic drug	0.423***	(0.147)
Brand drug	0.275***	(0.0399)
Unlimited Drug Coverage	0.105***	(0.0355)
Dental	-0.0318	(0.0362)
Routine Eye Exam	-0.172***	(0.0278)
Glasses	-0.134***	(0.0349)
Hearing Aids	0.204***	(0.0340)
Hearing Exam	0.0982***	(0.0370)
Nursing Home Copay up to 20 Days	-0.00173***	(0.000590)
Nursing Home Copay up to 100 Days	-0.00209***	(0.000383)
ER Copay	-0.000702	(0.00108)
ER Coinsurance	-0.108***	(0.0125)
ER Worldwide Coverage	0.145***	(0.0417)
Primary Physician Copay	-0.00111	(0.00247)
Primary Physician Coinsurance	-0.0446***	(0.00679)
Specialist Copay	-0.00262*	(0.00137)
Specialist Coinsurance	0.0424***	(0.00631)
Inpatient Copay up to 5 Days	9.36e-05	(6.99e-05)
Inpatient Copay up to 90 Days	0.00159***	(0.000275)
Inpatient Coinsurance	-0.0581***	(0.00454)
Quality: ease of getting referral to specialists	-0.00873	(0.0159)
Quality: overall rating of health plan	0.202***	(0.0177)
Quality: overall rating of health care received	-0.0731***	(0.0235)
Quality: doctors communicate well	-0.0378*	(0.0206)
Number of plans offered by a Firm-county-year	0.0268***	(0.00311)
Year FE	Yes	
Firm - county FE	Yes	

Table 3.11: Estimates for Parameters in Mean Utility (δ_{jmt})

insurer-level characteristics, such as advertising quantities and dummy variables for large insurers, will not be affected by our focus on an individual's choice of insurer because these characteristics are constant across each insurer's plans.

3.6.2 Cost

Table 3.13 displays estimates for marginal costs of providing insurance to an enrollee whose specification is given in (3.3). The most important parameter here is the coefficient for health status, which is measured as expected Medicare reimbursement costs. The

Variables	Estimates	Std. Err.
Health×Drug coverage	-0.112	(0.140)
Health×Inpatient copay	-1.2e-5	(2.3e-4)
Health×Skilled nursing facility copay	0.002	(0.001)
Health×Emergency care copay	0.002	(0.002)
Health×Primary care physician copay	-2.3e-4	(0.007)
Health×Specialist copay	-0.005	(0.004)
Health×How easy to get a referral for SP	-0.066**	(0.034)
Health×Overall rating health plan	0.016	(0.037)
Health×Secure Horizon	-7.5e-4	(0.135)
Health×United Healthcare	0.016	(0.119)
Health×Kaiser Permanente	-0.083	(0.142)
Health×Blue Cross Blue Shield	-0.107	(0.074)
Health×Aetna	0.013	(0.102)
Health×Humana	-0.149	(0.140)
Health×Health Net	-0.107	(0.142)
Medicaid×Outside option	1.464***	(0.189)
Employer benefits×Outside option	2.049***	(0.048)
Age×Outside option	-0.107	(2.695)
Age-squared×Outside option	0.046	(0.017)

Table 3.12: Estimates for Preference Heterogeneity

coefficient is very precisely estimated, and its effect is that a one-dollar increase in expected Medicare claims cost leads to an increase of \$0.86 for an MA insurer. This means that the average health status of an insurer’s enrollees is an important determinant of the insurer’s cost of providing insurance, which will create strong incentives to risk-select healthy individuals.

The marginal cost of providing insurance also depends on other characteristics. Notably, county-level characteristics are important determinants of marginal cost. We find that marginal cost increases with population density and with the percentage of the population that lives in urban areas. It may be because counties, which are densely populated and urban, are usually more expensive to operate in. Moreover, the higher the number of hospital beds and skilled nursing facilities, the lower the marginal cost, which is consistent with the finding that these factors determine the relative bargaining power of managed-care firms when setting reimbursement rates to providers (Ho 2009).

Variables	Estimates	Std. Err.
Expected Health Expenditure (h)	0.865***	(0.0320)
Dental	21.00***	(4.532)
Routine eye exam	10.39**	(4.972)
Skilled nursing facility copay	-0.669***	(0.0985)
Emergency care copay	-1.027***	(0.221)
Primary care doctor copay	1.878***	(0.403)
Specialist copay	-0.860***	(0.236)
How easy to get a referral for SP	17.30***	(2.638)
Overall rating health plan	-13.42***	(3.021)
Population density	0.00337***	(0.000221)
Percentage of urban population	0.365***	(0.0565)
No. of hospital beds	-0.395***	(0.0910)
No. of skilled nursing facility	-37.97***	(10.88)
Insurer Dummy	Yes	
Year Dummy	Yes	

Table 3.13: Estimates for Marginal Costs of Providing Insurance

Table 3.14 presents estimates for marginal costs of advertising. The estimates show that local TV advertising costs increase an insurer's marginal cost of advertising and that different firms potentially have different costs of advertising, possibly because the firms face different resource constraints..

Variables	Estimates	Std. Err.
Local TV Advertising Cost	0.491***	(0.123)
Secure Horizon	0.039	(0.135)
United Healthcare	-0.271***	(0.119)
Kaiser Permanente	0.572***	(0.142)
Blue Cross Blue Shield	-1.123***	(0.074)
Aetna	-0.272***	(0.102)
Humana	-0.103	(0.140)
Health Net	-.557***	(0.142)
Standard Deviation of ζ (σ_ξ)	1.356***	(0.579)
Year Dummy	Yes	

Table 3.14: Estimates for Marginal Costs of a Unit of Advertising

3.7 Counterfactual Experiments

With the estimated model, we conduct counterfactual analyses to understand the impacts of advertising on the MA market and how incentives for risk selection affect insurers' advertising decisions.

3.7.1 Ban of Advertising

In this counterfactual analysis, we simulate an equilibrium of the model where advertising is banned. The simulation has two purposes. First, we investigate how advertising affects the choices made by consumers and insurers, and how it affects over-payments by the government. Second, we study how much advertising can account for the selection of healthier individuals into MA.

In implementing this counterfactual analysis, we force each insurer's advertising quantity to zero and let insurers re-optimize their premiums. The results are presented in Table 3.15. We refer to the observed equilibrium in the data as the baseline. The ban on advertising decreases overall MA enrollment by 4% and decreases demand for insurers having above-average advertising expenditures in the baseline by 9%. Although a decrease in demand would usually lead to a lower premium, the ban on advertising does not have a large effect on premiums, which decrease by less than a dollar on average. The negligible effect of advertising on premiums results from the fact that advertising attracts relatively healthy individuals, which lower the costs of providing insurance. With the ban, MA enrollees become less healthy on average, resulting in a larger increase in average health expenditures for insurers having a relatively large amount of advertising in the baseline. For these insurers, an increase in average expected Medicare claims cost is about \$14, which is about 43% of the average premium charged by these insurers. Such an increase in the cost of providing insurance will offset incentives to lower premiums that result from the reduction in demand caused by the lack of advertising.

Table 3.16 presents the results on consumers' welfare. We calculate two different mea-

		Baseline	Ban on Ad
All firms (N = 4864)	Share of beneficiaries	0.243	0.236 (-4%)
	Average Premium (\$)	32.4	31.5
	Average Health Status (\$)	402.3	408.5
	Over-payment per enrollee (\$)	136.9	130.8 (-4%)
Insurers with Above-average ad (N = 881)	Share of beneficiaries	0.101	0.092 (-9%)
	Average Premium (\$)	32.4	31.1
	Average Health Status (\$)	407.4	421.1
	Over-payment per enrollee (\$)	150.4	137.6 (-8%)

Table 3.15: Ban on Advertising

Note: A share of beneficiaries is the fraction of the total Medicare beneficiaries who choose any MA insurers or insurers with above-average advertising spending.

asures of consumers' surplus. In the first measure, we include the effects of advertising on utility whereas we exclude these effects in the second measure. The first measure of welfare is consistent with the informative and complementary view of advertising.³⁹ The informative view holds that advertising provides information about the existence of a product or (unobserved) characteristics of a product that is difficult to be unobserved before consuming the product. As mentioned in the section for the demand model, the effect of advertising on indirect utilities in the model will capture an increase in expected utility due to advertising.⁴⁰ The complementary view holds that consumers receive a higher utility from a product when the product is advertised, which reflects a positive image or greater prestige generated by advertising (Stigler and Becker 1977; Becker and Murphy 1993). Therefore, according to these views, advertising will have a direct impact on an individual's indirect utility from an insurer. When consumers' surplus is calculated according to these views of advertising, we find that consumer welfare decreases because consumers do not receive utility from advertising with the ban and because the ban does not reduce premiums much. The second measure of welfare is supposed to capture the part of utility derived from insurer characteristics other than advertising, which is consistent with the persuasive view of advertising. This view holds that advertising does

³⁹For a discussion of different views of advertising and their welfare implications, see a survey by Bagwell (2007).

⁴⁰For examples, see Stigler (1961); Nelson (1974); Butters (1977); Schmalensee (1977); Grossman and Shapiro (1984); Kihlstrom and Riordan (1984); and Milgrom and Roberts (1986).

not add any real value to consumers (Bagwell, 2007). When consumers' welfare is calculated according to the persuasive view, we find that advertising increases consumers' welfare because advertising just distorts a consumer's decision according to the persuasive view. However, the welfare could have increased even more if the ban on advertising had decreased premiums by a greater amount.

		Baseline	Ban on Ad
Case 1	$rh_i < 25\%$	\$112.8	\$108.3
	$rh_i > 75\%$	\$135.7	\$133.9
	Overall	\$116.6	\$114.8
Case 2	$rh_i < 25\%$	\$101.9	\$108.3
	$rh_i > 75\%$	\$131.6	\$133.9
	Overall	\$109.5	\$114.8

Table 3.16: Consumer's Surplus with a Ban on Advertising

Note: In case 1, the calculation of consumer surplus included the effects of advertising on utility. In case 2, however, we exclude the effects of advertising on utility in the calculation of consumer surplus. That $rh_i < 25\%$ refers to the group of individuals whose relative health status rh_i is below the 25th percentile in the distribution of relative health status. That is, this group is the healthiest. That $rh_i > 75\%$ refers to the group of individuals whose relative health status rh_i is above the 75th percentile in the distribution of relative health status. That is, this group is the most unhealthy.

Now we turn to the second purpose of this counterfactual analysis, which is to investigate how much advertising accounts for the selection of healthier individuals into MA (which is called "advantageous selection", as opposed to adverse selection). In the baseline, MA enrollees are healthier than traditional Medicare enrollees. According to Table 3.17, the average health status of enrollees in traditional Medicare, in terms of Medicare claims cost, is higher than that of MA enrollees by \$60.6. The difference in average health status between the two groups decreases by 15% with the ban on advertising. This means that advertising accounts for 15% of advantageous selection into MA, and that the rest of the selection can be explained by preference heterogeneity for MA plans. In other words, although preference heterogeneity is a more important determinant of advantageous selection into MA, advertising by MA insurers reinforces the direction of selection.

	Baseline	Ban on Ad
Health Status of Enrollees in traditional Medicare (\$)	462.9	460.3
Health Status of Enrollees in MA (\$)	402.3	408.5
Difference between traditional Medicare and MA(\$)	60.6	51.8 (-15%)
Over-payments per MA enrollee (\$)	136.9	130.8 (-4%)
Over-payments per a random beneficiary (\$)	104.3	104.3
Additional over-payments per MA enrollee (\$)	32.6	26.5 (-19%)

Table 3.17: Health Compositions in traditional Medicare vs MA (Ban on Advertising)
Note: The numbers in the third row is the difference between the first and second row, and an additional over-payment is the difference between an over-payment per MA enrollee and over-payment per a random beneficiary.

Because advertising reinforces advantageous selection into MA, it leads to over-paying of capitation payments to MA plans. In the data, MA plans are over-paid even for a random Medicare beneficiary, as reported in Table 3.16. A reason for this over-payment is that capitation payments were higher than average traditional Medicare costs during this period. Moreover, capitation payments are calculated based on Medicare costs of beneficiaries in traditional Medicare, who are less healthy than MA enrollees. Because over-payments exist even with a random selection into MA, we calculate additional over-payments caused by a non-random selection into MA and compare how these additional over-payments change with the ban on advertising. We find that advertising accounts for 19% of additional over-payments per MA enrollee, and that the rest of the average additional over-payment is attributable to preference heterogeneity between healthy and unhealthy individuals for MA.

3.7.2 Risk Adjustment

In this counterfactual analysis, we simulate the effects of a perfectly risk-adjusted capitation payment on the MA market equilibrium in order to investigate how incentives for risk selection affect an insurer’s choices. A perfectly risk-adjusted capitation payment is a capitation payment that perfectly accounts for variation in health expenditures across individuals having different health statuses. In this counterfactual analysis, let $\widetilde{cap}_{ct}(h)$

denote the new capitation payment in county c in year t that directly depends on an individual's health status h in terms of Medicare claims cost. We assume that:

$$\widetilde{cap}_{ct}(h) = h + const_{ct}. \quad (3.14)$$

That is, the difference between a capitation payment to an MA insurer and an individual's health status is constant for individuals having different health statuses. An important choice we need to make in this counterfactual analysis is the choice of $const_{ct}$ because it determines the overall generosity of a capitation payment. In order to make the results of this counterfactual analysis comparable to the baseline, we choose $const_{ct}$ to be the average of the over-payments per MA enrollee in each county-year in the baseline. That is, noting that $cap_{ct}(z)$ is a capitation payment in the baseline that depends on individual characteristic z ,

$$const_{ct} = E[cap_{ct}(z) - h | d_{ct}(z) = 1].$$

Expectation is taken over individual characteristics z , and $d_{ct}(z)$ is an indicator that equals one if an individual with characteristic z chooses any MA plan in county c in year t in the baseline. This new capitation payment structure changes amounts of over-payments for individuals with different h but keeps the average over-payment unchanged.

We simulate insurers' premiums and advertising quantities in the new environment, and the results are presented in Table 3.18. The risk-adjusted capitation payments have large effects on insurers' choices. The average advertising expenditure decreases by 30.7%, and the average premium increases from \$32.4 to \$51.1. The results are similar for insurers whose advertising expenditures were above the average in the baseline. The average advertising expenditure by these insurers decreases by 27.8%, and the average premium increases from \$32.4 to \$63.5.

The large decrease in advertising expenditure results from a decrease in marginal profits from enrolling healthy individuals. With the perfect risk-adjustment considered

		Baseline	Risk Adjustment
All firms (N = 4864)	Ad expenditure (\$)	78.2K	53.7K (-30.7%)
	Premium (\$)	32.4	51.1
	Share of Beneficiaries	0.243	0.221 (-9%)
	Expected health expenditures (\$)	402.3	406.7
	Over-payment per enrollee (\$)	140.2	140.2
Firms with Above-average ad (N = 881)	Ad expenditure (\$)	392.4K	282.4K (-27.8%)
	Premium (\$)	32.4	63.5
	Share of Beneficiaries	0.101	0.091 (-8.8%)
	Expected health expenditures (\$)	407.4	411.9
	Over-payment per enrollee (\$)	150.4	145.4

Table 3.18: Risk Adjustment

in this counterfactual analysis, capitation payments decrease for healthy individuals and increase for unhealthy individuals. Because advertising has a greater effect on healthier individuals, the perfect risk-adjustment will result in a decrease in marginal profit from an additional unit of advertising, which will lead to a decrease in advertising spending. This finding highlights the importance of risk selection in driving incentives for MA insurers to advertise.

The decrease in revenues from healthy individuals due to the perfect risk-adjustment also leads to increases in premiums. Given our finding that healthy individuals prefer MA more than less healthy individuals even without advertising, MA enrollees are relatively healthy even with the lower advertising expenditure caused by the perfect risk-adjustment. Because the risk-adjustment reduces revenues from enrolling healthy individuals for MA insurers, the insurers increase premiums to compensate for the decrease in revenues. Another factor that contributes to the increase in premiums is that unhealthier individuals are less sensitive to premiums. Because unhealthy individuals now become more profitable to insure, insurers will have incentives to increase premiums to exploit their relative insensitivity to premiums.

Due to the decrease in advertising and the increase in premiums, overall MA enrollment decreases by about 9%, and MA enrollees become less healthy on average. The average over-payment per MA enrollee does not change very much because the constant term in

(3.14) was chosen to be equal to the average over-payment in the baseline. However, the average over-payment for insurers having above-average advertising in the baseline decreases because their enrollees are healthier than those of other insurers because they still advertise more than other insurers even in the new environment. The increase in premiums results in a reduction in consumers' welfare, which is presented in Table 3.19. Because the magnitude of the increase in premiums is large, consumers' welfare decreases, regardless of individual health status and whether we include the effects of advertising on utility. The changes in insurers' choices due to the risk-adjustment also leads to a less healthy pool of MA enrollees, which results in a decrease in the difference in health status between enrollees in MA and enrollees in traditional Medicare by 11%. Lastly, the average additional over-payment in the new environment does not change because the constant term in (3.14) was chosen to match the average over-payment in the baseline.

		Baseline	Risk Adjustment
Case 1	$rh_i < 25\%$	\$112.8	\$96.8
	$rh_i > 75\%$	\$135.7	\$128.3
	Overall	\$116.6	\$110.9
Case 2	$rh_i < 25\%$	\$101.9	\$99.1
	$rh_i > 75\%$	\$131.6	\$129.5
	Overall	\$109.5	\$107.3

Table 3.19: Consumer's Surplus with Risk Adjustment

Note: In case 1, the calculation of consumer surplus included the effects of advertising on utility. In case 2, however, we exclude the effects of advertising on utility in the calculation of consumer surplus. That $rh_i < 25\%$ refers to the group of individuals whose relative health status rh_i is below the 25th percentile in the distribution of relative health status. That is, this group is the healthiest. That $rh_i > 75\%$ refers to the group of individuals whose relative health status rh_i is above the 75th percentile in the distribution of relative health status. That is, this group is the most unhealthy.

3.8 Conclusion

This is the first study to quantify the effects of advertising on risk selection and competition in health insurance markets and to investigate how incentives for risk selection

	Baseline	Risk Adjustment
Health Status of Enrollees in traditional Medicare (\$)	462.9	459.8
Health Status of Enrollees in MA (\$)	402.3	406.1
Differences between traditional Medicare and MA(\$)	60.3	53.7 (-11%)
Over-payments per MA enrollee (\$)	136.8	138.5
Over-payments per a random beneficiary (\$)	104.3	138.5
Additional over-payments per MA enrollee (\$)	32.6	0

Table 3.20: Health Risk Compositions in traditional Medicare vs MA (Risk Adjustment)
Note: The numbers in the third row is the difference between the first and second row, and an additional over-payment is the difference between an over-payment per MA enrollee and over-payment per a random beneficiary.

affect insurers' advertising expenditures. We document strong incentives for risk selection by insurance companies in MA due to an imperfect risk adjustment of capitation payments, and we also show how the incentives for risk selection vary over different regions. We present descriptive evidence that MA insurers advertise more in regions where risk selection is more profitable. For the main analysis, we develop and structurally estimate an equilibrium model that incorporates strategic advertising by insurers. The estimates suggest that advertising increases overall demand with a larger effect on healthier individuals. With a counterfactual analysis where advertising is banned, we find that advertising accounts for 15% of the selection of healthier individuals into MA. By reinforcing the selection of healthier individuals into MA, advertising reduces the costs of MA insurers and keeps premiums from increasing although advertising increases demand for MA insurers. By implementing a perfectly risk-adjusted capitation payment, moreover, we also find that incentives for risk selection can account for about 30% of advertising spending in the data, which highlights an important link between advertising and risk selection.

Appendix A

Appendix to Chapter 1

A.1 Numerical Algorithm to Solve the Equilibrium of the Benchmark Model

In this appendix, we describe the numerical algorithm used to solve the equilibrium of the benchmark model in Section 1.4.

1. **(Discretization of Productivity)**. Discretize the support of productivity $[\underline{p}, \bar{p}]$ into N finite points $\{p_1, \dots, p_N\}$, and calculate the probability weight of each $p \in \{p_1, \dots, p_N\}$ using $\Gamma(p)$.¹
2. **(Initialization)**. Provide an initial guess of the wage policy function and the health insurance offer probability $(w_0^0(p), w_1^0(p), \Delta^0(p))$ for all $p \in \{p_1, \dots, p_N\}$.
3. **(Iterations)**. At iteration $\iota = 0, 1, \dots$, do the following sequentially, where we index the objects in iteration ι by superscript ι :
 - (a) Given the current guess of the wage policy function and the health insurance offer probability $(w_0^\iota(p), w_1^\iota(p), \Delta^\iota(p))$, construct the offer distribution $F^\iota(w, x)$

¹See Kennan (2006) for a discussion about the discrete approximation of the continuous distributions. In our empirical application, we set $N = 200$; and set $p_1 = 0.1$ and $p_N = 6$. We also experimented with $N = 800$. The results are similar.

by using (1.34) and (1.33).

- (b) By using $F^t(w, x)$, numerically solve worker's strategy $\left(\underline{w}_h^x, \underline{s}_h^x(\cdot, \cdot), \underline{q}_h^x\right)$ and calculate U_h and $V_h(w_x^t(p), x)$ for $h \in \{U, H\}$, $x \in \{0, 1\}$, and p on support $[\underline{p}, \bar{p}]$. Moreover, calculate $V_h(w, x)$ for $w \in \mathcal{W}$, where \mathcal{W} is the discrete set of potential wage choices.²
- (c) Calculate unemployment u_h^t and employment distribution $e_h^{x,t} G_h^{x,t}(w_x^t(p))$ for all $p \in \{p_1, \dots, p_N\}$ by solving functional fixed point equations (1.14), (1.17), (1.21) and (1.25);³
- (d) Calculate $n_h^t(w^t(p), x)$ and $n^t(w^t(p), x)$ for all p by respectively using (1.26) and (1.27). Moreover, calculate $n^t(w, x)$ for $w \in \mathcal{W}$;
- (e) Update the firm's optimal policy $(w_0^{*t}(p), w_1^{*t}(p), \Delta^{*t}(p))$ for all p using (1.29) and (1.30);⁴
- (f) Given $(w_0^{*t}(p), w_1^{*t}(p))$, calculate $\pi_0^{*t}(p)$ and $\pi_1^{*t}(p)$ from (1.29) and (1.30) and obtain $\Delta^{*t}(p)$ by using (1.31).

4. (Convergence Criterion)

- (a) If $(w_0^{*t}(p), w_1^{*t}(p), \Delta^{*t}(p))$ satisfies $d(w_0^{*t}(p), w_0^t(p)) < \epsilon_{tol}$, $d(w_1^{*t}(p), w_1^t(p)) < \epsilon_{tol}$ and $d(\Delta^{*t}(p), \Delta^t(p)) < \epsilon_{tol}$ where ϵ_{tol} is a pre-specified tolerance level of convergence and $d(\cdot, \cdot)$ is a distance metric, then firm's optimal policy converges and we have an equilibrium.⁵

²The number discrete values of potential wage choices is set to 400 in our empirical application.

³Although we do not have a proof that the unique fixed point exists, we always find the unique solution regardless of initial guess of u_h and $e_h^x G_h^x(w(p))$.

⁴See Proposition 3 below for a numerical shortcut in the updating of $w_0^{t+1}(p)$ and $w_1^{t+1}(p)$.

⁵In solving for the equilibrium we set ϵ_{tol} to 1.0e-6.

(b) Otherwise, update $(w_0^{\iota+1}(p), w_1^{\iota+1}(p), \Delta^{\iota+1}(p))$ as follows:

$$\begin{aligned} w_0^{\iota+1}(p) &= \omega w_0^\iota(p) + (1 - \omega)w_0^{*\iota}(p), \\ w_1^{\iota+1}(p) &= \omega w_1^\iota(p) + (1 - \omega)w_1^{*\iota}(p), \\ \Delta^{\iota+1}(p) &= \omega \Delta^\iota(p) + (1 - \omega)\Delta^{*\iota}(p), \end{aligned}$$

for $\omega \in (0, 1)$ and continue Step 2 at iteration $\iota' = \iota + 1$.

Given our convergence criterion, it is clear that the convergence point of our numerical algorithm will correspond to steady state equilibrium of our model.

Proposition 3. *For each p , optimal wage policy must satisfy*

$$w_1(p) = \frac{\left[\begin{array}{l} (p - m_H^1) n_H(w_1(p), 1) + (pd - m_U^1) n_U(w_1(p), 1) \\ - \int_{p_1^*}^p [n_H(w_1(\tilde{p}), 1) + dn_U(w_1(\tilde{p}), 1)] d\tilde{p} - \pi_1(p_1^*) \end{array} \right]}{n_H(w_1(p), 1) + n_U(w_1(p), 1)} \quad (\text{A.1})$$

$$w_0(p) = \frac{\left[\begin{array}{l} pn_H(w_0(p), 0) + pdn_U(w_0(p), 0) \\ - \int_{p_0^*}^p [n_H(w_0(\tilde{p}), 0) + dn_U(w_0(\tilde{p}), 0)] d\tilde{p} - \pi_0(p_0^*) \end{array} \right]}{n_H(w_0(p), 0) + n_U(w_0(p), 0)}. \quad (\text{A.2})$$

where $p_x^* = \inf\{p \in [\underline{p}, \bar{p}] : n_H(w_x(p), x) > 0 \text{ and } n_U(w_x(p), x) > 0\}$ and

$$\begin{aligned} \pi_1(p_1^*) &= \left[\begin{array}{l} [p_1^* - w_1^*(p_1^*) - m_H^1] n_H(w_1^t(p_1^*), 1) \\ + [p_1^*d - w_1(p_1^*) - m_U^1] n_U(w_1^t(p_1^*), 1) \end{array} \right] \\ \pi_0(p_0^*) &= \left[\begin{array}{l} [p_0^* - w_0(p_0^*)] n_H(w_0(p_0^*), 0) \\ + [p_0^*d - w_0(p_0^*)] n_U(w_0(p_0^*), 0) \end{array} \right]. \end{aligned}$$

Proof. To prove Proposition 3, we first establish a lemma that:

Lemma 4. *For any distribution $F(w, x)$, $w_x(p)$ is increasing in p for each x .*

Proof. The proof is based on revealed preference argument. Choose any p and p' in $[\underline{p}, \bar{p}]$

such that $p > p'$ and fix $x \in \{0, 1\}$. Notice that

$$\begin{aligned}
\pi_x(p) &= \begin{bmatrix} (p - w_x(p) - xm_H^x) n_H(w_x(p), x) \\ + (pd - w_x(p) - xm_U^x) n_U(w_x(p), x) \end{bmatrix} - xC \\
&\geq \begin{bmatrix} (p - w_x(p') - xm_H^x) n_H(w_x(p'), x) \\ + (pd - w_x(p') - xm_U^x) n_U(w_x(p'), x) \end{bmatrix} - xC \\
&\geq \begin{bmatrix} (p' - w_x(p') - xm_H^x) n_H(w_x(p'), x) \\ + (p'd - w_x(p') - xm_U^x) n_U(w_x(p'), x) \end{bmatrix} - xC \\
&= \pi_x(p', k) \\
&\geq \begin{bmatrix} (p' - w_x(p) - xm_H^x) n_H(w_x(p), x) \\ + (p'd - w_x(p) - xm_U^x) n_U(w_x(p), x) \end{bmatrix} - xC,
\end{aligned}$$

where the second line comes from the fact that $w_x(p)$ is the optimal wage policy for a firm with productivity p and third line is implied by the assumption that $p > p'$. The fifth line is implied by the fact that $w_x(p)$ is the optimal policy for a firm with productivity p , not p' . Therefore, we have

$$(p - p') [n_H(w_x(p), x) + n_U(w_x(p), x)] \geq (p - p') [n_H(w_x(p'), x) + n_U(w_x(p'), x)].$$

Since $n_h(w, x)$ is increasing in w , this inequality holds if and only if $w_x(p) \geq w_x(p')$. \square

Now we complete the proof of Proposition 3. Define $p_x^* = \inf\{p \in [\underline{p}, \bar{p}] : n_H(w_x(p), x) > 0 \text{ and } n_U(w_x(p), x) > 0\}$ for $x = 0, 1$. From Lemma 4, $w_x(p)$ is increasing in p ; also $n_h(w, x)$ is increasing in w ; thus we have $n_H(w_x(p), x) > 0$ and $n_U(w_x(p), x) > 0$ for $p > p_x^*$. Define

$$\tilde{\pi}(w_x, x) \equiv \max_x \begin{bmatrix} (p - w_x - xm_H^x) n_H(w_x, x) \\ + (pd - w_x - xm_U^x) n_U(w_x, x) \end{bmatrix}.$$

Notice that the solution $w_x(p)$ is equal to the one defined in (1.29) and (1.30) and be independent of C . By applying envelope condition, we have

$$\tilde{\pi}'_x(p) = n_H(w_x(p), x) + dn_U(w_x(p), x)$$

for $p > p_x^*$. By taking integral over $[p_x^*, p]$, we then obtain

$$\tilde{\pi}_x(p) = \int_{p_x^*}^p n [{}_H(w_x(\tilde{p}), x) + dn_U(w_x(\tilde{p}), x)] d\tilde{p} + \tilde{\pi}_x(p_x^*).$$

By equating it with (1.29) and (1.30), we obtain (A.1) and (A.2). This is a form of wage policy which we utilize in our numerical algorithm. \square

A.2 Omitted Formula in Section 1.6.1

In this appendix, we provide the formula omitted in Section 1.6.1:

$$\begin{aligned} \Pr(h_{t+3} = U | x_t, x_{t+1}, x_{t+2}, h_t = H) &= \pi_{HH}^{x_t} \pi_{HH}^{x_{t+1}} (1 - \pi_{HH}^{x_{t+2}}) \\ &+ \pi_{HH}^{x_t} (1 - \pi_{HH}^{x_{t+1}}) \pi_{UU}^{x_{t+2}} \\ &+ (1 - \pi_{HH}^{x_t}) (1 - \pi_{UU}^{x_{t+1}}) (1 - \pi_{HH}^{x_{t+2}}) \\ &+ (1 - \pi_{HH}^{x_t}) \pi_{UU}^{x_{t+1}} \pi_{UU}^{x_{t+2}}; \end{aligned}$$

$$\begin{aligned} \Pr(h_{t+3} = H | x_t, x_{t+1}, x_{t+2}, h_t = U) &= (1 - \pi_{UU}^{x_t}) \pi_{HH}^{x_{t+1}} \pi_{HH}^{x_{t+2}} \\ &+ (1 - \pi_{UU}^{x_t}) (1 - \pi_{HH}^{x_{t+1}}) (1 - \pi_{UU}^{x_{t+2}}) \\ &+ \pi_{UU}^{x_t} (1 - \pi_{UU}^{x_{t+1}}) \pi_{HH}^{x_{t+2}} + \pi_{UU}^{x_t} \pi_{UU}^{x_{t+1}} (1 - \pi_{UU}^{x_{t+2}}); \end{aligned}$$

$$\begin{aligned} \Pr(h_{t+3} = U | x_t, x_{t+1}, x_{t+2}, h_t = U) &= (1 - \pi_{UU}^{x_t}) \pi_{HH}^{x_{t+1}} (1 - \pi_{HH}^{x_{t+2}}) \\ &+ (1 - \pi_{UU}^{x_t}) (1 - \pi_{HH}^{x_{t+1}}) \pi_{UU}^{x_{t+2}} \\ &+ \pi_{UU}^{x_t} (1 - \pi_{UU}^{x_{t+1}}) (1 - \pi_{HH}^{x_{t+2}}) + \pi_{UU}^{x_t} \pi_{UU}^{x_{t+1}} \pi_{UU}^{x_{t+2}}. \end{aligned}$$

A.3 Derivation of Likelihood Function

We will first derive the likelihood contribution of the labor market transitions of *unemployed workers*. Consider an unemployed worker at period 1 with health status is h_1 , who experiences an unemployment spells l and in period $l + 1$ transitions to a job (\tilde{w}, x) . Moreover, denote $h^l \equiv (h_1, h_2, \dots, h_l)$ be the realized history of health status between $j = 1$ to l . In our data scenario, we assume that the initial h_1 but we do not observe h_2, \dots, h_l . The likelihood contribution of observing such a transition is:

$$\frac{u_{h_1}}{M} \times \sum_{h^l \in \{H, L\}^l} \Pr(s_u(h^l)) \times \sum_{h_{l+1} \in \{H, L\}} \Pr(h_{l+1}|h_l) \times [\lambda_u f(\tilde{w}, 1)]^{\mathbf{1}(x=1)} \times [\lambda_u f(\tilde{w}, 0)]^{\mathbf{1}(x=0)} \quad (\text{A.3})$$

where

$$\Pr(s(h^l)) = \prod_{j=2}^l \left\{ \Pr(h_j|h_{j-1}) \times \left[(1 - \lambda_u) + \lambda_u \left(F\left(\frac{w_{h_j}^1}{w_{h_j}^1}, 1\right) + F\left(\frac{w_{h_j}^0}{w_{h_j}^0}, 0\right) \right) \right] \right\}$$

and $\mathbf{1}(x = 1)$ is an indicator function such that it takes one if we observe a transition to employment with $(\tilde{w}, 1)$ at period $l + 1$, and similarly $\mathbf{1}(x = 0)$ is an indicator function such that it takes one if we observe a transition to employment with $(\tilde{w}, 0)$ at period $l + 1$. To understand (A.3), note that the first term in (A.3), u_h/M , reflects the assumption that the initial condition of individuals is drawn from steady state worker distribution because u_h/M the probability that an unemployed worker with health status h is sampled. The second term in (A.3) is the probability that individual experiences l periods of unemployment with health status transitions (h_2, \dots, h_l) during the process; note that the term $\left[(1 - \lambda_u) + \lambda_u \left(F\left(\frac{w_{h_j}^1}{w_{h_j}^1}, 1\right) + F\left(\frac{w_{h_j}^0}{w_{h_j}^0}, 0\right) \right) \right]$ is the probability that the individual does not receive an offer or receives an offer that is lower than the relevant reservation wages $\frac{w_{h_j}^1}{w_{h_j}^1}$ or $\frac{w_{h_j}^0}{w_{h_j}^0}$. The third to fifth terms in (A.3) are the probability that his health transitions from h_l to h_{l+1} in period $l + 1$ and receive an acceptable offer (\tilde{w}, x) from the relevant density function $f(\tilde{w}, x)$.

We can similarly derive the likelihood contribution of the job dynamics of *employed*

workers. Consider an employed worker in period 1 with health status h_1 working on a job with compensation package (w, x) . Suppose that the worker experiences a job status changes in period $l + 1$, and denote h^l be the realized history of health status between $j = 1$ to l (h_1, h_2, \dots, h_l). We again assume that we observe h_1 but do not for $h_2 \dots, h_l$. For an employed worker, there are four possible changes in job status:

- [Event “*Job Loss*”] the individual experienced a job loss at period $l + 1$;
- [Event “*Switch 1*”] the individual transitioned to a job (\tilde{w}, x') such that $x' = x$ and the accepted wage is $\tilde{w} > w$;
- [Event “*Switch 2*”] the individual transitioned to a job (\tilde{w}, x') such that $x' = x$ and the accepted wage is $\tilde{w} < w$;
- [Event “*Switch 3*”] the individual transitioned to a job (\tilde{w}, x') such that $x' \neq x$ and the accepted wage is \tilde{w} .

The likelihood contribution is given by:

$$\frac{e_h^x g_h^x(w)}{M} \times \sum_{h^l \in \{H, L\}^l} \Pr(s_e(h^l)) \quad (\text{A.4a})$$

$$\times \sum_{h_{l+1} \in \{H, L\}} \Pr(h_{l+1}|h_l) \quad (\text{A.4b})$$

$$\times \begin{cases} \delta \left[(1 - \lambda_e) + \lambda_e \sum_{\tilde{x}} F(\underline{w}_{h_{l+1}}^{\tilde{x}}, \tilde{x}) \right] & \text{if Event is “Job Loss”} \\ \lambda_e f(\tilde{w}, x) & \text{if Event is “Switch 1”} \\ \delta \lambda_e f(\tilde{w}, x) & \text{if Event is “Switch 2”} \\ (1 - \delta) \lambda_e f(\tilde{w}, x') + \delta \lambda_e f(\tilde{w}, x') & \text{if Event is “Switch 3”,} \end{cases}$$

where

$$\Pr(s_e(h^l)) = \prod_{j=2}^l \left\{ \Pr(h_{j+1}|h_{j-1})(1 - \delta) \left[\begin{array}{c} (1 - \lambda_e) \\ + \lambda_e \left(F(w, x) + F(\underline{s}_{h_j}^{x'}(w, x), x') \right) \end{array} \right] \right\}$$

and $x' \neq x$. To understand (A.4a), note that similar to that in (1.37a), the first term in (A.4a), $e_h^x g_h^x(w)/M$, is the probability of sampling an employed worker with health status h working on a job (w, x) ; the second term in (A.4a) is the probability that individual stays with the job (w, x) for l periods of unemployment with health status transitions (h_2, \dots, h_l) during the process. The remaining two terms in (A.4b) express the likelihood of observing health transition from h_l to h_{l+1} in period $l + 1$ and one of the four job status change events. For example, the event “Job Loss” is observed in period $l + 1$ with probability $\delta \left[(1 - \lambda_e) + \lambda_e \sum_{\tilde{x}} F(\underline{w}_{h_{l+1}}^{\tilde{x}}, \tilde{x}) \right]$ because in order for a job loss to occur, the worker has to experience an exogenous shock that destroys the current match (which occurs with probability δ), and then does not get matched to another accepted job (which occurs with probability $(1 - \lambda_e) + \lambda_e \sum_{\tilde{x}} F(\underline{w}_{h_{l+1}}^{\tilde{x}}, \tilde{x})$). To understand the probability of event “Switch 2”, we note that in order for a worker to switch to a job (\tilde{w}, x') with $x' = x$ but $\tilde{w} < w$, the worker must have experienced a job separation (which occurs with probability δ), but is then lucky enough to find an acceptable job immediately, which happens with probability $\lambda_e f(\tilde{w}, x)$. The probability of the other job switch events are derived similarly.

A.4 Estimation Procedure

The following is the procedure we use to implement the GMM estimator in Section 1.6:

1. **(Initialization)** Initialize a guess of the parameter values θ ;
2. **(Solving for Equilibrium Offer Distribution)** Given the guess, solve equilibrium numerically using the algorithm we provided in Section A.1. Obtain the offer distribution $\hat{F}(w, x)$ from the equilibrium;
3. **(Calculating the Worker-Side Moments)** Use $\hat{F}(w, x)$ in place of $F(w, x)$ in the likelihood functions of the observed worker-side data based on (1.37) and (1.38),

and obtain the numerical derivative of likelihood with respect to parameters $\theta_1 \equiv (\lambda_u, \lambda_e, \delta, \gamma, \mu_h, b)$ and use them as a subset of the moments in (1.35);⁶

4. **(Calculating the Employer-Side Moments)** Use $\hat{F}(w, x)$ and other equilibrium elements obtained in (2) to calculate the employer-side moments listed in Section 1.6.4;
5. **(Iteration)** Evaluate the GMM objective (1.36) and iterate until it converges.

A.5 Equilibrium of the Counterfactual Economy

A.5.1 Steady State Equilibrium for the Post-Reform Economy

A steady state equilibrium for the post-reform economy is a list

$$\left\langle \left(\underline{w}_h^x, \underline{s}_h^x(\cdot, \cdot), \underline{q}_h^x, x_h^*, x_h^*(\cdot) \right), \left(u_h^x, e_h^x, G_h^x(w) \right), \left(w_x(p), \Delta(p), F(w, x) \right), R^{EX} \right\rangle$$

such that the following conditions hold:

- **(Worker Optimization)** Given $F(w, x)$ and R^{EX} ,
 - \underline{w}_h^x solves the unemployed workers' job acceptance decision problem for each $(h, x) \in \{U, H\} \times \{0, 1\}$;
 - $\underline{s}_h^x(\cdot, \cdot)$ solves the job-to-job switching problem for currently employed workers for each $(h, x) \in \{U, H\} \times \{0, 1\}$
 - \underline{q}_h^x describes the optimal strategy for currently employed workers regarding whether to quit into unemployment for each $(h, x) \in \{U, H\} \times \{0, 1\}$;
 - x_h^* and $x_h^*(\cdot)$ respectively solve (1.41) and (1.47) for $h \in \{H, U\}$.

⁶Note from (1.37) and (1.38), the likelihood function of the worker labor market transitions depends only on θ_1 , given $F(w, x)$.

- **(Steady State Worker Distribution)** Given workers' optimizing behavior described by $(\underline{w}_h^x, \underline{s}_h^x(\cdot, \cdot), \underline{q}_h^x, x_h^*, x_h^*(\cdot))$ and $F(w, x)$ and $R^{EX}, (u_h^x, e_h^x, G_h^x(w))$ satisfy the steady state conditions for worker distribution (details are provided in Section A.5.3).
- **(Firm Optimization)** Given $F(w, x), R^{EX}$ and the steady state employee sizes implied by $(u_h^x, e_h^x, G_h^x(w))$, a firm with productivity p chooses to offer health insurance, i.e., $x = 1$, with probability $\Delta(p)$ and chooses not to offer health insurance with probability $1 - \Delta(p)$, where $\Delta(p)$ is given by (1.31). Moreover, conditional on insurance choice x , the firm offers a wage $w_x(p)$ that solves (1.48) and (1.49) respectively for $x = 0$ and 1.
- **(Equilibrium Consistency)** The postulated distributions of offered compensation packages are consistent with the firms' optimizing behavior $(w_x(p), \Delta(p))$. Specifically, $F(w, x)$ must satisfy:

$$\begin{aligned}
 F(w, 1) &= \int_0^\infty \mathbf{1}(w_1(p) < w) \Delta(p) d\Gamma(p), \\
 F(w, 0) &= \int_0^\infty \mathbf{1}(w_0(p) < w) [1 - \Delta(p)] d\Gamma(p).
 \end{aligned}$$

- **(Equilibrium Condition in Insurance Exchange)** The premium in exchange is determined by (1.50).

A.5.2 Numerical Algorithm for the Counterfactual Policy Experiments

We use numerical methods to solve the equilibrium. The basic iteration procedure to solve the equilibrium for the counterfactual environment remains the same as that described in Section A.1, but there are two important changes. First, we need to find the fixed point of not only $(w_0(p), w_1(p), \Delta(p))$ but also R^{EX} , the premium in insurance exchange. Second,

because the penalty associated with employer mandate depends on size of the firm, for example, the threshold under the ACA for firms to pay penalty if they do not offer health insurance is 50; as a result we need to modify the algorithm to allow for a potential mass point of employer size just to the left of 50 (say, 49 workers) when we derive optimal wage policy $w_0(p)$.

Finally, the establishment of the health insurance exchange with community rating may result in multiple equilibria under some counterfactual policy experiments. In our numerical simulations, we sometimes find multiple equilibria and we will discuss their implication.

Because of the size dependent employer mandate, there may exist a mass point of wage offer w_{49} under which firm size is equal to 49 if $x = 0$:

$$n_H(w_{49}, 0) + n_U(w_{49}, 0) = 49.$$

Note that w_{49} is endogenously determined in equilibrium. We now provide our numerical algorithm:

1. Guess $(w_0(p), w_1(p), \Delta(p))$ for $p = p_1, \dots, p_N$ and R^{EX} .
2. Solve value function and employment distribution as before. Notice that there may exist an interval of productivity $[p^*, p^{**}]$ such that there is a mass point of wage offer $w_{49} = w_0(p) = w_0(p')$, for $p, p' \in [p^*, p^{**}]$.
3. Once we solve employment distribution, find w_{49} by linear interpolation of firm size distribution.
4. Update $w_1(p)$ as in the benchmark case. $w_0(p)$ is updated in the following way:
 - (a) From p_1 , solve $w_0(p)$ by maximizing $(p-w)(n_H(w, 0) + n_U(w, 0))$. If $n_H(w, 0) > 0$ and $n_U(w, 0) > 0$, then solve it by using the equation implied from the envelope condition, as before. Repeat this for $p_2, p_3 \dots$ as long as $w_0(p) < w_{49}$.

- (b) If we find p^* such that $w(p^*) > w_{49}$ where $w(p^*) = \arg \max(p^* - w)(n_H(w, 0) + n_U(w, 0))$, then from p^* , solve firm's problem by

$$\max \{\Pi_{49}(p), \Pi_{pe}(p)\}$$

where

$$\Pi_{49}(p) = (p - w_{49})(n_H(w_{49}, 0) + n_U(w_{49}, 0));$$

$$\Pi_{pe}(p) = \max_w (p - w)(n_H(w, 0) + n_U(w, 0)) - P_E(n_H(w, 0) + n_U(w_0, 0)).$$

- (c) If we find p^{**} such that $\Pi_{49}(p^{**}) < \pi_{pe}(p^{**})$, then for the remaining $p > p^{**}$, evaluate $w_0(p)$ by

$$w_0(p) = \frac{\left[\begin{array}{c} pn_H(w_0(p), 0) + pdn_U(w_0(p), 0) - \\ \int_{p^{**}}^p [n_H(w_0(\tilde{p}), 0) + dn_U(w_0(\tilde{p}), 0)] d\tilde{p} - \pi_0(p^{**}) \end{array} \right]}{n_H(w_0(p), 0) + n_U(w_0(p), 0) - P_E(n_H(w_0(p), 0) + n_U(w_0(p), 0))},$$

the derivation of which basically follows Proposition 3 in benchmark case but reflects the existence of employer mandate.

5. Update R^{EX} using (46). Continue it until it converges.

In Step 4 (c) we utilize the following lemma.

Lemma 5. *Suppose that there exists a p^{**} such that $\Pi_{49}(p^{**}) < \Pi_{pe}(p^{**})$. Then, for any $p > p^{**}$, $\Pi_{49}(p) < \Pi_{pe}(p)$.*

Proof. Proof is by contradiction. Suppose that there exists $p' > p^{**}$ such that $\Pi_{49}(p') \geq \Pi_{pe}(p')$. Then, notice that

$$\Pi_{49}(p') \geq \Pi_{pe}(p') \geq \Pi_{pe}(p^{**}) > \Pi_{49}(p^{**})$$

where the second inequality is from the revealed preference argument. Therefore, we must have:

$$\Pi_{49}(p') - \Pi_{49}(p^{**}) > \Pi_{pe}(p') - \Pi_{pe}(p^{**}). \quad (\text{A.5})$$

The left hand side of (A.5) is:

$$\Pi_{49}(p') - \Pi_{49}(p^{**}) = (p - p^{**}) (n_H(w_{49}, 0) + n_U(w_{49}, 0)).$$

The right hand side of (A.5) is:

$$\begin{aligned} & \Pi_{pe}(p') - \Pi_{pe}(p^{**}) \\ = & \{(p' - w(p')) [n_H(w(p'), 0) + n_U(w(p'), 0)] - P_E(n(w(p')))\} \\ & - \{(p^{**} - w(p^{**})) [n_H(w(p^{**}), 0) + n_U(w(p^{**}), 0)] - P_E(n(w(p^{**})))\} \\ \geq & (p' - w(p^{**})) [n_H(w(p^{**}), 0) + n_U(w(p^{**}), 0)] - P_E(n(w(p^{**}))) \\ & - \{(p^{**} - w(p^{**})) [n_H(w(p^{**}), 0) + n_U(w(p^{**}), 0)] - P_E(n(w(p^{**})))\} \\ = & (p - p^{**}) [n_H(w(p^{**}), 0) + n_U(w(p^{**}), 0)]. \end{aligned}$$

Since $w(p^{**}) > w_{49}$, we must have $\Pi_{pe}(p') - \Pi_{pe}(p^{**}) > \Pi_{49}(p') - \Pi_{49}(p^{**})$. A contradiction. □

A.5.3 Steady State Distribution of Employment in Counterfactual Experiments

We provide a derivation of steady state employment distribution used in our counterfactual policy experiments. Note that worker's insurance status can be $x \in \{0, 2\}$ for the unemployed and $x \in \{0, 1, 2\}$ for the employed, because of the insurance exchange.

First, we define the resource constraints of the economy:

$$\sum_{h \in \{U, H\}} (u_h^0 + u_h^2 + e_h^0 + e_h^1 + e_h^2) = M. \quad (\text{A.6})$$

Next, we provide the determination of the unemployment. The inflow into unemployment with health status h and insurance status x is given by: if $x = x_h^*$,

$$\begin{aligned} [u_h^x]^+ &\equiv (1 - \rho) [\delta(1 - \lambda_e) + \delta\lambda_e(F(\underline{w}_h^1, 1) + F(\underline{w}_h^0, 0))] \\ &\quad \times [e_h^0\pi_{hh}^0 + e_h^1\pi_{hh}^1 + e_{h'}^0\pi_{hh'}^0 + e_{h'}^1\pi_{hh'}^1 + e_h^2\pi_{hh}^2 + e_{h'}^2\pi_{hh'}^2] \\ &\quad + (1 - \rho) \sum_{x=0,2} u_{h'}^x \pi_{hh'}^x [1 - \lambda_u (1 - F(\underline{w}_h^1, 1) - F(\underline{w}_h^0, 0))] + M\rho\mu_h \\ &\quad + (1 - \rho)(1 - \delta)e_{h'}^1 G_{h'}^1(\underline{w}_h^1) \pi_{hh'}^1 \\ &\quad \times [1 - \lambda_e (1 - F(\underline{w}_h^1, 1) - F(\underline{w}_h^0, 0))] \\ &\quad + (1 - \rho)(1 - \delta) (e_{h'}^0 G_{h'}^0(\underline{w}_h^0) \pi_{hh'}^0 + e_{h'}^2 G_{h'}^2(\underline{w}_h^0) \pi_{hh'}^2) \\ &\quad \times [1 - \lambda_e (1 - F(\underline{w}_h^1, 1) - F(\underline{w}_h^0, 0))] ; \end{aligned}$$

otherwise, $[u_h^x]^+ = 0$.

The outflow from unemployment with health status h and insurance status x is described as follows. If $x = x_h^*$ where x_h^* is defined in (1.41),

$$[u_h^x]^- \equiv u_h^x \{ \rho + (1 - \rho) [\pi_{h'h}^x + \pi_{hh}^x \lambda_u (1 - F(\underline{w}_h^1, 1) - F(\underline{w}_h^0, 0))] \} ;$$

otherwise, $[u_h^x]^- = 0$. Then, in a steady-state we must have

$$[u_h^x]^+ = [u_h^x]^-, h \in \{U, H\}.$$

Now we provide the steady state equation for workers employed on jobs (w, x) with health status h . Note that the inflow of workers with health status h on jobs $(w, 1)$, denoted by $[e_h^1(w)]^+$, is given by:

$$\begin{aligned}
[e_h^1(w)]^+ &\equiv (1-\rho)f(w,1) \left\{ \begin{array}{l} \lambda_u (u_h^0 \pi_{hh}^0 + u_{h'}^0 \pi_{hh'}^0 + u_h^2 \pi_{hh}^2 + u_{h'}^2 \pi_{hh'}^2) \\ + (1-\delta)\lambda_e \left[\begin{array}{l} \sum_{x'=0,2} e_h^{x'} G_h^{x'}(s_h^0(w,1)) \pi_{hh}^{x'} \\ + \sum_{x'=0,2} e_{h'}^{x'} G_{h'}^{x'}(s_{h'}^0(w,1)) \pi_{hh'}^{x'} \\ + \pi_{hh}^1 e_h^1 G_h^1(w) + \pi_{hh'}^1 e_{h'}^1 G_{h'}^1(w) \end{array} \right] \\ + \delta\lambda_e \left(\begin{array}{l} e_h^0 \pi_{hh}^0 + e_{h'}^0 \pi_{hh'}^0 + e_h^1 \pi_{hh}^1 \\ + e_{h'}^1 \pi_{hh'}^1 + e_h^2 \pi_{hh}^2 + e_{h'}^2 \pi_{hh'}^2 \end{array} \right) \end{array} \right\} \\
&+ (1-\rho)(1-\delta)e_{h'}^1 g_{h'}^1(w) \pi_{hh'}^1 \\
&\times \left[1 - \lambda_e \left(1 - \tilde{F}_h(w,1) \right) \right],
\end{aligned}$$

where $h' \neq h$ and $\tilde{F}_h(w,1)$ is defined by

$$\tilde{F}_h(w,1) = F(w,1) + F(s_h^0(w,1),0).$$

Denote the outflow of workers with health status h from jobs $(w,1)$ by $[e_h^1(w)]^-$, and it is given by

$$[e_h^1(w)]^- \equiv e_h^1 g_h^1(w) \left\{ \rho + (1-\rho) \left[\pi_{h'h}^1 + \pi_{hh}^1 \left(\delta + \lambda_e (1-\delta) \left(1 - \tilde{F}_h(w,1) \right) \right) \right] \right\}. \quad (\text{A.7})$$

The steady state condition requires that

$$[e_h^1(w)]^+ = [e_h^1(w)]^- \text{ for } h \in \{U, H\} \text{ and for all } w \text{ in the support of } F(w,1). \quad (\text{A.8})$$

Similarly, the inflows of workers with health status h into jobs $(w,0)$, denoted by

$[e_h^0(w)]^+$, are given as follows. If $x_h^*(w) = 0$, where $x_h^*(w)$ is defined in (1.47),

$$\begin{aligned}
[e_h^0(w)]^+ &= f(w, 0)(1 - \rho) \\
&\times \left\{ \begin{array}{l} \lambda_u [u_h \pi_{hh}^0 + u_{h'} \pi_{hh'}^0 + u_h^2 \pi_{hh}^2 + u_{h'}^2 \pi_{hh'}^2] \\ + \lambda_e (1 - \delta) \left[\begin{array}{l} e_h^1 G_h^1(\underline{s}_h^1(w, 0)) \pi_{hh}^1 + e_{h'}^1 G_{h'}^1(\underline{s}_{h'}^1(w, 0)) \pi_{hh'}^1 \\ e_h^0 G_h^0(w) \pi_{hh}^0 + e_{h'}^0 G_{h'}^0(w) \pi_{hh'}^0 \\ + e_h^2 G_h^2(w) \pi_{hh}^2 + e_{h'}^2 G_{h'}^2(w) \pi_{hh'}^2 \end{array} \right] \\ + \delta \lambda_e \left(\begin{array}{l} e_h^1 \pi_{hh}^1 + e_{h'}^1 \pi_{hh'}^1 + e_h^0 \pi_{hh}^0 \\ + e_{h'}^0 \pi_{hh'}^0 + e_h^2 \pi_{hh}^2 + e_{h'}^2 \pi_{hh'}^2 \end{array} \right) \end{array} \right\} \\
&+ (1 - \rho)(1 - \delta) \sum_{x=0,2} e_{h'}^x g_{h'}^x(w) \pi_{hh'}^x \\
&\times \left[1 - \lambda_e \left(1 - \tilde{F}_h(w, 0) \right) \right],
\end{aligned}$$

where $h \neq h'$ and $\tilde{F}_h(w, 0)$ is defined by

$$\tilde{F}_h(w, 0) = F(w, 0) + F(\underline{s}_h^1(w, 0), 1);$$

and $[e_h^0(w)]^+ = 0$ otherwise. The outflows of workers with health status h from jobs $(w, 0)$, denoted by $[e_h^0(w)]^-$, are given by:

$$[e_h^0(w)]^- = e_h^0 g_h^0(w) \left\{ \rho + (1 - \rho) \left[\pi_{h'h}^0 + \pi_{hh}^0 (\delta + (1 - \delta) \lambda_e (1 - \tilde{F}_h(w, 0))) \right] \right\}.$$

The steady state condition thus requires that

$$[e_h^0(w)]^+ = [e_h^0(w)]^- \text{ for } h \in \{H, U\} \text{ and for all } w \text{ in the support of } F(w, 0).$$

Similarly, the inflows of workers with health status h into jobs $(w, 2)$ with health insurance through exchange, denoted by $[e_h^2(w)]^+$, are given as follows. If $x_h^*(w) = 2$,

where $x_h^*(w)$ is defined in (1.47),

$$\begin{aligned}
[e_h^2(w)]^+ &= f(w, 0)(1 - \rho) \\
&\times \left\{ \begin{array}{l} \lambda_u [u_h \pi_{hh}^0 + u_{h'} \pi_{hh'}^0 + u_h^2 \pi_{hh}^2 + u_{h'}^2 \pi_{hh'}^2] \\ + \lambda_e (1 - \delta) \left[\begin{array}{l} e_h^1 G_h^1(\underline{s}_h(w, 0)) \pi_{hh}^1 + e_{h'}^1 G_{h'}^1(\underline{s}_{h'}(w, 0)) \pi_{hh'}^1 \\ e_h^0 G_h^0(w) \pi_{hh}^0 + e_{h'}^0 G_{h'}^0(w) \pi_{hh'}^0 \\ + e_h^2 G_h^2(w) \pi_{hh}^2 + e_{h'}^2 G_{h'}^2(w) \pi_{hh'}^2 \end{array} \right] \\ + \delta \lambda_e \left(\begin{array}{l} e_h^1 \pi_{hh}^1 + e_{h'}^1 \pi_{hh'}^1 + e_h^0 \pi_{hh}^0 \\ + e_{h'}^0 \pi_{hh'}^0 + e_h^2 \pi_{hh}^2 + e_{h'}^2 \pi_{hh'}^2 \end{array} \right) \end{array} \right\} \\
&+ (1 - \rho)(1 - \delta) \sum_{x=0,2} e_{h'}^x g_{h'}^x(w) \pi_{hh'}^2 \\
&\times \left[1 - \lambda_e \left(1 - \tilde{F}_h(w, 0) \right) \right],
\end{aligned}$$

where $h \neq h'$ and $\tilde{F}_h(w, 0)$ is defined by

$$\tilde{F}_h(w, 0) = F(w, 0) + F(\underline{s}_h^1(w, 0), 1);$$

and $[e_h^2(w)]^+ = 0$ otherwise. The outflows of workers with health status h from jobs $(w, 0)$, denoted by $[e_h^0(w)]^-$, are given by:

$$[e_h^2(w)]^- = e_h^2 g_h^2(w) \left\{ \rho + (1 - \rho) \left[\pi_{h'h}^2 + \pi_{hh}^2 (\delta + (1 - \delta) \lambda_e (1 - \tilde{F}_h(w, 0))) \right] \right\}.$$

The steady state condition thus requires that

$$[e_h^2(w)]^+ = [e_h^2(w)]^- \text{ for } h \in \{H, U\} \text{ and for all } w \text{ in the support of } F(w, 0).$$

These steady state conditions pin down the distribution of employment and are used to calculate firm size distribution as in the benchmark case.

Appendix B

Appendix to Chapter 2

Appendix

B.1 Numerical Algorithm

I describe the numerical algorithm which is used to solve the equilibrium of pre-ACA model described in section 2.2. I discretize the support of Γ , $[\underline{p}, \bar{p}]$, into finite points. Then, I solve the equilibrium by the following fixed-point algorithm.

1. First, I provide an initial guess of the skill price and the fraction of firms offering ESHI ($\langle \theta_{0,0}^{ed}(p), \theta_{1,0}^{ed}(p) \rangle^{ed}, \Delta_0(p)$) for all p on support $[\underline{p}, \bar{p}]$.
2. At iteration $\iota = 0, 1, \dots$, I do the following sequentially, where I index the objects in iteration ι by superscript ι :
 - (a) Given the current guess of the health insurance costs and the health insurance offer probability ($\langle \theta_{0,\iota}^{ed}(p), \theta_{1,\iota}^{ed}(p) \rangle^{ed}, \Delta_\iota(p)$), I construct an offer distribution of compensation package $F^{ed}(\theta, INS)$.
 - (b) Then, I numerically solve individual value functions backwards from the period $T - 1$. The main obstacle of solving individual life cycle problems is the large

size of state spaces. There, to speed up the computation, I apply Keane and Wolpin (1994)'s interpolation method.

- (c) Given the value function, I solve the steady state distribution $\hat{g}_t(\tilde{X}, INS, \theta)$ and $u_t(\tilde{X})$ sequentially from age $t = 1$.
- (d) Using $\hat{g}_t(\tilde{X}, INS, \theta)$ and $u_t(\tilde{X})$, I solve $(\langle \hat{\theta}_0^{ed}(p), \hat{\theta}_1^{ed}(p) \rangle^{ed}, \hat{\Delta}(p))$ for each p using (2.17), (2.18), and (2.14).

3. After completing the step (d) at iteration ι , I check whether the equilibrium object converges.

- (a) If $(\langle \theta_{0,\iota}^{ed}(p), \theta_{1,\iota}^{ed}(p) \rangle^{ed}, \Delta(p))$ satisfies $d(\theta_{0,\iota}^{ed}(p), \hat{\theta}_0^{ed}(p)) < \epsilon_{tol}$, $d(\theta_{1,\iota}^{ed}(p), \hat{\theta}_1^{ed}(p)) < \epsilon_{tol}$ for $ed \in \{NC, C\}$ and $d(\Delta^{\iota^*}(p), \Delta_\iota(p)) < \epsilon_{tol}$ where ϵ_{tol} is a pre-specified tolerance level of convergence and $d(\cdot, \cdot)$ is a distance metric, then firm's optimal policy converges and we have an equilibrium.

- (b) Otherwise, update $(\langle \theta_{0,\iota+1}^{ed}(p), \theta_{1,\iota+1}^{ed}(p) \rangle^{ed}, \Delta_{\iota+1}(p))$ as follows:

$$\begin{aligned}\theta_{0,\iota+1}^{ed}(p) &= w\theta_{0,\iota}^{ed}(p) + (1-w)\hat{\theta}_0^{ed}(p), \\ \theta_{1,\iota+1}^{ed}(p) &= w\theta_{1,\iota}^{ed}(p) + (1-w)\hat{\theta}_1^{ed}(p) \\ \Delta_{\iota+1}(p) &= w\Delta_\iota(p) + (1-w)\hat{\Delta}(p),\end{aligned}$$

for the pre-specified weight $w \in (0, 1)$ and continue Step 2 at iteration $\iota' = \iota + 1$.

B.2 Omitted Derivations in the Counterfactual Experiments

B.2.1 Value Function

Consider a non-employed worker having characteristics $\tilde{X}_t = (ed, type, E_t, h_t)$ and insurance status $INS \in \{0, 2\}$. Then, his value function is defined by

$$V_0^t(\tilde{X}_t, INS) = \mathbb{E}_{\epsilon_t^m} \left[\begin{array}{c} \max_{x_t} U_t(C_t(x_t, \epsilon_t^m, \tilde{X}_t, INS), 0, h_t) \\ +\beta \sum_{\hat{h}} \Pr(h_t = h | \tilde{X}_t, x_t, \epsilon_t^m) \lambda_u^{ed} \hat{V}_0^t(\tilde{X}_{t+1}) \\ +\beta \sum_{\hat{h}} \Pr(h_t = \hat{h} | \tilde{X}_t, x_t, \epsilon_t^m) (1 - \lambda_u^{ed}) V_0^{t+1}(\tilde{X}_{t+1}) \end{array} \right], \quad (\text{B.21})$$

subject to budget constraint (2.19) where $\tilde{X}_{t+1} = (ed, type, E_t, \hat{h})$ and

$$\hat{V}_0^t(\tilde{X}_{t+1}) = \int \mathbb{E}_{\epsilon_t^n} \max \left\{ \begin{array}{c} V_0^{t+1}(\tilde{X}_{t+1}) + \epsilon_t^n, \\ V_1^{t+1}(\tilde{X}_{t+1}, \theta, INS) \end{array} \right\} dF^{ed}(\theta, INS)$$

$$V_0^{t+1}(\tilde{X}_{t+1}) = \max \left\{ V_0^{t+1}(\tilde{X}_{t+1}, 0) + \epsilon_t^{HIX}, V_0^{t+1}(\tilde{X}_{t+1}, 2) \right\}.$$

ϵ_t^{HIX} is i.i.d. preference shock to purchase health insurance from HIX. Note that the main change is a possibility that the unemployed can purchase health insurance from HIX.

Similarly, consider an employed worker having characteristics $\tilde{X}_t = (ed, type, E_t, h_t)$ and insurance status $INS \in \{0, 1, 2\}$ and who are offered skill price θ . His value function is also defined by:

$$V_1^t(\tilde{X}_t, \theta, INS) = \mathbb{E}_{\epsilon_t^m} \left[\begin{array}{c} \max_{x_t} U_t(C_t(x_t, \epsilon_t^m, \tilde{X}_t, \theta, INS), 1, h_t) \\ +\beta \sum_{\hat{h}} \Pr(h_t = \hat{h} | \tilde{X}_t, x_t, \epsilon_t^m) \delta^{ed} \hat{V}_1^t(\tilde{X}_{t+1}) \\ +\beta \sum_{\hat{h}} \Pr(h_t = \hat{h} | \tilde{X}_t, x_t, \epsilon_t^m) (1 - \delta^{ed}) \tilde{V}_1^t(\tilde{X}_{t+1}, INS, \theta) \end{array} \right] \quad (\text{B.22})$$

subject to budget constraint (2.19) where $\tilde{X}_{t+1} = (ed, type, E_t + 1, \hat{h})$ and $\hat{V}_1^t(\tilde{X}_{t+1})$ and $\tilde{V}_1^t(\tilde{X}_{t+1}, INS, \theta)$ are defined as follows:

$$\begin{aligned} \hat{V}_1^t(\tilde{X}_{t+1}) &= (1 - \lambda_e^{ed})V_0^{t+1}(\tilde{X}_{t+1}) \\ &+ \lambda_e^{ed} \int \mathbb{E}_{\epsilon_t^n} \max \left\{ \begin{array}{l} V_0^{t+1}(\tilde{X}_{t+1}) + \epsilon_t^n, \\ V_1^{t+1}(\tilde{X}_{t+1}, \theta', INS') \end{array} \right\} dF(\theta', INS') \end{aligned}$$

where $V_0^{t+1}(\tilde{S}_{t+1}^0)$ is defined above; $V_1^{t+1}(\tilde{X}_{t+1}, \theta', INS')$ is defined as follows. If the health insurance is offered from the employer, then

$$V_1^{t+1}(\tilde{X}_{t+1}, \theta', INS') = V_1^{t+1}((\tilde{X}_{t+1}, \theta', 1)),$$

Otherwise, it takes the form

$$V_1^{t+1}(\tilde{X}_{t+1}, \theta', INS') = \max \left\{ V_1^{t+1}(\tilde{X}_{t+1}, \theta, 0) + \epsilon_t^{HIX}, V_1^{t+1}(\tilde{X}_{t+1}, \theta, 2) \right\}.$$

Similarly,

$$\begin{aligned} \tilde{V}_1^t(\tilde{X}_{t+1}, \theta, INS) &= (1 - \lambda_e^{ed})\mathbb{E}_{\epsilon_t^n} \max \left\{ \begin{array}{l} V_0^{t+1}(\tilde{X}_{t+1}) + \epsilon_t^n, \\ V_1^{t+1}(\tilde{X}_{t+1}, \theta, INS) \end{array} \right\} \\ &+ \lambda_e^{ed} \int \mathbb{E}_{\epsilon_t^n} \max \left\{ \begin{array}{l} V_0^{t+1}(\tilde{X}_{t+1}) + \epsilon_t^n, \\ V_1^{t+1}(\tilde{X}_{t+1}, \theta, INS), \\ V_1^{t+1}(\tilde{X}_{t+1}, \theta', INS') \end{array} \right\} dF(\theta', INS') \end{aligned}$$

The difference between $V_1^{t+1}(\tilde{X}_{t+1}, \theta, INS)$ and $V_1^{t+1}(\tilde{X}_{t+1}, \theta', INS')$ is that the compensation package in the former is determined by the current employer, but by a potential employer in the latter. The terminal value is the same as (2.6).

B.2.2 Steady State Worker Distribution

The steady state distribution of $g_t(\tilde{X}, \theta_{INS}^{ed}, INS)$ and $u_t(\tilde{X}, INS)$ are characterized as follows. First, define $\tilde{g}_t(\tilde{X}, \theta^{ed}, NINS)$ as the measure of employed workers having characteristics (t, \tilde{X}) who are offered skill price θ but not offered health insurance:

$$\tilde{g}_t(\tilde{X}, \theta, NINS) = g_t(\tilde{X}, \theta^{ed}, 0) + g_t(\tilde{X}, \theta^{ed}, 2).$$

It is determined by

$$\begin{aligned} \frac{\tilde{g}_t(\tilde{X}, \theta, NINS)}{1+n} &= \sum_{INS'=0,2} \sum_{h_{t-1}} g_{t-1}(\tilde{X}_{-1}^A, \theta, INS') \\ &\quad \times \mathbb{E}_{\epsilon_t^m} \left[\Pr(h_t = \hat{h} | \tilde{X}_{-1}^A, \theta, INS', \epsilon_t^m) \right] \\ &\quad \times \left[\begin{array}{l} (1 - \delta^{ed})(1 - \lambda_e^{ed}) \Pr(\Omega_1^E(\theta, 1, \tilde{X})) \\ + (1 - \delta^{ed})\lambda_e^{ed} (\Pr(\Omega_2^E(\theta, 1, \tilde{X})) \Pr(\Omega_1(\theta, 1, \tilde{X}))) \end{array} \right] \\ &+ \sum_{h_{t-1}} \sum_{INS'} u_{t-1}(\tilde{X}_{-1}^B, INS') \\ &\quad \times \mathbb{E}_{\epsilon_t^m} \left[\Pr(h_t = \hat{h} | \tilde{X}_{-1}^B, \epsilon_t^m) \right] \Pr(\Omega_1^E(\theta, 1, \tilde{X})) \\ &\quad \times \lambda_u^{ed} f^{ed}(\theta, 1) \\ &+ \sum_{h_{t-1}} \sum_{INS'} \int g_{t-1}(\tilde{X}_{-1}^A, \theta', INS') \\ &\quad \times \mathbb{E}_{\epsilon_t^m} \left[\Pr(h_t = \hat{h} | \tilde{X}_{-1}^A, \theta', INS', \epsilon_t^m) \right] \mathbf{1}(\theta, 1, \theta', INS') d\theta' \\ &\quad \times (1 - \delta^{ed})\lambda_e^{ed} f^{ed}(\theta, 1) \\ &+ \sum_{h_{t-1}} \sum_{INS'} \int g_{t-1}(\tilde{X}_{-1}^A, \theta', INS') \\ &\quad \times \mathbb{E}_{\epsilon_t^m} \left[\Pr(h_t = \hat{h} | \tilde{X}_{-1}^A, \theta', INS', \epsilon_t^m) \right] \Pr(\Omega_1^E(\theta, 1, \tilde{X})) d\theta' \\ &\quad \times \delta^{ed}\lambda_e^{ed} f^{ed}(\theta, 1) \end{aligned} \tag{B.23}$$

where $\tilde{X}_{-1}^A = (ed, type, E_{t-1}, \hat{h})$ and $\tilde{X}_{-1}^B = (ed, type, E_t, \hat{h})$ are individual characteristics in the last period for the employed and non-employed respectively which can turn into \tilde{X} in this period, and $\Pr(\Omega_1^E(\tilde{X}, \theta, INS))$ is the probability that individuals with \tilde{X} prefer

to work a job with compensation package (θ^{ed}, INS) over being non-employed, which can formally be expressed as

$$\Pr(\Omega_1^E(\tilde{X}, \theta, INS)) = \Pr(\tilde{V}_0^t(\tilde{X}) + \epsilon_t^n < V_1^t(\tilde{X}, \theta^{ed}, INS))$$

where the value of choosing non-employed is

$$\tilde{V}_0^t(\tilde{X}) = \mathbb{E}[\max\{\tilde{V}_0^t(\tilde{X}, 0), \tilde{V}_0^t(\tilde{X}, 2) + \epsilon_t^{HIX}\}]$$

and $\Pr(\Omega_2^E(\theta^{ed}, INS))$ is the probability that individuals who receive a job offer from other firms decide to stay the current job:

$$\Pr(\Omega_2^E(\theta, INS, \tilde{X})) = F^{ed}(\theta, INS) + F^{ed}(\tilde{\theta}_{\widehat{INS}}(\tilde{X}, \theta), \widehat{INS})$$

for $\widehat{INS} \neq INS$ where $\tilde{\theta}_{\widehat{INS}}(\tilde{X}, \theta)$ is threshold skill price which can be defined as

$$V_1^t(\tilde{X}, \theta, INS) = V_1^t(\tilde{X}, \tilde{\theta}_{\widehat{INS}}(\tilde{X}, \theta), \widehat{INS}).$$

$\mathbf{1}(\theta, INS, \theta', INS')$ is the indicator function such that individuals prefer to take an offer from (θ, INS) over (θ', INS') :

$$\mathbf{1}(\theta, INS, \theta', INS') = \begin{cases} 1 & \text{if } V_1^t(\tilde{X}, \theta, INS) > V_1^t(\tilde{X}, \theta', INS') \\ 0 & \text{otherwise} \end{cases}$$

Then, one can characterize $g_t(\tilde{X}, \theta, INS)$ as

$$g_t(\tilde{X}, \theta, 2) = \mathbb{E}_{\epsilon_t^{HIX}}[\Pr[V_1^t(\tilde{X}, \theta, 2) + \epsilon_t^{HIX} > V_1^t(\tilde{X}, \theta, 0)]]\tilde{g}_t(\tilde{X}, \theta, NINS)$$

and $g_t(\tilde{X}, \theta, 0) = \tilde{g}_t(\tilde{X}, \theta, NINS) - g_t(\tilde{X}, \theta, 2)$. $g_t(\tilde{X}, \theta, 1)$ is derived in the same way as in the main text. Therefore, I omit its derivation here.

One can characterize the determinants of steady state measures of non-employed at age t with characteristics \tilde{X} , $\tilde{u}_t(\tilde{X})$, as follows. First, define $\tilde{u}_t(\tilde{X})$ as the measure of non-employed workers with characteristics (t, \tilde{X}) which satisfy

$$\tilde{u}_t(\tilde{X}) = u_t(\tilde{X}, 0) + u_t(\tilde{X}, 2)$$

$$\begin{aligned}
\frac{\tilde{u}_t(\tilde{X})}{1+n} &= \sum_{INS'=0.2} \sum_{h_{t-1}} u_{t-1}(\tilde{X}_{-1}^B, INS') \\
&\quad \times \mathbb{E}_{\epsilon_t^m} \left[\Pr(h_t = \hat{h} | \tilde{X}_{-1}^B, \epsilon_t^m) \right] (1 - \lambda_u^{ed}) \\
&+ \sum_{INS'=0.2} \sum_{h_{t-1}} u_{t-1}(\tilde{X}_{-1}^A, INS') \\
&\quad \times \mathbb{E}_{\epsilon_t^m} \left[\Pr(h_t = \hat{h} | \tilde{X}_{-1}^A, \epsilon_t^m) \right] \Pr(\Omega_1^U(\tilde{X})) \lambda_u^{ed} \\
&+ \sum_{h_{t-1}} \sum_{INS} \int g_{t-1}(\tilde{X}_{-1}^A, \theta, INS) \\
&\quad \times \mathbb{E}_{\epsilon_t^m} \left[\Pr(h_t = \hat{h} | \tilde{X}_{-1}^A, \epsilon_t^m) \right] \Pr(\Omega_2^U(\tilde{X}, \theta^{ed}, INS)) dF(p')(1 - \delta^{ed}) \\
&+ \sum_{h_{t-1}} \sum_{INS} \int g_{t-1}(\tilde{X}_{-1}^A, \theta, INS) \\
&\quad \times \mathbb{E}_{\epsilon_t^m} \left[\Pr(h_t = \hat{h} | \tilde{X}_{-1}^A, \epsilon_t^m) \right] d\theta \delta^{ed} (1 - \lambda_e^{ed}) \\
&+ \sum_{h_{t-1}} \sum_{INS} \int g_{t-1}(\tilde{X}_{-1}^A, \theta, INS) \\
&\quad \times \mathbb{E}_{\epsilon_t^m} \left[\Pr(h_t = \hat{h} | \tilde{X}_{-1}^A, \epsilon_t^m) \right] \Pr(\Omega_1^U(\tilde{X})) d\theta \delta^{ed} \lambda_e^{ed} \tag{B.24}
\end{aligned}$$

where \tilde{X}_{-1}^A and \tilde{X}_{-1}^B are defined as above, and $\Pr(\Omega_1^U(\tilde{X}))$ is the probability that the non-employed with characteristics \tilde{X} decides to turn down the offer:

$$\Pr(\Omega_1^U(\tilde{X})) = \int \Pr(\tilde{V}_0^t(\tilde{X}) + \epsilon_t^n > V_1^t(\tilde{X}, \theta^{ed}, INS)) dF(\theta^{ed}, INS).$$

$\Pr(\Omega_2^U(\tilde{X}, \theta^{ed}, INS))$ is the probability that the employed with characteristics \tilde{X} and a

job with compensation package (θ^{ed}, INS) decides to quit into the non-employed:

$$\Pr(\Omega_2^U(\tilde{X}, \theta^{ed}, INS)) = \Pr(\tilde{V}_0^t(\tilde{X}) + \epsilon_t^n > V_1^t(\tilde{X}, \theta^{ed}, INS)).$$

Then, one can characterize $u_t(\tilde{X}, INS)$ as

$$u_t(\tilde{X}, 2) = \mathbb{E}_{\epsilon_t^{HIX}} [\Pr [V_0^t(\tilde{X}, 2) + \epsilon_t^{HIX} > V_0^t(\tilde{X}, 0)]] \tilde{u}_t(\tilde{X})$$

and $u_t(\tilde{X}, 0) = \tilde{u}_t(\tilde{X}) - u_t(\tilde{X}, 2)$.

B.3 Parameterization of Policy Parameters

I describe the approach to parameterize the stylized version of the ACA in the model. The approach follows Aizawa and Fang (2013) which examines the impact of ACA on labor market outcomes, but several modifications are made to fit the model environment in this study.

B.3.1 Penalties associated with individual mandate

The tax penalty on the uninsured in the ACA (from 2016 when the law is fully implemented) is set that the uninsured need to pay a tax penalty of the greater value of \$695 per year or 2.5% of the taxable income above the Tax Filing Threshold (TFT), which can be written as:

$$IM^{ACA}(y) = \max \{0.025 \times (y - \text{TFT}_{2011}), \$695\}, \quad (\text{B.35})$$

where y is annual income.

I adjust the above formula in several dimensions. First, I adjust the scale of policy parameters to fit the 2007 economic environment. I estimated the model using data sets

in 2004-2007 where the price level is normalized to 2007 value, while the ACA policy parameters are chosen to suit the economy in 2011. It is well known that the U.S. health care sector has a very different growth rate than that of overall GDP; in particular, there are substantial increases in medical care costs relative to GDP. Thus I need to appropriately adjust the policy parameters in the ACA to make them more in line with the U.S. economy in 2007. I implement the adjustment as follows: the \$695 amount is adjusted by the ratio of the 2007 Medical Care CPI (CPI_Med_2007) relative to the 2011 Medical Care CPI (CPI_Med_2011); I choose this adjustment given the idea that the penalty amount \$695 is chosen to be proportional to 2011 medical expenditures. I then multiply it by 1/3 to reflect the fact that the period-length in the model is four-month. Second, I adjust the TFT_2011 by the ratio of 2007 CPI of all goods (CPI_All_2007) relative to the 2011 CPI of all goods (CPI_All_2011). I also multiply it by 1/3 to reflect the choice of the four-month model period in this study.¹ Finally, I adjust the percentage 2.5% by the differential growth rate of medical care and GDP, i.e., multiply it by the relative ratio of $\frac{CPI_Med_2007}{CPI_All_2007}$ and $\frac{CPI_Med_2007}{CPI_All_2007}$. With these adjustments, the tax penalties on the uninsured are parameterized as:

$$IM^{ACA}(y) = \max \left\{ \begin{array}{l} 0.025 \times \frac{\left(\frac{CPI_Med_2007}{CPI_All_2007}\right)}{\left(\frac{CPI_Med_2011}{CPI_All_2011}\right)} \times \left(y - \frac{1}{3}TFT_2011 \right) \\ \times \frac{CPI_All_2007}{CPI_All_2011} \\ \frac{1}{3} \times \$695 \times \frac{CPI_Med_2007}{CPI_Med_2011} \end{array} \right\},$$

where w is four-month income in dollars.

B.3.2 Penalties associated with employer mandate

Tax penalties on employers in the ACA are set that firms with 50 or more full-time employees that do not offer coverage need to pay a tax penalty of \$2,000 per full-time

¹I obtain CPI data for medical care and all goods both from Bureau of Labor Statistics website: <http://www.bls.gov/cpi/data.htm>.

employee per year, excluding the first 30 employees from the assessment.² That is,

$$EM^{ACA}(l) = (l - 30) \times \$2,000. \quad (\text{B.36})$$

As in the case of individual mandate, I first adjust the above formula by first scaling the \$2,000 per-worker penalty using the ratio of the 2007 Medical Care CPI relative to the 2011 Medical Care CPI and by multiplying it by 1/3 to reflect our period-length of four months instead of a year, i.e., for $l \geq 50$,

$$EM^{ACA}(l) = \frac{1}{3} \left[(l - 30) \times \$2,000 \times \frac{\text{CPI_Med_2007}}{\text{CPI_Med_2011}} \right].$$

While it is ideal to apply this formula precisely in the model, to simplify the numerical algorithm, I approximate this penalty function as a differentiable function by removing the discontinuity, by following the approximation technique used by MaCurdy, Green, and Paarsch (1990).

B.3.3 Medicaid expansion

ACA stipulates that individuals with income below 133% of Federal Poverty Level (FPL) are able to enroll in the free public insurance Medicaid. While it is ideal to model this threshold carefully, given my sample selection those who are below 133% of Federal Poverty Level (FPL) tend to be non-employed. Moreover, there is a certain technical difficulty to model this threshold as described below. Therefore, to simplify the analysis, I assume that only non-employed individuals will be covered by Medicaid. As I mention in Section 2.6.1, I also consider the case that Medicaid expansion is not implemented.

²In July 2013, the government decided to postpone the implementation of the employer mandate until 2015.

B.3.4 Premium subsidies in HIX

In the ACA, federal premium subsidies are available to individuals who purchase health insurance from HIX if their incomes are less than 400 % of the Federal Poverty Level (FPL), denoted by FPL400³. The premium subsidies will be set on a sliding scale such that the premium contributions are limited to a certain percentage of income for specified income levels. If an individual's income is at 133% of the FPL, denoted by FPL133, premium subsidies will be provided so that the individual's contribution to the premium is equal to 3.5% of his income; when an individual's income is at FPL400, his premium contribution is set to be 9.5% of the income. If his income is above FPL400, he is no longer eligible for premium subsidies. Note that the premium support rule as described in ACA creates a discontinuity at FPL133: individuals with income below FPL133 receives free Medicaid, but those at or slightly above FPL133 have to contribute at least 2.3% of his income to purchase health insurance from HIX. To avoid this discontinuity issue, I instead adopt a slightly modified premium support formula as follows:

$$S(y, R^{HIX}(t)) = \begin{cases} \max \left\{ \begin{aligned} &R^{HIX}(t) - \left[0.0350 + 0.060 \frac{3y}{\text{FPL400}}\right] y \\ &\times \frac{\text{CPLMed.2007}}{\text{CPLMed.2011}}, 0 \end{aligned} \right\} & \text{if } y < \frac{\text{FPL400}}{3} \\ 0, & \text{otherwise,} \end{cases} \quad (\text{B.37})$$

where w is four-month income.

As I mention in Section 2.6.1, at this stage, it is unclear whether premium subsidies will be given to individuals with less than FPL133 who live in states where Medicaid is no expanded. In this study, I assume that those individuals will obtain health insurance at zero premium from HIX.

³I assume that FPL is defined for a single person. In 2007, it is \$11,200 annually.

B.3.5 Age-based pricing regulation

In the ACA, they set the maximum premium ratio between the oldest and the youngest $\omega_A = 3$. Therefore, if it binds,

$$\omega_{AGE} R^{HIX}(1) = R^{HIX}(T).$$

The issue is that it does not specify how the premium can vary over age. For simplicity, I assume that the premium is linearly increasing in t . Given this restriction, in order to satisfy the market clearing condition in HIX, insurance premia in HIX must be determined as follows:

$$\begin{aligned} R^{HIX}(t) &= R^{HIX}(1) + (\omega_{AGE} - 1) \frac{(t-1)}{T-1} R^{HIX}(1) \\ R^{HIX}(1) &= (1 + \xi_{HIX}) R_{HIX}^* \left[\frac{T-1}{\omega_{AGE} - 1 + T - 1 + \frac{(\omega_{AGE}-1)(T+1)}{2}} \right] \end{aligned} \quad (\text{B.38})$$

where R_{HIX}^* is the pooled premium among the all the participants in HIX:

$$R_{HIX}^* = (1 + \xi_{HIX}) \sum_t \sum_{\tilde{X}} \int E[m_*^{\tilde{X},t} | INS_t = 2] g_t(\tilde{X}, \theta, 2) d\theta.$$

To understand the role of age based pricing regulation, I also consider the case where there is no regulation: in such a case, premia are determined as

$$R^{HIX}(t) = (1 + \xi_{HIX}) \sum_{\tilde{X}} \int E[m_*^{\tilde{X},t} | INS_t = 2] g_t(\tilde{X}, \theta, 2) d\theta. \quad (\text{B.39})$$

That is, each age group consists of separate risk pool.

Appendix C

Appendix to Chapter 3

C.1 Constructing Health Status

In our analysis, an individual's health status is measured as expected Medicare Parts A and B claims cost if the individual were to receive insurance from traditional Medicare (Medicare Parts A and B). In order to construct the variable, we use the individual-level data. Because individuals in MA are not directly covered by traditional Medicare, information on Medicare Part A and B claims is available only for those in traditional Medicare. Therefore we need to impute predicted Medicare costs for MA enrollees. Our construction of the health status variable has two steps:

1. First, using beneficiaries in traditional Medicare, we estimate two equations that relate Medicare claims costs to an extensive list of health status and demographic characteristics.
2. We calculate predicted claims cost for traditional Medicare enrollees using the estimates. We impute the predicted Medicare claims costs for MA enrollees in the data, using their observed health and demographic characteristics and the estimates obtained in the first step.

First Step In the first step, we estimate two equations that relate an individual’s realized Medicare claims cost to an extensive list of health and demographic variables. In the first equation, we estimate the probability that an individual ever incurs positive Medicare claims cost. Approximately 5.6% of individuals have zero claims cost in a given year, and we account for the possibility of zero health expenditure using the following logistic regression:

$$Prob(y > 0|x) = \frac{\exp(x\beta_1)}{1 + \exp(x\beta_1)}. \quad (C.11)$$

y denotes an individual’s Medicare claims cost, and x is a vector of health and demographic characteristics. For x , we include an extensive list of health variables such as self-reported health status, whether an individual has difficulties in activities of daily living (ADL) and instrumental activities of daily living (IADL), and histories of diseases such as cancer, heart disease, diabetes, etc. We also include the average Medicare claims cost for each county and year to control for regional differences in health care costs. In the end, we include 76 variables in x . Parameter β_1 is estimated with maximum likelihood, and the results are presented in Table C.1.

Using the second equation, we estimate the relationship between an amount of Medicare claims cost and health characteristics for individuals having positive claims costs. We estimate the following equation:

$$\begin{aligned} \log(y) &= x\beta + \epsilon \\ \epsilon &\sim N(0, (z\gamma)^2) \end{aligned}$$

where y and x are the same as in the first equation; and z is a subset of x that includes self-reported health status, whether an individual is living in a skilled nursing facility, average Medicare claims cost for each county, and interaction terms between county-level average Medicare claims costs and other variables in Z . We estimate parameters β and

γ with the method of moments. The first set of moments is:

$$E[\log(y)|x, y > 0] = x\beta_2.$$

The second set of moments is:

$$E[y|z, y > 0] = \exp\left(x\beta_2 + \frac{(z\gamma)^2}{2}\right).$$

The right-hand side of the second condition is derived from the assumption that ϵ is normally distributed. The first set of moments will pin down β_2 , and the second set of moments will pin down γ . The estimates are presented in Table C.2.

Note that we make an implicit assumption here that ϵ is independent of the logistic error term for equation (C.11). This means that a correlation between $Prob(y > 0|x)$ and $E[y|x]$ only depends on x , not on the error terms. Although it is possible to allow for correlated error term, we make such an assumption for simplicity.

Second Step Given estimates of parameters $\hat{\beta}_1$, $\hat{\beta}_2$, and $\hat{\gamma}$, we calculate predicted Medicare claims cost for each individual. Because y is not observed only for individuals in MA, we have to impute predicted Medicare claims for MA enrollees using the estimates. An important assumption we make for the imputation is that x contains all relevant health characteristics of an individual. That is, individuals in MA and traditional Medicare are not different in unobserved health, conditional on x . This assumption implies that ϵ is a purely random shock to claims costs, and individuals do not select on ϵ when choosing between MA and traditional Medicare. Without this assumption, the imputation of predicted Medicare claims costs for MA enrollees will not be valid. Although it is possible that x may not capture all relevant health characteristics, the large number of variables in x would minimize the role of unobserved health characteristics.

We calculate predicted Medicare claims cost in the following way:

$$\begin{aligned} E[y|z] &= Prob(y > 0|x) \times E[y|x, y > 0] \\ &= \frac{\exp(x\hat{\beta}_1)}{1 + \exp(x\hat{\beta}_1)} \times \exp\left(x\hat{\beta}_2 + \frac{(z\hat{\gamma})^2}{2}\right). \end{aligned}$$

Dependent Variable: Dummy for Positive Medicare Claims Cost		
Variables	Coefficient	Std. Err.
Black	-0.781***	(0.0764)
Hispanic	-0.540***	(0.0957)
Living in a nursing home	1.924**	(0.815)
Health status: excellent	-0.286*	(0.153)
Health status: very good	0.00109	(0.149)
Health status: good	0.159	(0.145)
Health status: fair	0.354**	(0.147)
Difficulty using phone	-0.123	(0.106)
Difficulty light housework	-0.0820	(0.134)
Difficulty heavy housework	0.384***	(0.0809)
Difficulty preparing meals	0.545***	(0.162)
Difficulty shopping	-0.0864	(0.125)
Difficulty handling bills	-0.307**	(0.123)
Difficulty bathing	0.187	(0.131)
Difficulty dressing	-0.283*	(0.167)
Difficulty eating	-0.282	(0.206)
Difficulty stooping	-0.156	(0.109)
Difficulty walking	-0.0743	(0.0826)
Difficulty using toilet	0.304*	(0.178)
History with skin cancer	0.533***	(0.0726)
History with other cancers	0.622***	(0.0783)
History of high blood pressure	0.583***	(0.0501)
History of heart attack	0.219***	(0.0845)
History of angina pectoris	0.342***	(0.0963)
History of other heart conditions	0.420***	(0.0757)
History of stroke	0.253***	(0.0896)
History of rheumatoid arthritis	0.0372	(0.0972)
History of arthritis	0.411***	(0.0500)
History of diabetes	0.675***	(0.0827)
County-level Medicare cost	-0.00965***	(0.00182)
County-level Medicare cost × Nursing home	0.00123	(0.00139)
County-level Medicare cost × Age	0.000131***	(2.42e-05)
Medicaid	0.745***	(0.0961)
Employer-sponsored insurance benefit dummy	0.363***	(0.0550)
Observations	44,088	
Pseudo R-squared	0.158	

Table C.1: Logit Regression for Positive Medicare Claims Cost

Note: Other controls included are dummy variables for various groups of age, gender, interactions of age and gender, income, education, marital status, self-reported health status compared to a year ago. The number of variables included in this logit regression is 78.

Source: Medicare Current Beneficiary 2000–2004.

Dependent Variable: Medicare Claims Cost		
Variables	Coefficient	Std. Err.
Black	-0.0420	(0.0333)
Hispanic	-0.0149	(0.0461)
Living in a nursing home	0.513***	(0.194)
Health status: excellent	-0.580***	(0.0525)
Health status: very good	-0.347***	(0.0489)
Health status: good	-0.140***	(0.0469)
Health status: fair	-0.0758*	(0.0457)
Difficulty using phone	-0.188***	(0.0374)
Difficulty light housework	0.0469	(0.0396)
Difficulty heavy housework	0.204***	(0.0238)
Difficulty preparing meals	0.143***	(0.0447)
Difficulty shopping	-0.00237	(0.0380)
Difficulty handling bills	-0.0195	(0.0411)
Difficulty bathing	0.199***	(0.0380)
Difficulty dressing	0.0639	(0.0463)
Difficulty eating	-0.0271	(0.0633)
Difficulty stooping	-0.111***	(0.0335)
Difficulty walking	0.0477*	(0.0264)
Difficulty using toilet	0.108**	(0.0493)
History with skin cancer	0.240***	(0.0199)
History with other cancers	0.437***	(0.0214)
History of high blood pressure	0.104***	(0.0181)
History of heart attack	0.228***	(0.0256)
History of angina pectoris	0.224***	(0.0263)
History of other heart conditions	0.284***	(0.0208)
History of stroke	0.124***	(0.0266)
History of rheumatoid arthritis	0.147***	(0.0270)
History of arthritis	0.174***	(0.0180)
History of diabetes	0.348***	(0.0210)
County-level Medicare cost	0.00109*	(0.000627)
County-level Medicare cost × Nursing home	0.00108***	(0.000282)
County-level Medicare cost × Age	1.87e-05**	(7.95e-06)
Medicaid	0.0820**	(0.0323)
Employer-sponsored insurance benefit dummy	-0.0112	(0.0179)
Observations	41,603	
R-squared	0.249	

Table C.2: Regression of Medicare Claims Costs on Health Characteristics

Note: For this regression, only the individuals with positive Medicare claims costs are included. Other controls included are dummy variables for various groups of age, gender, interactions of age and gender, income, education, marital status, self-reported health status compared to a year ago. The number of variables included in this logit regression is 78.

Source: Medicare Current Beneficiary 2000–2004.

Bibliography

- ACKERBERG, D. A. (2001): “Empirically Distinguishing Informative and Prestige Effects of Advertising,” *RAND Journal of Economics*, 316–333.
- (2003): “Advertising, Learning, and Consumer Choice in Experience Good Markets: an Empirical Examination,” *International Economic Review*, 44, 1007–1040.
- AIZAWA, N. AND H. FANG (2013): “Equilibrium Labor Market Search and Health Insurance Reform,” Working Paper 18698, National Bureau of Economic Research.
- AKERLOF, G. A. (1978): “The Economics of ”Tagging” as Applied to the Optimal Income Tax, Welfare Programs, and Manpower Planning,” *The American Economic Review*, 68, 8–19.
- BAGGER, J., F. FONTAINE, F. POSTEL-VINAY, AND J.-M. ROBIN (2013): “Tenure, Experience, Human Capital and Wages: A Tractable Equilibrium Search Model of Wage Dynamics,” *Forthcoming, American Economic Review*.
- BAGWELL, K. (2007): “The Economic Analysis of Advertising,” *Handbook of Industrial Organization*, 3, 1701–1844.
- BAICKER, K., A. FINKELSTEIN, J. SONG, AND S. TAUBMAN (2013): “The Impact of Medicaid on Labor Force Activity and Program Participation: Evidence from the Oregon Health Insurance Experiment,” Working Paper 19547, National Bureau of Economic Research.

- BAJARI, P., H. HONG, AND A. KHWAJA (2011): “A semiparametric analysis of adverse selection and moral hazard in health insurance contracts,” Tech. rep., Working Paper.
- BARLEVY, G. (2008): “Identification of Search Models using Record Statistics,” *The Review of Economic Studies*, 75, 29–64.
- BATATA, A. (2004): “The Effect of HMOs on Fee-For-Service Health Care Expenditures: Evidence from Medicare Revisited,” *Journal of health economics*, 23, 951–963.
- BAUHOFF, S. (2012): “Do health plans risk-select? An audit study on Germany’s Social Health Insurance,” *Journal of Public Economics*, 96, 750–759.
- BECKER, G. S. AND K. M. MURPHY (1993): “A simple theory of advertising as a good or bad,” *The Quarterly Journal of Economics*, 108, 941–964.
- BERRY, S., J. LEVINSOHN, AND A. PAKES (1995): “Automobile prices in market equilibrium,” *Econometrica*, 841–890.
- BLAU, D. M. AND D. B. GILLESKIE (2008): “The Role of Retiree Health Insurance in the Employment Behavior of Older Men,” *International Economic Review*, 49, 475–514.
- BLUNDELL, R. AND A. SHEPHARD (2012): “Employment, Hours of Work and the Optimal Taxation of Low-Income Families,” *The Review of Economic Studies*, 79, 481–510.
- BONTEMPS, C., J.-M. ROBIN, AND G. J. VAN DEN BERG (1999): “An Empirical Equilibrium Job Search Model with Search on the Job and Heterogeneous Workers and Firms,” *International Economic Review*, 40, 1039–1074.
- (2000): “Equilibrium Search with Continuous Productivity Dispersion: Theory and Nonparametric Estimation,” *International Economic Review*, 41, 305–358.
- BROWN, J., M. DUGGAN, I. KUZIEMKO, AND W. WOOLSTON (2012): “How Does Risk Selection Respond to Risk Adjustment? Evidence from the Medicare Advantage Program,” *NBER Working Paper*.

- BRUEGEMANN, B. AND I. MANOVSKII (2010): “Fragility: A Quantitative Analysis of the US Health Insurance System,” Working Paper, University of Pennsylvania.
- BUNDORF, M. K., J. LEVIN, AND N. MAHONEY (2012a): “Pricing and Welfare in Health Plan Choice,” *American Economic Review*, 102, 3214–48.
- (2012b): “Pricing and Welfare in Health Plan Choice,” *The American Economic Review*, 102, 3214–3248.
- BURDETT, K., C. CARRILLO-TUDELA, AND M. G. COLES (2011): “Human Capital Accumulation and Labor Market Equilibrium,” *International Economic Review*, 52, 657–677.
- BURDETT, K. AND D. T. MORTENSEN (1998): “Wage Differentials, Employer Size, and Unemployment,” *International Economic Review*, 39, 257–273.
- CAHUC, P., F. POSTEL-VINAY, AND J.-M. ROBIN (2006): “Wage Bargaining with On-the-Job Search: Theory and Evidence,” *Econometrica*, 74, 323–364.
- CARLIN, C. AND R. TOWN (2007): “Adverse selection, welfare and optimal pricing of employer-sponsored health plans,” *U. Minnesota Working Paper*.
- CEBUL, R. D., J. B. REBITZER, L. J. TAYLOR, AND M. E. VOTRUBA (2011): “Unhealthy Insurance Markets: Search Frictions and the Cost and Quality of Health Insurance,” *American Economic Review*, 101, 1842–71.
- CHETTY, R. (2006): “A New Method of Estimating Risk Aversion,” *The American Economic Review*, 96, 1821–1834.
- CHETTY, R. AND A. FINKELSTEIN (2012): “Social Insurance: Connecting Theory to Data,” Working Paper 18433, National Bureau of Economic Research.
- CHIAPPORI, P.-A. AND B. SALANIE (2000): “Testing for asymmetric information in insurance markets,” *Journal of political Economy*, 108, 56–78.

- CHING, A. T. AND M. ISHIHARA (2012): “Measuring the informative and persuasive roles of detailing on prescribing decisions,” *Management Science*, 58, 1374–1387.
- CLARK, C. R., U. DORASZELSKI, AND M. DRAGANSKA (2009): “The effect of advertising on brand awareness and perceived quality: An empirical investigation using panel data,” *Quantitative Marketing and Economics*, 7, 207–236.
- CLEMENS, J. (2013): “Regulatory Redistribution in the Market for Health Insurance,” Working Paper, UCSD.
- COHEN, A. AND L. EINAV (2007): “Estimating Risk Preferences from Deductible Choice,” *American Economic Review*, 97, 745–788.
- COLE, H. L., S. KIM, AND D. KRUEGER (2012): “Analyzing the Effects of Insuring Health Risks: On the Trade-off between Short Run Insurance Benefits vs. Long Run Incentive Costs,” Working Paper 18572, National Bureau of Economic Research.
- CONESA, J. C., S. KITAO, AND D. KRUEGER (2009): “Taxing Capital? Not a Bad Idea after All!” *American Economic Review*, 99, 25–48.
- COURTEMANCHE, C. J. AND D. ZAPATA (2012): “Does Universal Coverage Improve Health? The Massachusetts Experience,” NBER Working Paper No. 17893.
- CURRIE, J. AND B. C. MADRIAN (1999): “Health, health insurance and the labor market,” in *Handbook of Labor Economics*, ed. by O. C. Ashenfelter and D. Card, Elsevier, vol. 3, Part C, 3309–3416.
- DAFNY, L. AND D. DRANOVE (2008): “Do report cards tell consumers anything they don’t already know? The case of Medicare HMOs,” *The Rand journal of economics*, 39, 790–821.
- DEY, M. S. AND C. J. FLINN (2005): “An Equilibrium Model of Health Insurance Provision and Wage Determination,” *Econometrica*, 73, 571–627.

- DIZIOLI, A. AND R. B. PINHEIRO (2012): “Health Insurance as a Productive Factor,” Working Paper, University of Colorado at Boulder.
- ECKSTEIN, Z. AND K. I. WOLPIN (1990): “Estimating a Market Equilibrium Search Model from Panel Data on Individuals,” *Econometrica*, 58, 783–808.
- EINAV, L. AND A. FINKELSTEIN (2011): “Selection in Insurance Markets: Theory and Empirics in Pictures,” *Journal of Economic Perspectives*, 25, 115–38.
- EINAV, L., A. FINKELSTEIN, AND M. R. CULLEN (2010a): “Estimating welfare in insurance markets using variation in prices,” *The Quarterly Journal of Economics*, 125, 877–921.
- EINAV, L., A. FINKELSTEIN, AND J. LEVIN (2010b): “Beyond Testing: Empirical Models of Insurance Markets,” *Annual Review of Economics*, 2, 311–336.
- EINAV, L., A. FINKELSTEIN, S. RYAN, P. SCHRIMPF, AND M. CULLEN (2013): “Selection on Moral Hazard in Health Insurance,” *American Economic Review*, 103.
- EINAV, L., A. FINKELSTEIN, AND P. SCHRIMPF (2010c): “Optimal Mandates and the Welfare Cost of Asymmetric Information: Evidence From the U.K. Annuity Market,” *Econometrica*, 78, 1031–1092.
- ERICSON, K. M. M. AND A. STARC (2012): “Pricing Regulation and Imperfect Competition on the Massachusetts Health Insurance Exchange,” Working Paper 18089, National Bureau of Economic Research.
- FANG, H. AND A. GAVAZZA (2011): “Dynamic Inefficiencies in an Employer-Provided Health Insurance System: Theory and Evidence,” *American Economic Review*, 101, 3047–3077.
- FANG, H., M. P. KEANE, AND D. SILVERMAN (2008): “Sources of Advantageous Se-

- lection: Evidence from the Medigap Insurance Market,” *Journal of Political Economy*, 116, 303–350.
- FARHI, E. AND I. WERNING (2013): “Insurance and Taxation over the Life Cycle,” *The Review of Economic Studies*, 80, 596–635.
- FIJOLEK, B. (2012): “Solidarity or Personal Responsibility? A Look at the Lessons Switzerlands Health Care System Can Teach the United States,” *Annals of Health Law*, 21, 1–13.
- FINKELSTEIN, A. AND K. MCGARRY (2006): “Multiple dimensions of private information: evidence from the long-term care insurance market,” *American Economic Review*, 96, 938–958.
- FINKELSTEIN, A., S. TAUBMAN, B. WRIGHT, M. BERNSTEIN, J. GRUBER, J. P. NEWHOUSE, H. ALLEN, K. BAICKER, AND THE OREGON HEALTH STUDY GROUP (2012): “The Oregon Health Insurance Experiment: Evidence from the first year,” *Quarterly Journal of Economics*, 127, 1057 – 1106.
- FLINN, C. AND J. J. HECKMAN (1982): “New Methods For Analyzing Structural Models Of Labor Force Dynamics,” *Journal of Econometrics*, 18, 115 – 168.
- FRENCH, E. AND J. B. JONES (2011): “The Effects of Health Insurance and Self-Insurance on Retirement Behavior,” *Econometrica*, 79, 693–732.
- GARICANO, L., C. LELARGE, AND J. VAN REENEN (2013): “Firm Size Distortions and the Productivity Distribution: Evidence from France,” Working Paper 18841, National Bureau of Economic Research.
- GARTHWAITE, C., T. GROSS, AND M. J. NOTOWIDIGDO (2013): “Public Health Insurance, Labor Supply, and Employment Lock,” Working Paper 19220, National Bureau of Economic Research.

- GERUSO, M. (2012): “Selection in Employer Health Plans: Homogeneous Prices and Heterogeneous Preferences,” Working Paper, Princeton University.
- GOEREE, M. S. (2008): “Limited information and advertising in the US personal computer industry,” *Econometrica*, 76, 1017–1074.
- GORDON, B. R. AND W. R. HARTMANN (2013): “Advertising effects in presidential elections,” *Marketing Science*, 32, 19–35.
- GOURIEROUX, C., A. MONFORT, AND E. RENAULT (1993): “Indirect inference,” *Journal of Applied Econometrics*, 8, S85–S118.
- GOURIO, F. AND N. A. ROYS (2012): “Size-Dependent Regulations, Firm Size Distribution, and Reallocation,” Working Paper 18657, National Bureau of Economic Research.
- GRUBER, J. (2000): “Health insurance and the labor market,” Elsevier, vol. 1, Part A of *Handbook of Health Economics*, 645 – 706.
- HACKMANN, M., J. KOLSTAD, AND A. KOWALSKI (2013): “Adverse Selection and an Individual Mandate: When Theory Meets Practice,” Worker Paper, University of Pennsylvania.
- HAI, R. (2013): “The Determinants of Rising Inequality in Health Insurance and Wages: An Equilibrium Model of Workers Compensation and Health Care Policies,” Working Paper, University of Pennsylvania.
- HALL, A. (2007): “The value of Medicare managed care plans and their prescription drug benefits,” .
- HANDEL, B. R., I. HENDEL, AND M. D. WHINSTON (2013): “Equilibria in Health Exchanges: Adverse Selection vs. Reclassification Risk,” Working Paper 19399, National Bureau of Economic Research.

- HANSEN, G. D., M. HSU, AND J. LEE (2012): “Health Insurance Reform: The Impact of a Medicare Buy-In,” Working Paper 18529, National Bureau of Economic Research.
- HASTINGS, J. S., A. HORTAÇSU, AND C. SYVERSON (2013): “Advertising and competition in privatized social security: the case of Mexico,” .
- HAUSMAN, J. A. (1996): “Valuation of new goods under perfect and imperfect competition,” in *The economics of new goods*, University of Chicago Press, 207–248.
- HENDREN, N. (2013): “Private Information and Insurance Rejections,” *Econometrica*, 81, 1713–1762.
- HO, K. (2009): “Insurer-Provider Networks in the Medical Care Market,” *The American Economic Review*, 393–430.
- IMBENS, G. W. AND T. LANCASTER (1994): “Combining Micro and Macro Data in Microeconomic Models,” *Review of Economic Studies*, 61 (4), 655–680.
- JOLIVET, G., F. POSTEL-VINAY, AND J.-M. ROBIN (2006): “The Empirical Content of the Job Search Model: Labor Mobility and Wage Distributions in Europe and the US,” *European Economic Review*, 50 (4), 877–907.
- KAISER FAMILY FOUNDATION AND HEALTH RESEARCH AND EDUCATIONAL TRUST (2009): “Employer Health Benefits, 2009 Annual Survey,” <http://ehbs.kff.org/2009.html>.
- KAPLAN, G. (2012): “Inequality and the Life Cycle,” *Quantitative Economics*, 3, 471–525.
- KEANE, M. P. AND K. I. WOLPIN (1994): “The Solution and Estimation of Discrete Choice Dynamic Programming Models by Simulation and Interpolation: Monte Carlo Evidence,” *The Review of Economics and Statistics*, 76 (4), 648–672.

- KENNAN, J. (2006): “A Note on Discrete Approximations of Continuous Distributions,” University of Wisconsin-Madison and NBER.
- KHWAJA, A. (2010): “Estimating Willingness to Pay for Medicare Using a Dynamic Life-cycle model of Demand for Health Insurance,” *Journal of Econometrics*, 156(1), 130–147.
- KOLSTAD, J. T. AND A. E. KOWALSKI (2012a): “Health Reform, Health Insurance, and Selection: Estimating Selection into Health Insurance Using the Massachusetts Health Reform,” *American Economic Review Papers and Proceedings*.
- (2012b): “The Impact of Health Care Reform On Hospital and Preventive Care: Evidence from Massachusetts,” *Journal of Public Economics*, forthcoming.
- (2012c): “Mandate-Based Health Reform and the Labor Market: Evidence from the Massachusetts Reform,” Working paper, University of Pennsylvania.
- KUZIEMKO, I., K. MECKEL, AND M. ROSSIN-SLATER (2013): “Do Insurers Risk-Select Against Each Other? Evidence from Medicaid and Implications for Health Reform,” *NBER Working Paper*.
- LANG, K. (2010): “Worker Sorting, Taxes and Health Insurance Coverage,” Working Paper, Boston University.
- LANGWELL, K. M. AND J. P. HADLEY (1989): “Evaluation of the Medicare competition demonstrations.” *Health Care Financing Review*, 11, 65.
- LOW, H., C. MEGHIR, AND L. PISTAFERRI (2010): “Wage Risk and Employment Risk over the Life Cycle,” *American Economic Review*, 100, 1432–67.
- LUSTIG, J. (2011): “Measuring welfare losses from adverse selection and imperfect competition in privatized medicare,” *Manuscript*.

- MACURDY, T., D. GREEN, AND H. PAARSCH (1990): “Assessing Empirical Approaches for Analyzing Taxes and Labor Supply,” *The Journal of Human Resources*, 25, 415–490.
- MEGHIR, C., R. NARITA, AND J.-M. ROBIN (2012): “Wages and Informality in Developing Countries,” NBER Working Paper No. 18347.
- MEHROTRA, A., S. GRIER, AND R. A. DUDLEY (2006): “The relationship between health plan advertising and market incentives: Evidence of risk-selective behavior,” *Health Affairs*, 25, 759–765.
- MELLO, M. M., S. C. STEARNS, E. C. NORTON, AND T. C. RICKETTS (2003): “Understanding biased selection in Medicare HMOs,” *Health services research*, 38, 961–992.
- MICHELACCI, C. AND H. RUFFO (2011): “Optimal Life Cycle Unemployment Insurance,” Working Paper, CEMFI.
- MILGROM, P. AND J. ROBERTS (1986): “Price and advertising signals of product quality,” *The Journal of Political Economy*, 796–821.
- MOSCARINI, G. AND F. POSTEL-VINAY (2012): “Stochastic Search Equilibrium,” Working Paper, Yale University.
- MULLIGAN, C. B. (2013a): “Average Marginal Labor Income Tax Rates under the Affordable Care Act,” Working Paper 19365, National Bureau of Economic Research.
- (2013b): “Is the Affordable Care Act Different from Romneycare? A Labor Economics Perspective,” Working Paper 19366, National Bureau of Economic Research.
- NATIONAL COMMISSION ON FISCAL RESPONSIBILITY AND REFORM (2010): “The Moment of Truth: Report of the National Commission on Fiscal Responsibility and Reform,” .

- NELSON, P. (1974): “Advertising as information,” *The Journal of Political Economy*, 82, 729–754.
- NEUMAN, P., E. MAIBACH, K. DUSENBURY, M. KITCHMAN, AND P. ZUPP (1998): “Marketing HMOs to Medicare beneficiaries,” *Health Affairs*, 17, 132–139.
- NEVO, A. (2001): “Measuring market power in the ready-to-eat cereal industry,” *Econometrica*, 69, 307–342.
- NOSAL, K. (2012): “Estimating switching costs for medicare advantage plans,” *Unpublished manuscript, University of Mannheim*.
- OZKAN, S. (2011): “Income Differences and Health Care Expenditures over the Life Cycle,” Working Paper, Federal Reserve Board.
- PAKES, A., J. PORTER, K. HO, AND J. ISHII (2011): “Moment inequalities and their application,” .
- PAPAGEORGE, N. W. (2012): “How Does Medical Innovation Create Value? Health, Human Capital and the Labor Market,” Working Paper, Johns Hopkins University.
- PASHCHENKO, S. AND P. PORAPAKKARM (2013): “Quantitative analysis of health insurance reform: Separating regulation from redistribution,” *Review of Economic Dynamics*, 16, 383 – 404.
- PETRIN, A. (2002): “Quantifying the Benefits of New Products: The Case of the Minivan,” *Journal of Political Economy*, 110, pp. 705–729.
- POHL, V. (2012): “Medicaid and the Labor Supply of Single Mothers: Implications for Health Care Reform,” Working Paper, Yale University.
- POPE, G. C., J. KAUTTER, R. P. ELLIS, A. S. ASH, J. Z. AYANIAN, M. J. INGBER, J. M. LEVY, J. ROBST, ET AL. (2004): “Risk adjustment of Medicare capitation payments using the CMS-HCC model,” .

- POSTEL-VINAY, F. AND J.-M. ROBIN (2002): “Equilibrium Wage Dispersion with Worker and Employer Heterogeneity,” *Econometrica*, 70, 2295–2350.
- ROBIN, J.-M. AND S. ROUX (2002): “An Equilibrium Model of the Labour Market with Endogenous Capital and Two-Sided Search,” *Annales d’économie et de Statistique*, 67/68, 257–308.
- RUST, J. AND C. PHELAN (1997): “How Social Security and Medicare Affect Retirement Behavior in a World of Incomplete Markets,” *Econometrica*, 65, 781–831.
- SELDEN, T. AND B. M. GRAY (2006): “Tax Subsidies for Employer-Related Health Insurance: Estimates for 2006,” *Health Affairs*, 25, 1568–1579.
- SHEPHARD, A. (2012): “Equilibrium Search and Tax Credit Reform,” Working Paper, University of Pennsylvania.
- STARC, A. (2012): “Insurer pricing and consumer welfare: Evidence from medigap,” *Manuscript*.
- STIGLER, G. J. AND G. S. BECKER (1977): “De gustibus non est disputandum,” *The American Economic Review*, 67, 76–90.
- TOWN, R. AND S. LIU (2003): “The welfare impact of Medicare HMOs,” *RAND Journal of Economics*, 719–736.
- VAN DE VEN, W. P. AND R. P. ELLIS (2000): “Chapter 14 Risk adjustment in competitive health plan markets,” Elsevier, vol. 1, Part A of *Handbook of Health Economics*, 755 – 845.
- VAN DEN BERG, G. J. AND G. RIDDER (1998): “An Empirical Equilibrium Search Model of the Labor Market,” *Econometrica*, 66, 1183–1221.
- WEINZIERL, M. (2011): “The Surprising Power of Age-Dependent Taxes,” *The Review of Economic Studies*, 78, 1490–1518.

WOLPIN, K. I. (1992): “The Determinants of Black-White Differences in Early Employment Careers: Search, Layoffs, Quits, and Endogenous Wage Growth,” *Journal of Political Economy*, 100, 535–560.

YAMAGUCHI, S. (2010): “Job Search, Bargaining, and Wage Dynamics,” *Journal of Labor Economics*, 28, 595–631.