THREE ESSAYS IN HEALTH INSURANCE COVERAGE

by

Matthew S. Rutledge

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Economics)
in The University of Michigan
2010

Doctoral Committee:

Associate Professor Daniel S. Silverman, Chair
Professor Charles C. Brown
Professor Thomas C. Buchmueller
Professor Jeffrey A. Smith

To Kara and Jack

ACKNOWLEDGEMENTS

This dissertation would not be possible without the guidance and mentorship of my committee, Dan Silverman, Charlie Brown, Tom Buchmueller, and Jeff Smith. I greatly appreciate their help, patience, and availability throughout the last few years.

Perhaps no issue has received more attention in public policy circles over the last decade than health insurance, and I would not be a part of that conversation if I hadn't, nearly immediately after arriving in Ann Arbor, encountered the Economic Research Initiative on the Uninsured, a Robert Wood Johnson Foundation-sponsored research group. Hanns Kuttner, Mary Harrington Naud, Paula Song, Karoline Mortensen, Wendy Zellers, Jacqui Hinchey, Chris Jacobs and, especially, Helen Levy were instrumental in providing me an education in the institutions and existing research on health care and health insurance, and I am grateful for their encouragement and feedback. And I doubt there exists a better mentor, in health economics or any social science, than Catherine McLaughlin, my coauthor on the second chapter of this volume; for taking a chance on an unfunded first-year grad student interested in macro public finance, introducing me not only to health insurance research but also the very people doing that interesting work, and providing invaluable advice on work and life, I am forever in her debt.

So much of what I learned about health economics came from my experience as a Graduate Student Instructor in the Health Management and Policy Department of the Michigan School of Public Health. I would like to thank Rich Hirth, Edward Norton, Daniel Eisenberg, Catherine McLaughlin, Omar Alshanqeety, and Edward Okeke for immersing me in health institutions and research. I am also grateful to the HMP Classes of 2009 and 2010 for their friendliness, interesting discussions, and enthusiasm; I got as much, if not more, out of our time together as they did.

It is no exaggeration to say that I wouldn't have survived my coursework, let alone finish this dissertation, without the help of my classmates and friends, who provided so much advice and assistance to these chapters directly and indirectly: Zoë McLaren, Brian Rowe, Osborne Jackson, Shanthi Ramnath, Stephan Lindner, Mike Gough, Doug Smith, Ken Kashiwase, Owen Kearney, Ryan Michaels, Brian Kovak, Angus Chu, Andres Cooper, Laura Kawano, Jon Lanning, Brooke Helppie, Marta Murray-Close, Joanne Hsu, Gabriela Contreras, Daniel Cooper, Kate Anderson, Josh Congdon-Hohman, Nick Montgomery, Ben Keys, Brian Cadena, Alex Resch, Jody Schimmel, Josh Cherry, Dmitry Lubensky, Todd Pugatch, Jesse Gregory, and Brad Hershbein. I am lucky that I encountered Luke Shaefer even before his arrival at UM School of Social Work; though our joint work is not a part of this volume, I am excited about where it is going. Thanks also to Christine Cheu, Paul Kreiter, Curt McFall, Cara Lewis, Nick Emptage, Edward Czilli and everyone at the ICPSR Summer Program, Clint Carter and Maggie Levenstein at the Michigan Census Research Data Center, Ray Kuntz and the group at AHRQ, Mary Braun and Ronda Abrigo from the Economics Department, the Michigan Econ softball, flag football, and broomball teams, and everyone else who has helped me the last six years, either personally or professionally.

I recently became a father, so I am only now realizing all that goes into raising a child from zygote to cute and helpless infant to pride-worthy adult. Thanks and love to my parents, Yvette Rutledge and Steve and Kathy Rutledge, for all that they have done and continue to do. Thanks to my in-laws, Doug and Suzanne Fitzgerald, not just for their daughter, but also for their support, especially over the last year. I would also like to thank my siblings and siblings-in-law, Kate, Rachael, Meaghan, Kelly, Vinnie, and Dean, and all of the other members of the family, and our pseudo-family from Auburn, BC, and the Boston Fed, that we got to see all-too-sparingly the last six years; I know we'll make up for lost time soon.

I just met Jackson Fitzgerald Rutledge on December 16, 2009, and yet I feel like no one has had as much of an impact on my time at Michigan as he has. I can't imagine another grad student being happier to wake up at 6 a.m. every morning, but then again, most grad students don't get to see Jack's smile first thing. Big thanks and lots of love, buddy, and I look forward to watching you grow up into whatever you want to be — as long as you don't become an economist.

My most profound thanks to my beautiful wife, best friend, and partner in everything, Kara. She selflessly followed me 800 miles, leaving behind family, friends, and jobs, dealt with late nights and working weekends and mood swings and confidence deficits and, most impressively, the constant flow of economics jargon into casual conversation. I know the last six years are but a drop in the bucket of our life together, but I am eternally grateful for the sacrifice, support, affection, and inspiration.

TABLE OF CONTENTS

DEDICATION ii

ACKNOWLEDGEMENTS iii

LIST OF FIGURES viii

LIST OF TABLES ix

CHAPTER

I. Introduction 1

II. Asymmetric Information and the Generosity of Employer-Sponsored Health Insurance 5

2.1 Introduction 5
2.2 Related Literature 8
2.3 Data 11
2.4 Theoretical Framework 12
2.5 Health Insurance Generosity 18
2.5.1 Indications of Adverse Selection 21
2.6 Empirical Strategy 25
2.6.1 Estimation Sample 26
2.6.2 Specification 27
2.6.3 Weights 30
2.7 Results 31
2.7.1 Expenditures 31
2.7.2 Utilization 35
2.7.3 Direct Measures of Selection 35
2.7.4 Robustness Checks 38
2.8 Conclusion 40
2.9 Appendix: Proof That Expected Value of Expenditures Increases with the Threshold Spending Value h'. 46
2.10 Figures and Tables 48

III. Decomposing the Growing Disparity in Health Insurance Coverage between Hispanics and Non-Hispanic Whites 59

3.1 Introduction 59
3.2 Previous Literature 62
3.3 Data 65

3.4 Trends in the Uninsured Rate . . . 67
3.5 Methods . . . 73
3.5.1 Method 1 - Regression-Adjusted Trend Line . . . 74
3.5.2 Method 2 - Pairwise Blinder-Oaxaca Change Decompositions . . . 77
3.6 Results . . . 80
3.6.1 Regression Adjusted Trend Line . . . 80
3.6.2 Blinder-Oaxaca Decompositions . . . 84
3.6.3 Decompositions by Insurance Source . . . 86
3.6.4 Decompositions by Hispanic Subgroup . . . 88
3.7 Discussion . . . 89
3.8 Conclusions . . . 95
3.9 Appendix: Inverse Probability Weighting for Time Series Graphs . . . 98
3.10 Figures and Tables . . . 100

IV. Another Moment to Consider: Health Plan Choice and the Distribution of Out-Of-Pocket Medical Spending . . . 118

4.1 Introduction . . . 118
4.2 Related Literature . . . 121
4.3 Theoretical Framework . . . 124
4.4 Data . . . 128
4.5 Empirical Implementation . . . 130
4.6 Results . . . 136
4.7 Conclusion . . . 141
4.8 Figures and Tables . . . 146

V. Conclusion . . . 158

BIBLIOGRAPHY . . . **161**

LIST OF FIGURES

Figure

2.1 Risk Distribution and Generosity at Two Hypothetical Firms With the Same Less-Generous Plan (C) . . . 48
2.2 Risk Distribution and Generosity at Two Hypothetical Firms With the Same More-Generous Plan (G) . . . 49
2.3 Total Expenditure by Threshold Spending Level (h') Decile and Plan Type . . . 49
2.4 Probability of Choosing the More Generous Plan, By Total Expenditure . . . 50
2.5 Total Expenditure by Absolute Generosity Decile and Plan Type . . . 50
2.6 Distribution of Total Expenditures, by Plan Type . . . 50
2.7 Average Total Medical Expenditure by Plan Month and Calendar Month . . . 51
3.1 Uninsured Rate By Ethnicity, Age 17 and Under . . . 100
3.2 Uninsured Rate By Ethnicity, Age 55 and Over . . . 100
3.3 Uninsured Rate by Racial and Ethnic Group, Age 18 to 55 . . . 101
3.4 Uninsured Rate by Hispanic Origin and Citizenship, Age 18 to 55 . . . 101
3.5 Uninsured Rate by Educational Attainment, Age 18 to 55 . . . 102
3.6 Uninsured Rate by Family Income as a Percent of the Federal Poverty Level, Age 18 to 55 . . . 103
3.7 Uninsured Rate by Age, Age 18 to 55 . . . 104
3.8 Uninsured Rate by Hispanic Subgroup and Citizenship, Age 18 to 55 . . . 105
3.9 Public Coverage Rate By Racial and Ethnic Group, Age 18 to 55 . . . 105
3.10 Employer-Sponsored Coverage in Own's Name Rate By Racial and Ethnic Group, Age 18 to 55 . . . 106
3.11 Hispanic - White Non-Hispanic Gap in Uninsured Rate . . . 107
3.12 Hispanic - White Non-Hispanic Gap in Public Coverage Rate . . . 107
3.13 Hispanic - White Non-Hispanic Gap in Private Coverage Rate . . . 108
3.14 Hispanic - White Non-Hispanic Gap in Rate of Own Employer-Sponsored Coverage 109
3.15 Hispanic - White Non-Hispanic Gap in Rate of Own Private non-ESI Coverage . . 109
3.16 Hispanic - White Non-Hispanic Gap in Rate of Private Coverage From Someone Else 110
4.1 Distribution of Differences between Selected and Non-Selected Plans . . . 147
4.2 Probability of Selecting Plan, By Decile of Difference with Most Generous (Lowest Co-Premium) Plan . . . 148

LIST OF TABLES

Table

2.1 Non-Financial Generosity Differences Between Plan Pairs . . . 51
2.2 Employee Contributions, Premiums, and Choosing the More Generous Plan . . . 52
2.3 Summary Statistics . . . 52
2.4 Selected Summary Statistics for Weighted Full Sample and Estimation Sample, Unweighted and Weighted . . . 53
2.5 Two-Part Regression of Total Expenditures on Absolute Generosity and Log Threshold Spending Level . . . 54
2.6 Two-Part Regression of Expenditures on Generosity and Premium Measures, by Type of Care . . . 54
2.7 Two-Part Regression of Total Expenditures on Absolute Generosity and Log Threshold Spending Level, By Employee Contribution Scheme and h' Level . . . 55
2.8 Poisson Regression of Visits on Absolute Generosity and Log Threshold Spending Level . . . 55
2.9 OLS Regressions of Numbers of Family Members with an Adverse Health Outcome 56
2.10 Two-Part Regression of Total Expenditures on Absolute Generosity and Log Threshold Spending Level, Offered Exactly Two Plans . . . 57
2.11 Two-Part Regression of Total Expenditures in the Middle Eight Months on Absolute Generosity and Log Threshold Spending Level . . . 57
2.12 Two-Part Regression of Total Expenditures on Absolute Generosity and Log Threshold Spending Level, by Coverage and Marital Status . . . 58
2.13 Two-Part Regression of Total Expenditures on Absolute Generosity and Log Threshold Spending Level, by Whether the Predicted Plan was Selected . . . 58
3.1 Survey of Income and Program Participation (SIPP) Panels . . . 111
3.2 Sample Means . . . 111
3.3 Predicted Hispanic - White Non-Hispanic Gaps in Uninsured Rate (Method 1) . . 113
3.4 Linear Time Trends in the Uninsured Rate Gap between Hispanics and Non-Hispanic Whites . . . 114
3.5 Blinder-Oaxaca Decompositions for the Change in the Hispanic-White Non-Hispanic Uninsured Rate Gap . . . 114
3.6 Blinder-Oaxaca Decompositions for Annual Hispanic-White Non-Hispanic Uninsured Rate Gap . . . 115
3.7 Blinder-Oaxaca Decompositions for the Change in Gap with Non-Hispanic Whites 116
3.8 Blinder-Oaxaca Decompositions for the Change in Gap with Mexicans . . . 117
4.1 Plan Characteristics by Out-of-Pocket Mean and Variance Quartile . . . 149
4.2 Employee Premium Sharing Arrangements, By Coverage Type . . . 149
4.3 Summary Statistics for Unweighted and Weighted Samples . . . 150
4.4 Variation in Non-Financial Plan Elements between Selected Plan and Each Non-Selected Offered Plan . . . 151
4.5 Mean Derivatives from Nested Logit Regression of Plan Choice, Single Coverage Only . . . 152

4.6 Mean Derivatives from Nested Logit Regression of Plan Choice, Single Coverage Only, Splitting The Second Moment . . . 153
4.7 Mean Derivatives from Nested Logit Regression of Plan Choice, Single Coverage Only, Different Premium for Each Plan . . . 154
4.8 Mean Derivatives from Nested Logit Regression of Plan Choice, Single Coverage Only, Worst Plan Has Lowest Premium . . . 155
4.9 Mean Derivatives from Nested Logit Regression of Plan Choice, Family Coverage Only with No Working Spouse . . . 156
4.10 Mean Derivatives from Conditional Logit Regression of Plan Choice, Single Coverage Only . . . 157

CHAPTER I

Introduction

The health insurance reform debate, which resulted in the passage of the Patient Protection and Affordable Care Act and the Health Care and Education Reconciliation Act, both signed into law in late March, 2010, had two general goals. One was extending coverage to millions of uninsured Americans. The other was controlling the growth of health care costs, with an aim to making health care more affordable and making health insurance more protective against catastrophic financial losses associated with illness or injury. With these dual goals, policymakers focused on both the extensive margin of health insurance coverage — moving the uninsured into any form of coverage — and the intensive margin — how much coverage one has, and how insurance influences the medical spending of the insured.

This dissertation consists of three distinct essays related to both aspects of the health insurance policy debate; two focus on the intensive margin of health insurance coverage, and one examines the extensive margin. The three essays cover some of the most relevant issues in health insurance coverage, including whether more generous coverage leads to more spending due to moral hazard, how different are the enrollees in more and less generous plans (i.e., adverse selection) and how that might influence insurance premiums, how interracial disparities in health insurance coverage have changed over time and why, and how responsive consumers are to particular plan characteristics, including the level and variability of the consumer's projected burden of out-of-pocket medical spending under each plan.

In the first essay, I devise a new test for whether adverse selection and/or moral hazard influence the national market for employer-sponsored health insurance. Before enrollment, employees likely know much more about their health needs in the coming year than the insurer, which in most cases must charge the same premium to every employee for any particular health plan, regardless of their past, present, or future utilization. In addition, the employee and her care provider likely know more than the insurer about the necessity of each unit of medical care received during the coverage year. The employees' information advantage before and during enrollment threatens to induce inefficiency due to adverse selection and moral hazard, respectively. While these effects are well understood theoretically, in practice they are difficult to distinguish, as each results in enrollees in the more generous plan spending more on medical care. I propose a model that produces two empirical predictions to separately identify these effects. First, a more financially generous health plan, irrespective of the rejected alternatives, induces increased spending due to moral hazard. Second, controlling for the selected plan's generosity, a positive correlation between expenditures and the level of spending that leads employees to prefer the more generous plan suggests adverse selection. Then, using data from the Medical Expenditure Panel Survey from 1996 to 2001 on health plan offerings from a national sample of employers, I confirm that more generous coverage leads to increased spending on medical care due to moral hazard, although the effect is weaker in managed care plans than in traditional, non-managed plans. While there is some evidence that individuals with a medical condition enroll in more generous plans, spending in those plans is not significantly higher, suggesting that adverse selection does not lead to diverging premiums.

In the second essay, written with Catherine G. McLaughlin of Mathematica Policy Research and the University of Michigan School of Public Health, we decompose how differences in citizenship, education, and labor market outcomes have contributed to the growing disparity in the uninsured rate between Hispanics and non-Hispanic Whites. Over the last 25 years, the uninsured rate has risen dramatically for Hispanics, while the proportion of most other racial and ethnic groups without health insurance coverage has been fairly stable. Hispanics, on average, are less likely to be citizens and have lower educational attainment and worse labor market outcomes, and there is evidence that these differences between Hispanics and non-Hispanics have grown over time. Using pooled panels of the Survey of Income and Program Participation from 1983 to 2007, we find that differences in citizenship and education can explain about half of the growth in the uninsured rate gap between Hispanics and non-Hispanic Whites, but even after controlling for other observable characteristics, more than half of the growing disparity, or more than one million extra uninsured Hispanics, remains unexplained.

In the final essay, I examine the relative influence of the elements of a health plan's "price" on the probability that it is selected from the set of plans offered by one's employer. In health insurance, price includes not just the upfront cost to the consumer, the premium, but also both the level and the variance out-of-pocket spending associated with the care the plan insures. In addition, I consider how plan characteristics beyond financial coverage for basic care, including access to providers, non-standard coverage like mental health and long term care, and plan's market share among one's coworkers, affect plan choice. Using the same data as the first essay and accounting for the differences between managed care and less restrictive plans, I find that employees demand a plan with lower expected out-of-pocket costs, but that the variance of a plan's out-of-pocket spending has no effect on plan choice. Employees are only weakly responsive to their share of the premium, but do tend to choose the most popular plan at a firm, suggesting the importance of unobserved plan quality differences and/or the employees opting for the default plan.

CHAPTER II

Asymmetric Information and the Generosity of Employer-Sponsored Health Insurance

DISCLAIMER: The research in this paper was conducted while the author was a Special Sworn Status researcher of the U.S. Census Bureau at the Michigan Census Research Data Center. Any opinions and conclusions expressed herein are those of the author and do not necessarily represent the views of the U.S. Census Bureau. All results have been reviewed to ensure that no confidential information is disclosed.

2.1 Introduction

Seventy-five percent of Americans under the age of 65 obtain their health insurance through their (or a family member's) employer (DeNavas-Walt et al., 2009). The main advantage of the employer insurance system is the ability to pool risk over a group that gathers for reasons that likely have little to do with health. However, there may be no insurance market where information is more asymmetric between buyer and seller. Because of nondiscrimination laws, privacy concerns, and administrative difficulty, health insurers must set plan premiums without knowing any individual employee's family health history, recent medical care utilization, or often even basic demographic information like gender, age, or family size. Usually, the most the insurer knows about any one individual in a firm is whether it is covering her alone or her and some unspecified number of family members.

A consequence of asymmetric information is that moral hazard and adverse selection can unravel the efficiency gains from selling insurance to an employee group rather than individual by individual. Because the insurer cannot properly monitor which units of care are medically necessary, more generous coverage may lead to enrollees purchasing care that the consumer values below their full cost, reducing social welfare. The inflated health care spending due to moral hazard drives up insurers' costs and makes health insurance more expensive for all.

In addition, employees with private information about their high expected health needs have an incentive to choose more generous plans, driving up the premium for these plans and forcing healthier employees into plans that provide less protection against financial risk. Rothschild and Stiglitz (1976) show that this self-selection may lead to the adverse selection "death spiral," where a plan that enrolls only the highest risk employees becomes too expensive for the employer to offer. This results in a welfare loss to both high- and low-risk enrollees: there are fewer insurance options available, and those plans that do exist have less generous coverage than enrollees would prefer.

In this paper, I take advantage of variation across employers in the generosity of plans in the employees' choice set, and information about both accepted and rejected plans, to separately identify the influence of moral hazard and adverse selection. The separate estimation of these two phenomena stands in contrast to previous studies of other insurance markets that are restricted to a joint test of adverse selection and moral hazard (Chiappori and Salanie, 2000).

I model the employee's decision to enroll in the more generous plan offered to her by her employer, and two empirical predictions emerge. First, a more generous plan leads to more spending on medical care due to moral hazard; the employee's spending decision is independent of her plan's premium and the premium and generosity of other plans the employee had been offered. Second, an employee's incentive to choose higher- or lower-generosity plans depends on the difference in out-of-pocket spending between her offers. Low spending employees prefer the less generous plan with the lower premium, but above some (firm-specific) spending threshold, the more generous plan is preferable. If adverse selection occurs, as the threshold spending level increases, enrollees in both the more and less generous plans should be less healthy and higher-spending on average.

This model leads to clear testable implications for employer-sponsored health insurance, the broadest insurance market in the U.S. I produce the first nationwide estimates of the impact of moral hazard and adverse selection between plans of the same type, either health maintenance organizations (HMOs) or more traditional, less-managed care, on this market. Like existing studies on a small number of firms (including Feldman et al., 1989, and Cutler and Reber, 1998), I consider how differences in direct health measures suggest adverse selection between plans of the same type. Unlike these studies, I also estimate how differences in generosity and premiums correlate with the difference in expenditures between plan enrollees, which is more likely to lead to the adverse selection "death spiral."

My estimates, using Medical Expenditure Panel Survey data from 1996 to 2001, suggest that moral hazard leads to substantially more spending on medical care in non-HMO plans, though spending does not increase with generosity in an HMO plan. There is no evidence for adverse selection between plans of the same type; an increase in the threshold spending level does not lead to an increase in observed spending nor more adverse health outcomes among enrollees. There is weak evidence, that

HMO plans are less likely than non-HMOs to enroll employees with family members suffering from a chronic condition or over the age of 50, consistent with adverse selection between plan types.

These results are relevant to at least two aspects of the current health insurance debate. One important goal of reform is to reduce the growth of health care costs; the lower level of moral hazard-induced spending among HMO enrollees suggests that tighter management of care among all plans may reduce the ever-growing spending on health care in the U.S. economy. Officials also aim to reduce the often exorbitant premiums in the small business and individual insurance markets, restricting premiums based on personal characteristics, including age, eliminating coverage exemptions for pre-existing conditions, and generally making these markets more like the existing employer insurance system. While adverse selection resulting from this increase in private information could lead to rapidly increasing premiums or the elimination of the most generous plans, my estimates suggest that this outcome is unlikely.

2.2 Related Literature

Economists have long theorized that insurance contracts that are improperly priced due to private information can lead to spiraling prices, unequal access, and market failure (Arrow, 1963). Because efforts to prevent illness or injury, or limit the amount spent on the patient beyond what is medically necessary, are difficult if not impossible to monitor, insurers include deductibles and coinsurance to reduce moral hazard, rather than fully cover medical care (Pauly, 1968). Furthermore, Rothschild and Stiglitz (1976) show that when insurers are unable to differentiate between high- and low-risk enrollees, insurance markets are prone to adverse selection: in equilibrium, either low-risk individuals are unable to fulfill their desire to fully insure, or no insurance contract is offered to either group. The market is thus forced to decide on the proper tradeoff between protecting against the financial risk of medical care episodes and moderating inefficiency.

The health insurance literature has generally found that moral hazard has a moderate but statistically significant impact on medical spending (Cutler and Zeckhauser, 2000). This consensus emerged after the RAND experiment (Manning et al., 1987), which featured an experimental design that avoided corrupting the moral hazard estimates with selection effects. The RAND experiment's conclusion was that consumers are responsive to out-of-pocket price, with a small but statistically significant elasticity (approximately -0.2).

While there has never been an analogous large-scale experiment for determining the impact of adverse selection, researchers have frequently used employer data on the composition of enrollees in their health plans to determine whether employees select a health plan based on private information about their health. Cutler and Zeckhauser (2000) survey this literature, and find that in most studies of a small number of firms, more comprehensive plans tend to enroll older workers and those more likely to have chronic conditions. Most of these studies focus on selection between managed care and more traditional plans (including Cutler and Reber, 1998), with managed care plans attracting the lower risk employees. A few others (including Ellis, 1989) examine selection within plans of the same type (i.e., managed care or traditional fee-for-service plans). Feldman et al. (1989) consider selection both across and within plan types, but they lack data on covered dependents, and only observe personal characteristics, not the eventual medical care expenditures by enrollees.[1]

[1] Other studies (Buchmueller and Feldstein, 1997, Strombom et al., 2002, and Cutler, Lincoln, and Zeckhauser, 2009) find evidence that unhealthy employees are also less likely to switch plans between years. I lack data on employees over time, so I am restricted to a static model of adverse selection.

These studies of individual firms nearly universally find evidence of adverse selection by individual characteristics, which are often correlated with, but not equivalent to, eventual financial risk. The problem with adverse selection is the "death spiral," where a plan that attracts only high-risk enrollees will have to charge ever-higher premiums that result in a high cost enrollee base. If costs are sufficiently managed by the insurer so that the difference in financial risk between plans remains within reason, then the death spiral is not inevitable (Feldman and Dowd, 1991). Furthermore, the initial high-generosity plan enrollees that are forced to switch to a lower-generosity plan are likely to bring up the average financial risk for their new plan, stabilizing the difference in costs between the plans.

A better measure of adverse selection is whether enrollees in the more generous plan spend more on medical care than enrollees in the less generous plan. The problem, as Pauly (1986) points out, is that higher generosity plans also encourage more spending due to moral hazard, so a positive correlation between coverage amount and medical care expenditure (controlling for variables that are common knowledge and can thus be used to price the contract) could suggest the presence of adverse selection, moral hazard, or both. The positive correlation test literature (Chiappori and Salanie, 2000) acknowledges this limitation, but provides a starting point for determining whether cost differences between plans in a national insurance market could lead to escalating premiums.[2]

Other researchers have suggested estimation techniques to separately identify adverse selection and moral hazard, but these methods carry downsides both in general and for employer-sponsored health insurance in particular. Structural estimation requires assumptions about the underlying utility maximization problem and the na-

[2] Chiappori and Salanie (2000) and Cawley and Philipson (1999) find no evidence for a positive correlation in the auto and life insurance markets, respectively. Finkelstein and Poterba (2004) find a positive correlation in the U.K. annuities market.

ture of private information about one's expenditure risk. Cardon and Hendel (2001) rule out adverse selection, though they do find evidence of moral hazard in line with the RAND experiment. Reduced form models may also allow for separating adverse selection from moral hazard, but these models require either exogenous variation in prices (Einav, Finkelstein and Cullen, 2008) or experience rating (Abbring et al. 2003); however, there are few, if any, natural experiments in the national employer insurance market in the U.S. during the late 1990's, while group health insurance premiums are almost never based on any individual enrollee's past utilization.

My paper employs a method for distinguishing the effects of adverse selection from the effects of moral hazard that does not rely on assumptions about consumers' preferences or the nature of the private information signal, nor does it require exogenous variation in the price of insurance. Instead, I use variation in the choice set of plans across a national sample of employers to determine what portion of an enrollee's spending can be attributed to the generosity of the plan she selects (moral hazard), and what can be attributed to her selecting that plan over alternatives based on her expectations about her future spending. Like the existing single-employer studies, I compare plans based on the composition of their enrollees, but I also consider the difference between plans in enrollees' ultimate financial risk, which would more directly lead to the adverse selection "death spiral."

2.3 Data

The Medical Expenditure Panel Survey (MEPS) is conducted by the Agency of Healthcare Research and Quality (AHRQ), a division of the U.S. Department of Health and Human Services. MEPS interviews households about health care expenditure and utilization, health status and conditions, and health insurance, including which family members were covered and when, and other demographic variables. Household respondents are surveyed every three to five months, a total of five times over their two years in the panel. A new panel has started each January since 1996.

If an individual or family receives medical care during their time in the survey, MEPS gathers more information on care events directly from the provider, including total expenditures on the patient's behalf and how the provider was reimbursed for that care, either by the patient out-of-pocket or by private insurance. The primary dependent variable in this analysis, total expenditure, is the sum of the private and self-paid expenditures from these events during the months when individuals in the family were covered by the selected employer-sponsored plan.

In addition, MEPS surveys business establishments and government units on the health insurance plans (up to four) that they offer their employees. MEPS collects characteristics of the business, eligibility rules for benefits, and details about the offered plans. From 1996 through 2001, except 2000, this sample includes employers of individuals in the household survey.[3] This allows for a link between a survey of the full set of insurance offerings of employers of varying size and industry from across the country to comprehensive data on medical care utilization and expenditures for 6,600 employees, including demographic data and measures of health status.[4]

2.4 Theoretical Framework

The analysis in this paper considers decisions made by the health insurance unit, which may consist of a single individual or multiple individuals in a family, subject

[3]MEPS documentation (AHRQ 1999) cautions that there was significant survey non-response from the employers, and thus there has been no effort to create weights that allow for nationally representative estimates from the Linked File, unlike the full MEPS sample.

[4]Only one spouse was surveyed in about 92 percent of two-earner households in this sample, but this is usually the spouse whose employer provided coverage to the other household members. The set of offers may also be less than complete if the firm offers more than four plans; government entities were asked about up to 36 plans, but I restrict the choice set to the four most popular plans of the type ultimately selected by the employee.

to eligibility rules. Single workers maximize their own individual utility. When the health insurance unit is a family, the worker receiving the offer of insurance maximizes expected family utility. In the analysis that follows, the "employee" refers to the individual in a single coverage plan, or the family in a family coverage plan. I take the decision to choose single coverage or family coverage as given.

A representative employee obtains utility from her consumption of health care, h, and goods other than health care. While she does not know exactly how much she (and her family) will spend on medical care during the coming year, she likely has a better estimate than her employer or the insurer, as she knows her personal health history, family history, the severity of any existing health conditions, and her preference for obtaining care. While the insurer may have data on her past usage of care if she has been enrolled previously, this information cannot be used to vary the premium by risk type; the premium for each policyholder in a given plan is equal to the expected average cost of all enrollees in that plan.[5]

At the beginning of the fiscal year, she will choose a health insurance plan from among the plans offered by her employer.[6] Her selected plan j is characterized by the amount she must contribute toward the premium, p_j, and the financial generosity of the plan, as summarized by the coinsurance rate, $0 \leq \delta_j \leq 1$, i.e., the share of h she must pay out-of-pocket.[7] The plan may also differ from other plans offered by her employer in non-financial elements: access to doctors, the breadth of in-network and out-of-network options for care, whether a referral is necessary to see a specialist, the

[5] Insurers do, however, charge different premiums for the same plans across employers. The plan's total premium depends on the region, industry, general demographic composition of the employer's workforce, and the collective utilization history of that employer's enrollees in the plan. However, the total premium paid on behalf of any worker, i.e., the sum of the employer- and employee-paid portions, does not vary by any individual characteristic. Conceivably, the employer could require higher-risk employees to contribute a larger amount toward the premium, but in practice this is rare; I exclude such plans from the analysis.

[6] I leave the modeling of the employer's decision to offer one or more plans of varying generosity for future research; see Moran, Chernew, and Hirth (2001) and Bundorf (2002).

[7] For simplicity, plan j is assumed to have a linear coinsurance schedule; a kinked schedule, as would be the case with a deductible, would not change the analysis.

inclusion of dental, vision, home health care, or mental health coverage, and coverage for particular procedures, such as mammograms or vaccinations. I restrict my focus to the financial generosity of the plans in the portfolio with respect to physician, hospital, and prescription drugg expenditures alone, as any two plans of the same type (e.g., all managed care plans) in a firm's portfolio tend to offer the same non-financial benefits.[8] A more financially generous plan will reimburse the care provider with a higher portion of every dollar spent on health care on the employee's behalf, leaving the employee with a smaller share; that is, plan j is defined to be more generous than plan k if $\delta_j < \delta_k$.

During the fiscal year, the employee's demand for medical care should depend only on the out-of-pocket price per unit she faces under selected plan j, δ_j; the more that the insurance plan reduces the out-of-pocket price below the true price, the more the employee will buy, due to moral hazard. This assumption follows directly from the concept of the sunk costs, which, according to traditional economic theory, are irrelevant in deciding how much of a product, in this case medical care, to buy. In this model, there are two sunk costs. The premium is a fixed cost (even if paid monthly or in each pay period, as the amount is a contractual obligation), and is therefore a sunk cost for the length of the contract. In addition, the opportunity cost of choosing one plan rather than another is also sunk during the contract period, and therefore irrelevant; that is, within the fiscal year, only the selected plan matters. Thus, moral hazard is influenced only by the absolute generosity of the selected plan — the more generous the employee's plan, the more she and her family will spend, irrespective of the rejected alternatives.[9]

[8]Table 4.4 shows how the selected plans differ from the rejected plans along non-financial elements. Any two HMO plans tend to be very similar; slightly more people prefer not to need referrals for specialist care, and prefer to have access to a chiropractor. Non-HMO plans are more different from each other; coverage for home health care and physicals is especially preferred. Non-HMO plan enrollees often choose a plan that requires referrals and has a deductible over plans, consistent with choosing PPO plans over FFS plans.

[9]Endogenous job selection could bias this correlation upward. A job applicant with high expected medical costs

Adverse selection, however, requires information on both the selected plan and the rejected alternatives, as those are the options between which the employee is self-selecting. In an insurance market with standardized options, or within one firm that is restricted to offering the same set of health insurance plans to all employees, the rejected alternatives are standard across enrollees, so their characteristics are of relatively little importance. In the national employer-sponsored health insurance market, however, the parameters of the rejected plans vary. The potential for adverse selection will be different at a firm that offers only two very generous plans compared to a firm that offers one very generous plan and one much more bare-bones plan, and very obviously different at a firm that offers just the one very generous plan and nothing else.

In this model, the distribution of risks is continuous and one-dimensional, based only on expected health care expenditure (as in Cutler and Reber, 1998), and risk type maps monotonically to the ex-post realization of medical spending.[10] Employees are offered two plans of the same type (e.g., managed care), with associated premiums that do not vary by individual but may differ between the plans, and no one chooses to remain uninsured.[11] For any fixed dollar amount of total spending on behalf of

may be attracted to a firm by its generous health plans (or one specific generous plan). Thus, a positive correlation between generosity and spending would result purely from adverse selection between jobs. However, there are several reasons to think that endogenous job selection does not invalidate this paper's results. For one, among all the uncertainty surrounding a job search, it is difficult to place a value on hypothetical plan offers, especially when all potential employers are likely touting their "generous health benefits." I also include firm characteristics, including industry and firm size, in the regression analysis, which may control for differences across firms. My sample is also limited to employees who take up an employer's insurance offer. Endogenous job selection is more of a concern when there is sorting between jobs that offer no health insurance versus jobs that offer at least one plan; even then, Hirth et al. (2006) and Monheit and Vistnes (1999) indicate that a substantial number of individuals who prefer not to have insurance are mismatched with firms that offer insurance, resulting in lower wages than would be ideal.

[10] Medical spending could be modeled as a function $h(\Psi)$, where Ψ represents the difference in monetary value between the more generous and less generous plans (Feldman and Dowd, 1982). Employees with low values of Ψ opt for C, and those with high values select G, and the analysis is no different than above. Medical spending may also be modeled as a function $h(\Psi, \epsilon)$, where ϵ is a stochastic variable accounting for uncertainty about one's health during the coming year. Employees with low Ψ will opt for plan C, but some will draw a large ϵ (bad health) and realize a large value for h; these employees will have low-generosity plans but high spending levels, biasing the expected positive correlation between absolute generosity and spending toward zero. A similar result holds for high Ψ employees in plan G who draw a low value of ϵ and thus spend very little.

[11] This restriction simplifies the exposition of the model but is not essential. If there are uninsured employees, the model assumes that they come from the left-most points on the risk distribution. The implications for the subsets choosing the less generous and more generous plans remain unchanged.

the employee, h, plan G requires less out-of-pocket spending than the less generous plan, C ($\delta_G h < \delta_C h$). Premiums for the two plans will depend on their respective average costs; unless G enrollees spend far less than C enrollees, enough to make up for the plan-paid advantage, plan G will have the higher premium, so $p_G > p_C$.

Adverse selection arises in this model because each employee chooses the plan that requires a lower total out-of-pocket spending at their expected expenditure level, h. If spending is likely to be low, then the lower coinsurance paid under plan G will not make up for the extra premium every month. However, if h will be high, then plan G's advantage in cost sharing will more than compensate for the higher premium. Below some threshold value of spending, h', plan C will be preferred, while above h', plan G results in a lower out-of-pocket spending amount, including premium. At h', the total out-of-pocket spending under the two plans are equal, so

$$p_G + \delta_G h' = p_C + \delta_C h'. \tag{2.1}$$

Solving for h',

$$h' = \frac{p_G - p_C}{\delta_C - \delta_G} = \frac{p_G - p_C}{(1 - \delta_G) - (1 - \delta_C)}, \tag{2.2}$$

which indicates that the threshold level of spending above which consumers would prefer the more generous plan is simply the ratio of the relative premium, or the difference in premium between the two plans, to the relative generosity, the difference in the plan-paid portion out of every dollar of medical spending.[12]

The MEPS dataset essentially draws a small number (often just one) of employees randomly from the risk distribution in a firm. The expected value of that draw will depend on the threshold spending level, h'. As h' increases, the expected value of h

[12]This framework implicitly assumes a risk neutral consumer. A risk averse employee would likely prefer plan G even at some h slightly below h'. In this case h', and its (theoretically) positive correlation with spending, would be overstated, so the correlation measured in the regression model is an upper bound on the true correlation. In the results section, I show that the measured correlation is not significantly greater than zero, so we can be confident that there is no correlation with h' and thus no adverse selection along the demand for financial generosity.

both above and below h' also increases.[13] The intuition is presented in Figure 2.1. Firm X and Firm Y have identical risk distributions, $f(h)$, in their employee pools, and both firms offer the same less-generous plan C requiring the same employee contribution.[14] The firms offer different more-generous plans: Firm X's plan G_X is more generous, and requires a larger employee contribution, than Firm Y's plan G_Y, and thus is further to the right on the generosity continuum. Because the generosity gap between the two plans is wider in Firm X, the threshold level of spending h' will also be further to the right in Firm X than in Firm Y, which means the randomly-selected X employee who chose plan C will have a larger expected value of h than the sampled Firm Y employee with the same plan.[15]

The exercise in Figure 2.1 summarizes the strategy for separately identifying moral hazard and adverse selection among employees who opt for their respective firm's lower generosity plan. A positive correlation between plan C's absolute generosity and the C enrollee's total medical care expenditure, all else (including the relative generosity and relative premium) equal, reflects the presence of the moral hazard incentive, because only the generosity of the selected plan matters. If unhealthy employees adversely select into more generous plans, the threshold level of risk, h', i.e., the ratio of plan C's relative premium to its relative generosity, should be positively correlated with spending, holding C's absolute generosity constant.[16]

The exercise is similar for plan G enrollees. In Figure 2.2, I compare the same

[13]See Appendix for the proof to this statement.

[14]Note that h is endogenous to plan choice, due to moral hazard, but in Figure 2.1, the relevant employees at each firm face the same moral hazard incentive, as both choose plan C. The same is true for the employees choosing plan G in Figure 2.2.

[15]In this intuitive model, plan G_X requires a larger employee contribution than plan G_Y, for two reasons. First, because the insurer reimburses the care provider with a higher proportion of any dollar of spending under plan G_X, the expected cost for any fixed dollar amount in any fixed population is higher for G_X than for G_Y. Second, as shown in the graph, the enrollees in plan G_X are higher risk on average, so the premium will need to be higher to reflect the higher average cost. The higher premium only reinforces the adverse selection — only the riskiest Firm X employees will choose to pay the much higher premium for G_X.

[16]Plan C should have a lower premium than plan G, and by definition is lower generosity, so both relative premium and relative generosity are negative, and thus the ratio is positive. The threshold value h' is always positive if plan G is more expensive than plan C.

Firm X to a different firm, Z, which also has the same risk distribution among its employees. Firms X and Z offer the same more-generous health plan G, but Firm Z's less-generous offering, C_Z, is more generous than at Firm X, requiring a greater premium. Because Firm X's plan C is further to the left, so too is the threshold spending level h', and therefore the sampled employee at Firm X should have a lower expected value of spending. Thus, holding constant the absolute generosity of the selected plan (in this case, G), a positive correlation between h' and the ultimate level of medical spending h suggests adverse selection.

In summary, in the national employer-sponsored health insurance market, the variation in both absolute and relative generosity and relative premium allows for the separate identification of moral hazard and adverse selection, respectively. The more generous the selected plan is, the more that moral hazard will induce the enrollee to spend on covered medical care; thus, a positive correlation between absolute generosity and medical expenditures suggests the influence of moral hazard. A positive correlation between expenditures and the threshold level of spending that separates those who prefer the more generous, more expensive plan and those who prefer the cheaper option, controlling for the moral hazard incentive, suggests adverse selection.

2.5 Health Insurance Generosity

A key contribution of this paper is the creation of a single measure that summarizes the financial generosity of a health insurance plan. This generosity measure simplifies a plan's often complex, non-linear coinsurance schedule, and allows for comparing plans over the range of one's possible out-of-pocket costs.

The simplest possible generosity measure is the total amount that the plan actually reimburses on behalf of the enrollee (and her family, where applicable) during the given year. This is not the measure I use, as there are a number of serious problems with this measure. First, the family's spending level is likely endogenous to the plan's characteristics; in fact, one of the hypotheses tested in this paper is exactly that: a plan with lower cost-sharing will likely lead to more spending, due to moral hazard. Second, this variable is only ascertained for the selected plan, not the rejected alternatives. Finally, for those plans with no deductible, there is a mechanical relationship between the chosen plan's actual insurer-paid amount and the total expenditure level, in effect tracing out the average plan-paid rate for that enrollee.

As an alternative, I calculate the generosity of any plan based on a sample of employees with similar risk profiles, and determine how much every plan offered to those employees would pay out if each member of the sample was enrolled in the plan, using the plan's deductible(s), co-payments, coinsurance, and maximums. The "absolute generosity" of a plan j is the mean percent of total spending that the plan would pay for these hypothetical enrollees.

In the model, the employee chooses a plan from among her portfolio of offers based on her expectations about her (and her family's) expenditures during the coming year. I do not observe this range of expectations. Instead, I define the range of her possible outcomes as the empirical distribution of the ex-post realization of expenditures by others in the MEPS linked employer-employee sample who are of a "similar" risk type ex-ante. This risk type is defined by three variables: the survey year, to account for the growth of spending over the six years; the employee's enrollment status, based on whether she chose a single or family plan; and whether anyone in the employee's family (or just herself, if she chose an individual plan) had a chronic condition or health limitation before or during the plan year, to account for her baseline risk.[17]

The N_r employees in subsample r were offered a total of J_r plans, between one and four plans per person. For each employee i in that subsample, I calculate how much plan $j \in J_r$ would have paid on i's behalf, based on i's actual sequence of care events and total amount spent, h_i. Plan j's generosity is then

$$g_j \equiv 1 - \delta_j = \frac{1}{N_r} \sum_{i=1}^{N_r} \frac{h_i - G_j(h_i)}{h_i}, \tag{2.3}$$

or the average plan-paid rate if every employee in subsample r were covered by plan j, where G_j is the (possibly non-linear) coinsurance schedule for plan j.[18]

There are other factors that will lead a consumer to pick one plan over another besides financial generosity and the employee contribution to the premium, including the available network of doctors, restrictions on access to specialists, and the availability of insurance for other types of care besides basic physician, hospital, and prescription drug coverage. Physician networks and specialist access will differ between health management organization (HMO) plans, preferred provider organization (PPO) plans, and fee-for-service (FFS) indemnity plans.[19] HMO plans make up about 47 percent of the plans chosen in this sample, while PPO plans make up 46 percent, and FFS plans are only 7 percent. Financial generosity may, then, have a different effect depending on whether the selected plan was an HMO, PPO, or FFS.

[17] Chronic conditions range from cancer and heart disease to asthma and diabetes. A respondent is "limited" if she reports a limitation during a battery of questions about activities of daily living and functioning. I include not just conditions that exist before the plan year begins, but also conditions that are diagnosed during the year, as employees may have some private information about undiagnosed conditions that may soon arise because of family history or early onset symptoms.

[18] In a regression of the generosity measure on a comprehensive list of plan characteristics, the R^2 is 0.63, and the coefficients for most elements are significantly different from zero at the 99 percent confidence level, and in the predicted direction: the plan is more generous when the deductible is zero or small, and when the coinsurance rates and co-payments for physician and hospital visits are low. These results suggest that the financial generosity of a plan relies on each cost-sharing element of the plan working in concert, rather than just one dominant plan characteristic.

[19] Generally, HMOs employ a network of health care providers, and within this network, care is covered at generous rates; care sought outside the network is generally not covered, except for emergencies and where the primary care physician has referred the patient to a specialist. PPOs pay for care received both in- and out-of-network, but in-network coinsurance rates are lower. FFS or traditional indemnity plans pay either a flat fee or percentage of the cost regardless of where the care is received. See Cutler, McClellan, and Newhouse (2000) or Moran, Chernew, and Hirth (2001) for more details.

Unfortunately, I am not able to directly compare the generosity of managed care (HMO) plans to less-managed plans like PPOs and FFS plans. Cutler, McClellan, and Newhouse (2000) find that the quantity of care does not vary much by plan type, but that the price paid to providers for that care varies greatly, with HMOs paying the least, FFS plans paying the most, and PPOs somewhere in between. Because HMOs have lower payouts to providers, premiums and cost-sharing for these plans need not be as high as PPO or FFS plans, and thus for the same dollar amount of spending, the HMO plans appear to be more generous. This runs counter to the assumption in other studies (Feldman and Dowd, 1982, Cutler and Reber, 1998) that the HMO plan is the less generous option that should attract lower risk employees. In addition, the total spending measure, the sum of the amounts paid by private insurance and the patient herself, will be lower for HMO plans that are able to negotiate lower prices for care.[20]

Because neither the expenditures nor the plans' generosity are comparable between HMO and non-HMO plans, I limit my comparisons to plans of the same type. If plans j and k are both HMOs, or both non-HMOs, then the relative generosity of the chosen plan j is $|g_j - g_k|$.[21]

2.5.1 Indications of Adverse Selection

The concept of adverse selection suggests that when facing a higher premium for more generous coverage than for a more restrictive plan, individuals with private information that they are likely to need large amounts of care will pay the higher

[20]If my calculation overstated the generosity of HMOs by a consistent percentage (for instance, the 11-17 percent difference between HMOs and PPOs in prices paid for the same procedure in Cutler, McClellan, and Newhouse), I could just deflate the HMO generosity by this same percentage to make the generosity measure comparable across plan types. Because total spending is lower for HMOs, however, I would have to estimate how much plans of each type would pay for every procedure ever performed in any care event, which is beyond the scope of this paper.

[21]Because FFS plans are a small portion of the total number of plans, I include them with PPO plans. While most of the non-HMO comparisons will be a PPO chosen over another PPO, I also classify a PPO chosen over a FFS plan, or vice versa, as a valid comparison.

price while lower risk individuals will opt not to. With only bad risks in the risk pool, the premium increases, driving out more good risks and further increasing the premium.

This would not be the case, however, if more generous coverage was not also more expensive. If one plan was more generous in every way — not just financially, but also with less restrictive access to specialists, a wider breadth for the provider network, and coverage for secondary care types like dental and mental health — yet cost the same as a less generous plan, there would be very little reason not to pick the more generous plan, whether one's expected expenditures were high or low. In this case, both high and low risks are in the same pool, and there is no adverse selection. (Moral hazard may still result in a positive correlation between insurance generosity and expenditures, however.)

In practice, more generous plans often do not require the higher employee-paid share of the premium (Table 2.2). Out of 3685 plan pairs (a comparison between the chosen plan and each of the rejected alternatives, in turn), 1438, or 39 percent, required a lower employee contribution from the more generous plan. Another 21 percent required the exact same contribution from both plans, with the majority of those cases paid entirely by the employer.

Not surprisingly, employees are more likely to choose the more generous plan if it requires the same or lower employee contribution than when the more generous plan has a higher price. Still, 38 percent chose the lesser plan when it had a higher employee contribution. The tradeoff faced by the employees who are offered a package with employee contributions that are not proportional to generosity is less clear. These individuals may sign up for plans that are more attractive on a non-financial level; for example, perhaps one's preferred doctor is in the network of only the plan that is less financially generous. It could also be that the less generous plan is the default option, which keeps costs low for the employer and insurer, especially when employees do not depart from this default.

This finding also extends to the total premium: the higher premium plan is often not more generous, and when the less generous plan has a higher premium, it is still selected nearly 39 percent of the time.[22] While a plan with a higher total premium may reduce the employee's cash wage more than a plan with a lower total price, it is unlikely that the employee feels the pinch of this incentive directly; thus, the employee is likely to respond more to the portion of the premium she pays rather than the total premium, which includes her employer's contribution. Therefore, I focus on the employee's contribution to the premium, not the total premium itself, in the estimation of adverse selection.[23]

The theoretical framework which leads to the derivation of h', the threshold level of spending above which employees should prefer the more generous plan, suggests an approach for dealing with employees who face a higher premium for the less generous plan. Recall that h' is equal to the ratio of the relative premium to the relative generosity of the selected plan. When the relative premium is the same for both plans, $h' = 0$, and the model predicts that the more generous plan will always be selected; if non-financial differences are minor or not greatly valued, the employee has no reason not to choose the more generous plan if it costs exactly the same as a lesser alternative, no matter what her expected spending level. The employee should also choose the more generous plan, regardless of how much she will ultimately spend, if

[22]While this finding may call the generosity measure into question, the R^2 on a regression of the total premium on plan characteristics is only 0.04, so total premium does not appear to hinge on the coinsurance schedule. This matches the counterintuitive finding in Cebul, Rebitzer, Taylor, and Votruba (2008) that higher premium plans are often not more generous because of the constant churning in both the employee-employer and the employer-insurer relationships, encouraging insurers to chase short-run profits at the expense of quality.

[23]Chernew, Frick, and McLaughlin (1997) find that employees are more responsive to their share of the premium than to the total premium itself.

the lesser plan has a employee contribution (where the relative premium is negative). Because the more generous plan is the preferred option at any spending level $h \geq 0$, I set $h' = 0$ for any employee facing the same or lesser price for the more generous plan.

On the other end of the spectrum, some plans have the same calculated generosity, but different premiums. In that event, h' should be infinite, as the denominator, relative generosity, is zero. This means that regardless of how much the employee plans to spend, the model predicts that "less generous" plan (in this case, the plan with the same generosity but a lower premium) is always preferred. In practice, because the relative generosity is often quite small (within one or two percentage points), I cap h' at the maximum level of expenditures observed in my sample, approximately \$201,000; because no employee's expected expenditure is anywhere close to this maximum, no one should pick the more generous plan when the relative premiums is relatively large and the relative generosity is relatively small (or zero).

Still, many employees with $h' = 0$ pick the less generous plan with the same or higher employee contribution to the premium, while many others with $h' = \$201,000$ pick the higher premium but barely more generous plan, as shown in Figure 2.3. Only 65 percent of HMO enrollees and 59 percent of non-HMO enrollees with a negative (paying more for the less generous plan) or zero (same premium for both plans) h' choose the more generous plan. The percent selecting the more generous of any two plans should be declining in h', but for both HMOs and non-HMOs, the graph is relatively flat, with a slight decrease for most deciles of non-HMO enrollees reversed by a large share of employees paying a higher premium even when the two plans have equal generosity ($h' = \$201,000$). According to the model that results in adverse selection in the insurance market, employees with large values of h' should never pick the more generous plan, but when the two plans are (approximately) equal in generosity, nearly two-thirds of the sample opt for the higher premium option.

If the market is characterized by adverse selection between plans of the same type (HMO or non-HMO), the percent picking the more generous plan would increase with expenditures. Figure 2.4 indicates this is not the case for non-HMOs, where the probability of selecting the more generous plan instead stays relatively flat throughout the expenditure distribution. For HMOs, however, there is a noticeable increase in the probability of selecting the more generous plan as total expenditure increases.

The model suggests that, unlike adverse selection, moral hazard depends only on the absolute generosity of a plan — the more generous the selected plan is, the more the employee (regardless of her risk type) will likely spend. This is exactly what I observe for non-HMO enrollees, and less so for HMO enrollees, in Figure 2.5. With each successive decile of absolute generosity, the average expenditure for that decile tends to increase; i.e., a linear trend line through all three series is positive and significantly different from zero. HMO expenditures are slightly lower for all but three deciles.[24]

2.6 Empirical Strategy

To test for the presence of adverse selection and moral hazard in the employer-sponsored insurance market, I estimate a series of regressions where the dependent variable is total expenditure, and the key independent variables are measures of absolute and relative generosity.

[24]Figure 2.5 by itself does not indicate moral hazard, however, because adverse selection may also lead to a positive correlation between the selected plan's generosity and expenditure. On the other hand, coupled with the evidence in Figures 2.3 and 2.4, and Table 2.2, that the conditions for adverse selection do not exist, Figure 2.5 provides strong evidence for moral hazard.

2.6.1 Estimation Sample

The sample for these regressions consists of pairs of plans of the same type (HMO or non-HMO), where one plan is the selected plan and the other is each rejected plan in turn. If an employee is offered plans A, B, and C, and chooses A, I would like to be able to measure the selection between A and the employee's second-best plan; unfortunately, we do not know whether B is preferred to C, or vice versa. Therefore, I include every comparison between the selected plan and a rejected plan of the same type, because it is unclear whether the selection of risk types is most relevant between selected plan A and rejected plan B, or between A and rejected plan C. This means that employees offered three plans are double-counted, while employees offered four plans are triple-counted.[25] Standard errors are clustered by individual, to account for double- and triple-counting, and for the few employees that appear in the linked employer-employee sample during both years of their observation by MEPS. I include as many as three plan pairs per person; for government employees who are asked about more than four plans, I retain only the selected plan and the three rejected options with the highest enrollment among employees at that establishment.

I do not include a plan pair where one plan is an HMO and the other is a non-HMO. That means that if the employee chose an HMO, I only consider relative generosity and premium between the selected plan and other HMO plans in her portfolio.

I exclude employees who are offered only one plan, and employees who are offered only one plan of the type they ultimately selected. That is, if the employee chose an HMO plan, and all of her rejected alternatives were non-HMO plans, I consider her portfolio to consist of only one plan, and she is excluded. For one-plan employees,

[25] As a robustness check, I also estimate the main result using just one pair per person; see section 4.6.

the other option is to go uninsured, where the generosity is $g = 0$, which means that absolute and relative generosity are equal. The absolute generosity coefficient, then, is also picking up the effect of adverse selection between insurance and being uninsured. Similarly, for an employee offered only one plan of the type she chose, the absolute generosity coefficient may pick up the effect of adverse selection across plan types.

Information is asymmetric in the employer-sponsored insurance market because very few plans allow the insurer to know, and charge different premiums based on, any characteristics of the individual employee aside from firm characteristics. MEPS asks employers in the establishment survey whether each plan's premiums vary according to age, gender, family size, wage, or some other factor. Approximately 83 percent of chosen plans indicated that they do not vary premiums based on any of these factors; I exclude the 17 percent of plans that have variable premiums.[26]

2.6.2 Specification

The medical expenditure variable in this paper, as in the existing literature, is right-skewed and has a small mass point at zero. Figure 2.6 displays the distribution of the natural logarithm of total expenditures for HMO and non-HMO enrollees. There is a small mass point at zero expenditures, while the rest of the distribution is normally distributed in log form. The HMO distribution is slightly to the left, and more likely to have zero expenditures; this could suggest that HMO enrollees are healthier, or that HMOs just pay out less for the same procedures.

Duan, Manning, Morris, and Newhouse (1983) suggest using a two-part model when medical expenditures is the dependent variable. In the first equation, I estimate

[26] Of those that did vary premiums by observable characteristics, the most popular pricing scheme was based on family size alone (8.2 percent of chosen plans), followed by "other reasons" alone (3.6 percent) and age alone (1.2 percent); no other combination of these factors totaled more than one percent of the sample of chosen plans.

a logit regression where the dependent variable is an indicator for whether total medical expenditure h for employee i (and her family, if she has family coverage) is positive,

$$\begin{aligned} I(h_i > 0) &= f(\alpha_1 g_j + \alpha_2 log(h') + X_i\xi + \nu_i) \\ &= f(\alpha_1 g_j + \alpha_2 log(\frac{p_j - p_k}{g_j - g_k}) + X_i\xi + \nu_i), \end{aligned} \tag{2.4}$$

where g_j and p_j is the absolute generosity and employee contribution, respectively, for plan j. The logit estimation accounts for the mass point at zero in the expenditure distribution.

In the second equation, I estimate a linear regression of the natural logarithm of medical care spending, $log(h)$, conditional on having non-zero expenditures, on the absolute generosity (g_j) of selected plan j and the natural logarithm of the threshold level of spending ($h' = \frac{p_j - p_k}{g_j - g_k}$) above which the more generous of the two options j and k should be preferred,

$$\begin{aligned} log(h_i|h_i > 0) &= \beta_1 g_j + \beta_2 log(h') + X_i\gamma + \epsilon_i \\ &= \beta_1 g_j + \beta_2 log(\frac{p_j - p_k}{g_j - g_k}) + X_i\gamma + \epsilon_i. \end{aligned} \tag{2.5}$$

The loglinear estimation accounts for the long right tail of the expenditure distribution.

The predicted expenditure, $\hat{h}$, will be the product of the predicted probability that the expenditure is positive from 2.4, $\hat{p}$, multiplied by the conditional predicted expenditure from the 2.5, $\widehat{h|h>0}$. Duan, et al., (1983) propose scaling the predicted expenditure by a smearing factor, $\hat{s}$, equal to the exponentiated residual from the loglinear regression, as the error term is likely to be non-normal. So the final predicted expenditure is

$$\hat{h} = \hat{p}(\widehat{h|h > 0})(\hat{s}). \tag{2.6}$$

The smearing factor simplifies the differentiation of equation (2.6) with respect to g and $log(h')$, in turn. The marginal effect of absolute generosity on total expenditure for employee i reduces to

$$\frac{\partial \hat{h}}{\partial g} = \frac{\hat{\alpha_1}}{1 + exp(\hat{\alpha_1} g_j + \hat{\alpha_2} log(h') + X_i \hat{\gamma})} \hat{h}. \tag{2.7}$$

Similarly, the marginal effect of $log(h')$ on total expenditure is the predicted expenditure, $\hat{h}$, times the logit coefficient for $log(h')$ divided by the denominator of the logistic function. In the results tables, I report the mean of this derivative evaluated for each individual, with standard errors calculated by bootstrap (500 iterations). The null hypothesis of no moral hazard assumes that the marginal effect of g on $\hat{h}$ is zero, while the null for no self-selection by risk type assumes that the marginal effect of $log(h')$ on $\hat{h}$ is zero.[27] As the measures of spending (h) and generosity (g) have different meanings due to the ability for HMOs to negotiate lower prices for the same services, I estimate this two-part model separately for HMO and non-HMO plan pairs.

X includes the few variables that insurers who sell coverage through employers may be able to include in their pricing formula, so that we control for any symmetric information in the transaction.[28] The only two elements that vary at the individual level in this matrix are an indicator for whether the policyholder has family coverage rather than single coverage and, in the non-HMO regression only, controls for whether the selected plan was a FFS plan or PPO plan. I also control for the characteristics

[27]Though both coefficients are expected to be positive, I perform two-sided tests; one possibility, contrary to my model, is that self-selection could be favorable, if higher-spending employees choose the less generous plan in their portfolio, perhaps because of liquidity constraints, higher discount rates, or higher risk tolerance.

[28]Nearly all employers in the sample employ only one MEPS Household Component respondent, so I cannot estimate a firm fixed effect.

of the firm and the local insurance market. These variables include the region of the country, whether the policyholder lives within a metropolitan area, whether the policyholder is a member of a union, the industry of the policyholder's firm, the firm's size, whether the firm self-insures, whether the firm offers health savings or flexible savings accounts, and year dummies.

2.6.3 Weights

The first two columns of Table 2.4 reinforce the warning in the MEPS documentation that the sample of employees from the MEPS household survey linked to the employer survey is not nationally representative, a problem made worse by my sample inclusion requirement that employees are offered more than one plan of their selected type (HMO or non-HMO). The first column presents summary statistics for the the full MEPS Household Component sample of working adults with insurance through their own employer, with means and standard deviations calculated using the population weights provided by MEPS. The second column features the unweighted means and standard deviations of my sample. The variables I have included in this table are those that are most different between the estimation subset and the full MEPS sample. White non-Hispanic, non-self employed, higher-income employees working in large establishments or the government tend to be most likely to have multiple health insurance offers of the same type, and are thus over-represented in my sample.

In order to produce results that are more easily interpretable and generalizable, I re-weight my sample to account for this over-representation. Using the full (linked and unlinked) household sample of employer-insured working adults, I create an indicator variable for whether the employee is in the estimation sample. I then estimate a logistic regression of the probability of inclusion on the gamut of personal and firm characteristics from the household survey, including the variables in Table 2.4

plus age categories, marital status, self-reported health status, education, region, urbanicity, and occupation as explanatory variables. An employee's weight is their MEPS-provided weight divided by the predicted probability of inclusion in the sample, so that the groups that are over-represented in the sample are under-weighted. The third column in Table 2.4 includes means and standard deviations calculated using the new weights; the means are much closer to the means reported in the first column.

Table 2.3 provides summary statistics for the regression variables of note, both unweighted and weighted. The standard deviations on the expenditure variables are very large relative to the means. On average the plans pay 75 percent of an employee's total annual medical care spending. The mean difference in generosity between plans is between 8 and 10 percentage points; the difference in generosity is smaller between HMO plans than between non-HMO plans. Conditional on the employee's premium contribution for the more generous plan being the higher than for the less generous plan, the relative employee contribution is approximately $450 per year. The mean threshold spending level, h', is lower for non-HMO plans because a higher percentage of the more generous of two non-managed care plans require the same, or even lower, employee contribution.

2.7 Results

2.7.1 Expenditures

In Table 2.5, I report the results of the regression of the natural logarithm of total annual medical expenditures on the selected plan's absolute generosity and the natural logarithm of the threshold spending level (h'), separately by whether the selected plan and rejected alternative are both HMO (top panel) or non-HMO plans (bottom panel). All regressions include plan and firm characteristics and year fixed effects, and are weighted to reflect non-random missing observations in the MEPS HC-IC Linked File.

Absolute generosity is positively associated with medical expenditure for both HMO and non-HMO plans, but it is statistically significant only for non-HMO enrollees. The logit coefficient (column 1) indicates that non-HMO enrollees are more likely to spend a positive amount on medical care the more generous is their selected health plan. Conditional on having positive spending, the loglinear coefficient (column 2) indicates that for every additional percentage point paid by the plan, the employee (and her family) enrolled in a non-HMO plan spends 2.1 percent more on medical care. The marginal effect (column 3), which accounts for both the incentive to spend some positive amount and the incentive to spend more thereafter, suggests a large response to absolute generosity; a small change in generosity leads to more than $460 in additional medical spending, approximately 14 percent of the average non-HMO family's spending.

The response is much smaller for HMO plans, and imprecisely estimated. Both the logit and loglinear coefficients for absolute generosity are positive but not significantly different from zero, and the point estimates are smaller than those estimated for non-HMO plan pairs. The marginal effect suggests that a small increase in the generosity of the selected plan increases spending by $180, or 6 percent of average spending, but the standard error on the marginal effect is too large to rule out the absence of an effect with any certainty.

The theoretical model predicts that, in an insurance market subject to adverse selection, expenditures will be positively correlated with the threshold level of spending (h') above which the insured would prefer the more generous of two insurance contracts. This does not appear to be the case between employer-sponsored health insurance plans of the same type. For HMO enrollees, the logit coefficient is positive and the loglinear coefficient is negative, while the opposite is true for non-HMO enrollees. Only the loglinear coefficient for non-HMO plans is positive and significant, but the marginal effect that takes into account zero spenders is actually negative; whereas theory would predict an increase, spending decreases by 1.5 percent with a small increase in the relative premium (or a small decrease in the relative generosity), though the effect is not statistically significant. The marginal effect is larger in absolute value and positive for HMO plans, but the standard error is also larger.

I can further decompose medical care spending to examine whether moral hazard or adverse selection differs by the type of care received (Table 2.6). The same pattern holds for each care type: moral hazard leads non-HMO employees to spend more on physician care, hospital care, and prescription drugs in more generous plans, while HMO enrollees do not have the same response. The marginal effect for hospital care is especially large; buoyed by the large increase in the probability of spending a positive amount on hospital care, a small increase in non-HMO plan generosity increases hospital spending by \$1531, more than doubling the average hospital expenditure. The marginal effects are much smaller, and statistically insignificant, for HMO plans. There is no evidence of adverse selection, as each marginal effect is not significantly different from zero.

These results indicate that moral hazard leads employees enrolling in more generous non-HMO, but not HMO, plans to spend more, and that the effect is substantively large for non-HMO enrollees. There is no detectable adverse selection between either two HMO nor two non-HMO plans, but this could be because so many portfolios require the same employee contribution for both plans, so employees of all risk types should select the more generous plan if its premium is no higher. The first two columns of Table 2.7 present regression results separately for employees who pay the same portion (in dollars or percent) of the total premium for both plans, and for those who contribute differently to each plan.

One of the assumptions in the model is that moral hazard does not depend on premiums, neither absolute nor relative, and these results seem to confirm that assumption: as I find in the main regression, the moral hazard incentive is weak for HMO enrollees in both groups (top panel), and the marginal effect of generosity on spending is positive and statistically significant, and of comparable magnitude, for both groups of non-HMO enrollees (bottom panel). The marginal effect of $log(h')$ on spending is a larger positive number for HMO and non-HMO enrollees that face a higher premium for a more generous plan than for those who face the same premium for both plans, but in neither case is it statistically significant.

In the third column of Table 2.7, I present the marginal effects from the two-part model just for those who pay a higher employee premium for the more generous plan (regardless of whether they chose the more or less generous plan);[29] this group has an h' that's greater than zero but less than infinity, so there is a range of expected spending where the less generous plan is optimal, and other values where the more generous plan should be preferred. The point estimates are similar, though because of the smaller sample size, the standard errors are larger. The pattern of results is the same: a large positive effect of absolute generosity on spending, and a small and insignificant effect of $log(h')$, which together suggest moral hazard but not adverse selection.

[29] The regression results are nearly identical between people who select the more or less generous plan.

2.7.2 Utilization

Moral hazard and adverse selection, if present, should be detectable not only in their effect on expenditures, but also in medical care utilization. A high-risk individual, in addition to spending more, will likely also seek care more often, so the total number of care events should be correlated with absolute and relative generosity exactly like expenditures. Whether physician visits or hospital visits or nights increases, however, is more ambiguous: more generous insurance may lead to substitution toward going to the hospital rather than the physician, or it might lead to more frequent office visits and fewer hospitalizations.

Because the number of care events in a year is a count variable, I use a Poisson specification.[30] The results in Table 2.8 suggest that non-HMO enrollees do increase the number of both physician visits and hospital nights due to moral hazard, while a generous HMO plan sees an increase only in physician visits. The marginal effect of $log(h')$ is small and insignificant, so employees do not appear to self-select by their frequency of care use. The effects are quite substantial: the marginal effect of increasing generosity slightly leads to nearly 14 more physician visits and four more nights in the hospital for non-HMO enrollees, and seven more visits to the physician in an HMO plan. Moral hazard thus seems to materialize as an increase in the frequency of medical care utilization, not just an increase in the amount spent by consumers and their care providers.

2.7.3 Direct Measures of Selection

In most of the existing literature that uses data from a single employer, or a sample of employers in a metropolitan area, to determine whether there is adverse selection between plans, the authors look for significant differences in the demographic and

[30] OLS coefficients are similar in significance and relative magnitude.

health composition of the enrollee pools in each plan. Fortunately, the MEPS data provides information on measures of health status, so I can directly test for the presence of adverse selection between plans of the same type (HMO or non-HMO). Furthermore, because the measurement of these health measures across plan types is not corrupted by HMO management's skill at negotiating lower prices, I can also test for whether adverse selection occurs across types, i.e., whether high risk employees choose non-HMO plans while healthier employees opt for an HMO. I display the results of these regressions in Table 2.9.

I estimate the correlation between absolute and relative plan generosity and, in turn, six different health measures. The first such measure is the number of people in the employee's family (or just the employee, if the plan covers just the individual) who have a chronic condition.[31] The results of the OLS estimation are reported in column 1.[32] The number of family members with a chronic condition is positively correlated with the threshold spending level (h') in HMO plans, but the effect is very small and not statistically significant; for non-HMO plans, the effect is even smaller and similarly insignificant.

Other measures of health are similarly uncorrelated, or have a very small positive correlation, with $log(h')$. Employees with family members who consider their own physical health as "fair" or "poor" or have a limitiation are statistically more likely to sort into the more generous of two plans, but the magnitude of the effect is quite small. Those with infants or relatives over age 50 (where the age profile of health spending begins to ascend rapidly) covered under their plan are no more likely to

[31]Because an individual may have more than one chronic condition, I also estimate a Poisson regression for the total number of conditions in the family. The results match in significance, but the magnitudes are larger than those reported in column 1.

[32]For all six health measures, Poisson coefficients are similar in significance. I also estimate a logistic regression for whether the employees has any family members with each of the six health outcomes, and the results are also similar.

choose the more generous plan. Thus, there is no direct evidence of adverse selection between plans of the same type.

Interestingly, the coefficient on absolute generosity is positive and significant for three out of the six measures among non-HMO enrollees; this may arise because conditions are more likely to be diagnosed and treated the more generous is the selected plan, as care is sought more often (moral hazard), or because sufferers of these conditions seek out firms that offer more generous plans (endogenous job selection). This does not appear to be the case for the HMO plans, however.

In section 2.5, I outlined why HMO and non-HMO plans cannot be compared in the total expenditure regressions; because HMOs are able to negotiate lower prices than non-HMOs for the same procedures, expenditures and generosity are not measured on the same scale. With these regressions involving the direct health status measures, however, there is no such limitation. I include both HMO and non-HMO enrollees in the same regression, and include an indicator variable equal to one if the plans in the pair were both HMO plans. The coefficient on the HMO indicator is given in the bottom panel of Table 2.9. If, as expected, the healthiest employees choose HMO plans, which are generally considered to be less generous than PPO or FFS plans when measured properly, then we should observe that each of these coefficients should be negative and significantly different from zero. I find that employees in HMO plans are less likely to have people in their families with chronic conditions (6 percent), or over the age of 50 (7 percent), but neither effect is statistically significant; the coefficient on the HMO indicator is also negative and fairly large (17 percent, statistically significant at the 90 percent confidence level), in the regressions where the dependent variable is the total number of chronic conditions in the family (not shown). HMOs are actually significantly more likely (at the 90th percentile) to have young children (3 percent); while children under age 2 are expensive to insure due to their frequent use of care, this would further indicate that HMOs enroll younger, generally healthy families. These results (which are similar to Cutler and Reber, 1998, and Feldman and Dowd, 1982), though weak statistically, suggest that enrollees in the comparatively less generous HMO plans are less risky to the insurer, as predicted by the model of adverse selection across plan types.

2.7.4 Robustness Checks

One plan comparison per employee The regressions above include every valid (same plan type) comparison between plans for each employee. That means that there are two observations for the employees who were offered three plans of the same type as the one they selected (i.e., the HMO plan they selected plus two other HMO options), and three observations for the employees who were offered four plans of the same type as the one they selected. The dependent variable is the same for each of the two or three observations for these employees, so there may be a concern that the positive correlation between expenditures and the absolute generosity of the selected plan, or between expenditures and the threshold level of spending that divides risk types, may be biased upwards.

The employees who are offered exactly two plans of the same type should be unaffected by the decision to double- and triple-count employees. The marginal effect of absolute generosity on total spending for just this group (Table 2.10) is smaller and insignificant for non-HMO enrollees, though the loglinear coefficient is positive and significant as expected; the small sample size, leading to a larger standard error in the logit regression, likely leads to the insignificance. Here, the marginal effect for HMO enrollees is actually positive and significant, and larger in magnitude than the marginal effect for non-HMO enrollees in Table 2.5, but this is entirely due to a large coefficient in the logit regression; the loglinear coefficient is negative and insignificant. The marginal effect of $log(h')$ is positive and significant for HMO plans despite a negative and significant coefficient in the loglinear regression, and negative and significant, counter to theory, for non-HMO plans. While these results are somewhat different than the main estimations, the disparities are mostly explained by the small sample size.

Timing of medical care If an enrollee in the more generous plan foresees switching to the less generous plan when the fiscal year ends, she may increase her spending now while the out-of-pocket price is low. Conversely, if the enrollee plans to switch to a more generous plan, she will delay spending until the following year. Both scenarios suggest that the rejected plan's generosity may actually influence spending during the plan year, contrary to the model, so the absolute generosity coefficient may be biased upward.

While I do not observe enough employees in multiple years to look at those who switch plans, I can look at the timing of spending for all employees in the sample. Figure 2.7 suggests that spending is relatively flat throughout the year, with no increase in the beginning nor at the end of a plan year. As a robustness check, I rerun the main regression with the log of total expenditures in the middle eight months of the plan year, excluding the first two and last two months, as the dependent variable. The results in Table 2.11 are very similar to the main results (though, not surprisingly, smaller in magnitude, as the spending is over a shorter period) suggesting that the results are not biased by delayed or accelerated care.

Coverage and marital status For most dual-earner couples, I observe only one of the spouses' choice set of health plans. The h' I calculate for these families may not represent the relevant choice between the first- and second-best plans, if the second-best plan was offered by the other spouse's employer. I therefore estimate the two-part regression model separately for unmarried people with single coverage, married people with single coverage, unmarried employees opting for family coverage, and married employees in family coverage. The results in Table 2.12 largely match the main estimation: more generous coverage leads to more spending for each group in non-HMO plans, and the coefficient on $log(h')$ is never significantly greater than zero statistically.

Validity of the theoretical model of plan choice The theoretical model predicts that employees who plan to spend more than h' should enroll in the more generous plan, while those planning to spend less than h' will find that the lower premium paid in the less generous plan is a better fit for their needs. Expected spending, even over a short period of time, is uncertain, but if I assume that consumers have perfect foresight, then I should see that employees who ultimately spent more than h' should enroll in the more generous plan, and vice versa. About 55 percent of the sample enrolls in the plan that is a better fit according to the theoretical model; these employees who select a non-HMO have a greater response to the moral hazard incentive than those who do not select the plan that the model predicts (Table 2.13). The marginal effect of $log(h')$ is actually negative, though insignificant, for both HMO and non-HMO enrollees whose plan choice fit the theoretical model, but positive and both statistically and substantively significant for HMO enrollees who do not fit the model.

2.8 Conclusion

In this paper, I suggest a new method for determining whether adverse selection and moral hazard affect the market for employer-sponsored health insurance, using the variation in health plan offers across employers. By controlling for both the absolute generosity of a selected plan and the level of spending above which the employee prefers the more generous option, I am able to separately identify the effects of moral hazard from adverse selection, respectively.

My results indicate that the moral hazard incentive significantly influences the employer health insurance market: the more (absolutely) generous the selected plan, the more the enrollee tends to spend on medical care. This moral hazard incentive is strongest in non-HMO plans, as expected. There is no evidence of adverse selection, however; the correlation between medical expenditures and the level of spending that makes the more generous plan optimal is never significantly greater than zero. Enrollees in more generous plans are also no more likely to have an adverse health outcome, even when they face a tradeoff between premiums and generosity, which suggests that there is no self-selection by risk type between plans of the same generosity. There is some evidence that adverse selection occurs between plan types, as HMO enrollees are less likely to have (or to cover family members with) a chronic condition or be over the age of 50 than enrollees in non-HMO plans, but because expenditures across plan types are not directly comparable, this evidence comes only from the direct measures of risk type, whereas the adverse selection "death spiral" is likely to arise from differences in expenditure (Feldman and Dowd, 1991).

The methodology used here can be applied to any insurance market where the quantity of coverage is variable. Most insurance plans do not fully cover the damage caused by a risky event, and even those that do may vary the deductible or other cost-sharing involved in paying out claims, so there are very few markets where this does not apply. To apply the distinction between absolute and relative generosity, the researcher would need data on not only the selected plan but also the other plans available in the market. This may be difficult in some markets where customers are allowed to shop around (and do not rely on an employer to do the shopping for them), but rules against insurance plans crossing state or national boundaries often make a market finite. Moreover, the calculated financial generosity may actually be a better measure of the perceived value to the enrollee in insurance markets other than health insurance, which is complicated by personal relationships with care providers.

While the moral hazard finding matches the qualitative conclusion of the RAND Health Insurance Experiment and most subsequent studies, the lack of adverse selection stands in contrast to most single-firm studies (though, importantly, not structural estimates like Cardon and Hendel (2001), which use data from a similarly broad sample of firms) and the prediction of the Rothschild-Stiglitz theoretical model, especially given the information gap between insurer and enrollee. This paper's results should be interpreted with some skepticism, due to limitations both with the data and the model.

Using health insurance plan data from a nationwide sample of firms allows for greater confidence in the external validity of the results, but has the disadvantage of incomplete and possibly inaccurate information on the quality of each plan. In this study, I focus entirely on the plan's implied financial burden on the employee for physician, hospital, and prescription drug expenditure, but other factors almost certainly influence the decision to select one plan over another.[33] One very important missing plan parameter, especially as HMOs and PPOs have come to dominate the market, is the breadth of the provider network under one's plan — in particular, whether the employee may keep seeing her, or her family member's, preferred care

[33] Finkelstein and McGarry (2006) and Fang, Keane, and Silverman (2008) suggest that self-selection occurs along other dimensions, including risk tolerance and cognition, and future work should take these important dimensions into account. Unfortunately, MEPS lacks direct measures of risk tolerance and cognition like those used in these two studies from the Health and Retirement Study. MEPS can, however, be linked to the National Health Interview Survey, which includes self-reported information on smoking, alcohol use, exercise, and use of preventive care.

provider under alternative plans. Employees may also desire coverage for types of care that have their own coinsurance schedule, if they're covered at all, such as mental health or home health care. Lacking information on the provider network and the coinsurance schedule for other care types, I cannot determine the extent to which employees are willing to trade off greater financial generosity in basic care for the ability to see a specific care provider or to be covered for a more specialized care type; this may partly explain the prevalence of employees who opt for a plan that has a higher premium and lower measured financial generosity than some alternative. More generally, I am forced to make the strong assumption that unobserved differences in (real or reputational) plan quality have the same impact on plan choice throughout the health distribution, whereas in single-firm studies, plan-specific quality is differenced away.[34]

It is also possible that there is adverse selection, but it does not manifest as predicted by the model. Aside from adverse selection along other plan elements, such as network breadth (sicker patients enrolling in broader, more expensive networks), mental health, or long term care, it is also possible that h' is not the relevant margin between two plans. One weakness of the model is the assumption that the employee chooses between exactly two plans of the same type; many employees have the choice of three or more plans, which means that there may be multiple h' thresholds along the distribution of potential health spending.[35] The model also assumes that the consumer has complete knowledge of the plan's financial parameters, but the effort cost of differentiating between similar and complex plans may be prohibitive, so that

[34]Only comparing plans of the same type (HMO or non-HMO) likely reduces the importance of unobserved plan differences, though this may be less true as non-HMO plans, especially PPOs, mimic the successes of HMO plans. Eliminating cross-type comparisons also reduces the sample to the subset of employees offered multiple plans of the same type; while I weight the observations to account for differences in observable characteristics between this subsample and the full MEPS sample, the weights may not fully account for how the relationship between the plan choice and care-seeking decisions may be different in this subsample.

[35]Risk aversion actually biases in the direction of finding evidence of adverse selection, so the assumption of risk neutrality is not qualitatively problematic; see note 12.

workers do not sort into the plan that is the best fit for their needs, even ex-ante. I also assume that the ex-post realization of medical expenditures is a good proxy for an enrollee's risk type, but biases may arise if employees systematically under- or overestimate their medical needs before the year.

Finally, I cannot estimate by how much moral hazard or adverse selection reduce social welfare, only whether they exist or not. The implication of my finding of insignificant adverse selection suggests that the social welfare impact is small, but I cannot begin to estimate exactly how small unless I can trace out demand and supply curves.[36]

While the reader is cautioned about taking these results at face value, adverse selection appears to be restricted to sorting between plan types, not plans within a type (e.g., HMO plan versus other HMO plans). Using managed care as a sorting mechanism allows low-risk consumers to get actuarially fair insurance without attracting the highest risks. At the same time, the employer-sponsored insurance system effectively subsidizes the less healthy, as the employee's contribution to the premium does not depend on her health status, and the tax system favors compensation in the form of health benefits. Feldman and Dowd (2000) suggest that the combination of sorting by managed care and subsidization of the unhealthy through the employer system strikes the right balance between efficiency and fairness. If there is no further adverse selection beyond managed care versus more traditional plans, then that balance is not upset.

The recent effort by the Obama Administration and Congress to reform the health insurance market for individuals, and for small business employees that enter the market without the benefit of diverse risk pools, use the large- and medium-firm

[36] The Einav-Finkelstein-Cullen (2008) approach appears promising, but in the national health insurance market during the 1996-2001 period studied here, there is no obvious exogenous price variation required to estimate marginal and average cost curves.

employer-sponsored insurance market as a model. One concern is that proposals that increase the amount of private information, including preventing screening for pre-existing conditions and capping the degree to which premiums can vary based on age, could result in adverse selection. The results in this paper suggest that, if the composition of the risk pool of small business employees and individuals is similar to employees in the existing employer insurance system, adverse selection will have only a minimal impact.

Moral hazard, on the other hand, may be inevitable, as even the most tightly managed health plan will lower the out-of-pocket cost below its actual marginal cost. Still, this paper shows that managed care plans have diminished the moral hazard incentive significantly over non-managed plans. Further monitoring of care, especially in non-HMO plans, will likely reduce costs, and allow for more affordable care for all.

2.9 Appendix: Proof That Expected Value of Expenditures Increases with the Threshold Spending Value h'.

Let $f(h)$ and $F(h)$ represent the probability density function and cumulative density function of spending, h, respectively. The researcher does not observe the threshold level of spending, h', that separates enrollees in high-generosity plan G from enrollees in low-generosity plan C, but does observe whether G or C was selected, and the ex-post realization of expenditures h.

The expected value of h for C enrollees is

$$E(h|h<h') = \frac{1}{F(h')}\int_{-\infty}^{h'} hf(h)dh.$$

Taking the derivative of $E(h|h<h')$ with respect to h' yields

$$\frac{\partial E(h|h<h')}{\partial h'} = -\frac{f(h')}{F(h')^2}\int_{-\infty}^{h'} hf(h)dh + \frac{1}{F(h')}(h'f(h')),$$

with the latter portion due to the fundamental theorem of calculus.

For the derivative to be positive,

$$\frac{h'f(h')}{F(h')} > \frac{f(h')}{F(h')^2}\int_{-\infty}^{h'} hf(h)dh.$$

Simplifying, $f(h')$ and $1/F(h')$ cancel from both sides, leaving

$$h' > \frac{1}{F(h')}\int_{-\infty}^{h'} hf(h)dh = E(h|h<h'),$$

which must be true, as the expected value must be less than the upper bound for that interval.

The expected value of h for G enrollees is

$$E(h|h>h') = \frac{1}{1-F(h')}\int_{h'}^{\infty} hf(h)dh.$$

Taking the derivative of $E(h|h>h')$ with respect to h' yields

$$\frac{\partial E(h|h<h')}{\partial h'} = \frac{f(h')}{(1-F(h'))^2}\int_{h'}^{\infty} hf(h)dh + \frac{1}{1-F(h')}(-h'f(h')),$$

with the latter portion due to the fundamental theorem of calculus.

For the derivative to be positive,

$$\frac{h'f(h')}{1-F(h')} < \frac{f(h')}{(1-F(h'))^2}\int_{h'}^{\infty} hf(h)dh.$$

Simplifying, $f(h')$ and $1/(1-F(h'))$ cancel from both sides, leaving

$$h' < \frac{1}{1-F(h')}\int_{h'}^{\infty} hf(h)dh = E(h|h > h'),$$

which must be true, as the expected value must be greater than the lower bound for that interval. $\square$

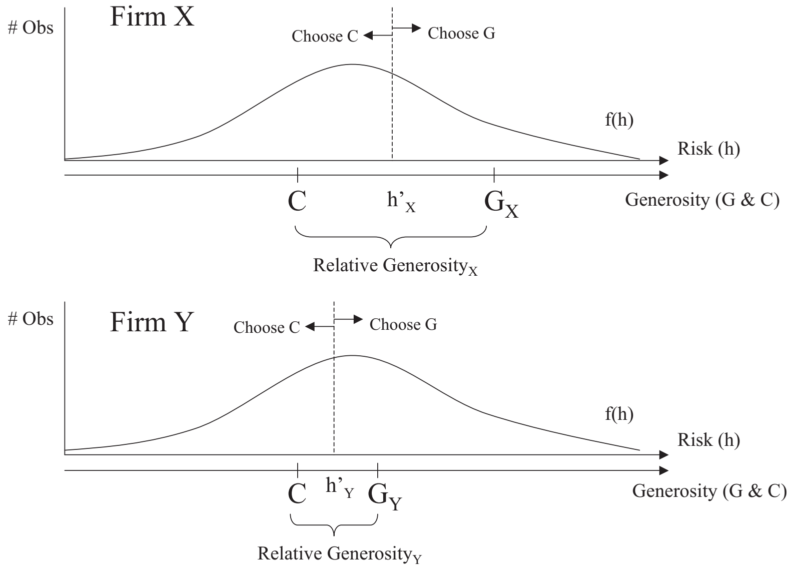


Figure 2.1: Risk Distribution and Generosity at Two Hypothetical Firms With the Same Less-Generous Plan (C)

2.10 Figures and Tables

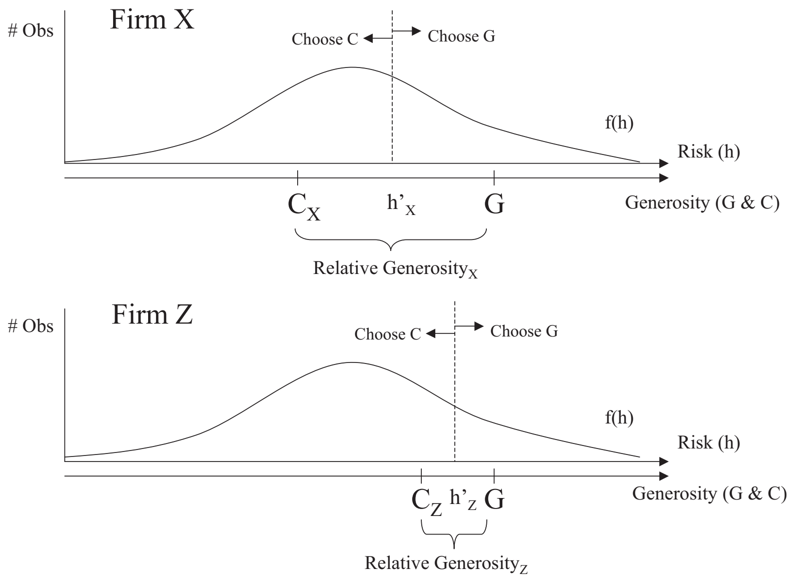


Figure 2.2: Risk Distribution and Generosity at Two Hypothetical Firms With the Same More-Generous Plan (G)

Figure 2.3: Total Expenditure by Threshold Spending Level (h') Decile and Plan Type

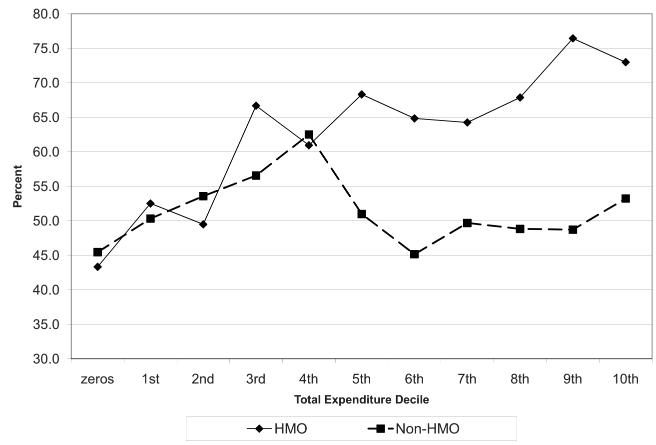


Figure 2.4: Probability of Choosing the More Generous Plan, By Total Expenditure

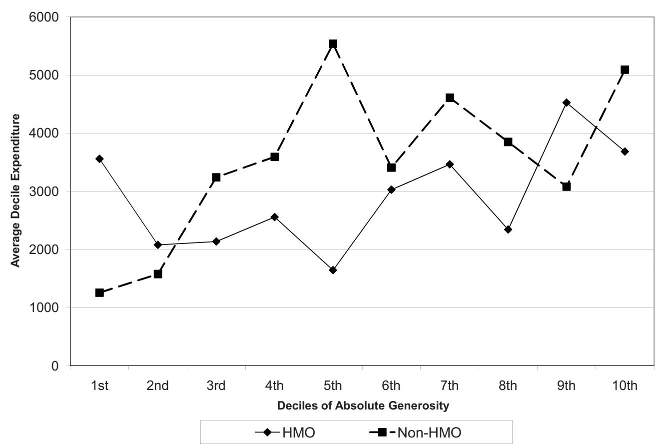


Figure 2.5: Total Expenditure by Absolute Generosity Decile and Plan Type

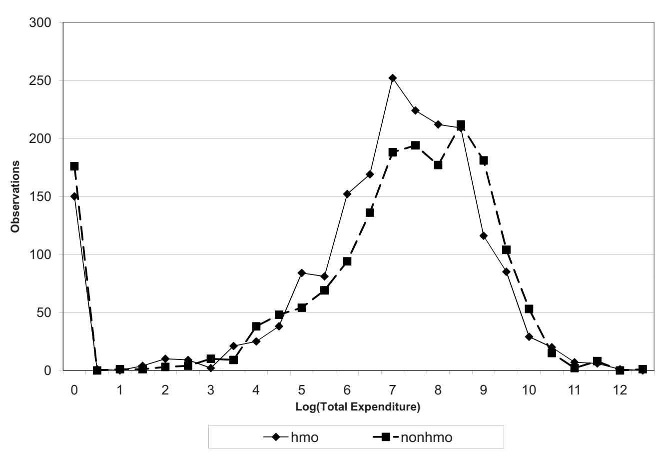


Figure 2.6: Distribution of Total Expenditures, by Plan Type

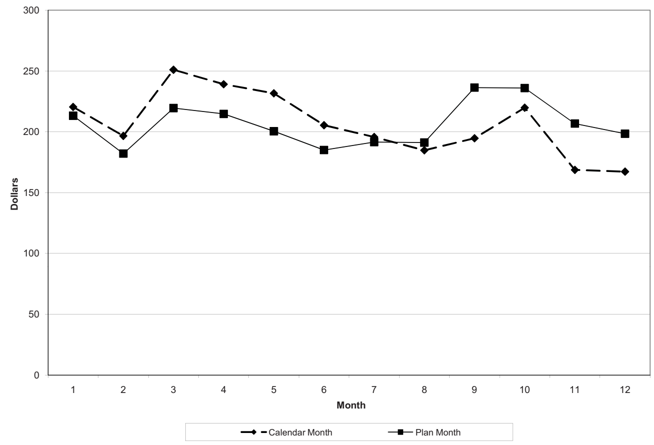


Figure 2.7: Average Total Medical Expenditure by Plan Month and Calendar Month

Table 2.1: Non-Financial Generosity Differences Between Plan Pairs

	HMO Pairs		Non-HMO Pairs	
	Selected Had	Rejected Had	Selected Had	Rejected Had
Deductible	1.4	1.0	25.0	16.2
Lifetime Maximum	1.5	1.8	1.6	2.1
Annual Plan-Paid Max	0.9	0.7	0.8	0.7
OOP Annual Limit	9.9	10.6	7.2	5.4
Refuse for Pre-Existing	1.2	1.6	4.7	4.3
Wait Period for Pre-Existing	1.7	2.5	3.9	4.7
Specialist Referrals	3.4	7.2	11.2	8.5
Mammograms	1.8	1.9	1.3	0.6
Chiropractor	14.1	8.4	8.0	31.5
Rx Included	1.9	1.7	2.4	1.2
Dental Included	7.3	8.0	10.6	8.4
Orthodontal Included	6.9	6.2	4.9	9.3
Inpatient Mental Incl.	2.3	1.9	1.9	1.7
Outpatient Mental Incl.	1.8	1.9	1.0	1.0
Substance Abuse	1.6	1.3	1.0	0.3
Life Insurance Incl.	0.3	0.3	0.4	0.1
Disability Insurance Incl.	0.1	0.2	0.3	0.2
Physicals	1.6	1.2	25.1	8.1
Pap Smears	1.6	1.0	2.2	1.4
Prenatal Coverage	1.0	0.6	3.8	0.8
Adult Immunizations	1.9	1.6	5.8	3.7
Child Immunizations	1.6	1.3	1.3	0.5
Well-Infant Care	1.6	1.5	2.2	1.0
Well-Child Care	1.4	1.2	5.7	2.7
Other Non-Physician	5.3	3.3	5.1	1.4
Nursing Home	1.9	1.6	1.1	1.0
Home Health	1.5	0.8	27.9	5.3

Note: Table entry is the percent of plans with that characteristic.

Table 2.2: Employee Contributions, Premiums, and Choosing the More Generous Plan

	All		HMO		Non-HMO	
	Plan Pairs	% More Gen	Plan Pairs	% More Gen	Plan Pairs	% More Gen
Employee Contributions						
More Gen Has Lower Contrib	1437	61.5	690	64.5	747	58.8
Both Zero Contribution	529	56.1	396	53.8	133	63.2
Same Contribution	227	65.2	134	65.7	93	64.5
More Gen Has Higher Contrib	1490	51.3	686	64.1	804	40.4
Total Premiums						
More Gen Has Lower Premium	1718	61.5	909	64.0	809	58.6
Same Premium	234	63.7	105	61.9	129	65.1
More Gen Has Higher Premium	1731	51.4	892	60.4	839	41.7
Total	3683	56.9	1906	62.2	1777	51.1

Table 2.3: Summary Statistics

	Unweighted		Weighted	
	HMO	Non-HMO	HMO	Non-HMO
Total Expenditure	2900.73	3522.39	2831.41	3293.58
	(6864.51)	(8351.29)	(6860.87)	(9291.98)
Physician Expenditure	1092.86	1428.92	1067.10	1210.09
	(1908.35)	(2645.68)	(1911.67)	(2287.88)
Hospital Expenditure	1199.02	1372.99	1224.20	1474.16
	(5658.72)	(6378.78)	(5738.07)	(7784.03)
Prescription Drug Expenditure	589.55	698.87	511.45	579.35
	(1304.39)	(1238.61)	(1364.44)	(1001.55)
Chose HMO	1	0	1	0
	(0)	(0)	(0)	(0)
Chose PPO	0	0.90	0	0.87
	(0)	(0.30)	(0)	(0.33)
Chose FFS	0	0.10	0	0.13
	(0)	(0.30)	(0)	(0.33)
Chose More Generous	0.62	0.51	0.58	0.52
	(0.48)	(0.50)	(0.49)	(0.50)
Absolute Generosity	0.83	0.66	0.80	0.63
	(0.12)	(0.17)	(0.17)	(0.20)
Relative Generosity, More Generous	0.04	0.14	0.05	0.15
	(0.10)	(0.15)	(0.13)	(0.16)
Relative Generosity, Less Generous	0.05	0.14	0.06	0.16
	(0.09)	(0.13)	(0.11)	(0.17)
Relative Premium, More Generous	394.24	541.06	394.24	541.06
	(713.48)	(670.00)	(713.48)	(670.00)
Relative Premium, Less Generous	284.87	505.64	284.87	505.64
	(442.13)	(644.83)	(442.13)	(644.83)
Threshold Spending Level (h')	17563	14219	17563	14219
	(48770)	(41881)	(48770)	(41881)
Number of observations	1907	1778	1877	1738

Note: Standard deviations in parentheses. Relative premium means are from the sample where the employee contribution for the more generous plan is larger than the less generous plan.

Table 2.4: Selected Summary Statistics for Weighted Full Sample and Estimation Sample, Unweighted and Weighted

	Full MEPS-HC Weighted Sample of Insured Workers	Estimation Sample (Unweighted)	Estimation Sample (Weighted)
White	0.769	0.663	0.770
	(0.421)	(0.473)	(0.421)
Black	0.112	0.158	0.115
	(0.315)	(0.365)	(0.319)
Hispanic	0.080	0.142	0.082
	(0.271)	(0.349)	(0.275)
Self Employed	0.044	0.011	0.047
	(0.205)	(0.104)	(0.212)
Establishment Size			
1-10	0.138	0.063	0.147
	(0.344)	(0.242)	(0.354)
11-25	0.111	0.075	0.104
	(0.314)	(0.263)	(0.305)
26-50	0.115	0.092	0.104
	(0.319)	(0.289)	(0.305)
51-100	0.130	0.127	0.139
	(0.336)	(0.333)	(0.346)
101+	0.429	0.571	0.431
	(0.495)	(0.495)	(0.495)
Industry (partial list)			
Construction	0.047	0.020	0.065
	(0.212)	(0.138)	(0.247)
Manufacturing	0.201	0.135	0.208
	(0.401)	(0.342)	(0.406)
Sales	0.130	0.056	0.100
	(0.336)	(0.229)	(0.300)
F.I.R.E.	0.070	0.054	0.067
	(0.255)	(0.226)	(0.250)
Professional Services	0.257	0.318	0.263
	(0.437)	(0.466)	(0.440)
Public Administration	0.077	0.222	0.081
	(0.267)	(0.416)	(0.273)
After-Tax Family Income			
Less than $25k	0.230	0.183	0.237
	(0.421)	(0.387)	(0.426)
$25k-50k	0.356	0.348	0.324
	(0.479)	(0.477)	(0.468)
$50k-75k	0.242	0.274	0.253
	(0.428)	(0.446)	(0.435)
$75k-100k	0.098	0.107	0.109
	(0.297)	(0.309)	(0.311)
$100k+	0.074	0.088	0.077
	(0.262)	(0.284)	(0.267)
Observations	32395	1996	1961

Note: Standard deviations in parentheses.

Table 2.5: Two-Part Regression of Total Expenditures on Absolute Generosity and Log Threshold Spending Level

	Logit		Log Linear		Marginal Effect	
HMO Enrollees						
Absolute Generosity	0.939		0.330		177.05	
	(1.337)		(0.379)		(202.27)	
Log(h')	0.018		-0.0092		3.306	
	(0.016)		(0.0077)		(4.098)	
N	1870		1732			
R^2	0.197		0.113			
Non-HMO Enrollees						
Absolute Generosity	2.742	***	2.112	***	467.38	***
	(1.015)		(0.421)		(149.45)	
Log(h')	-0.009		0.0221	**	-1.530	
	(0.021)		(0.0086)		(3.669)	
N	1719		1572			
R^2	0.336		0.226			

Note: Regressions weighted to account for non-random missing observations in HC-IC Linked File sample.
*** - Significantly different from zero at the 99% confidence level ** - 95% confidence level * - 90% confidence level

Table 2.6: Two-Part Regression of Expenditures on Generosity and Premium Measures, by Type of Care

	Marginal Effects		
	HMO	Non-HMO	
Ln(Physician Expenditures)			
Absolute Generosity	104.36	475.25	***
	(102.84)	(112.16)	
Log(h')	0.864	2.214	
	(1.837)	(3.023)	
N	1876	1726	
R^2 Logit	0.154	0.241	
R^2 LogLin	0.115	0.210	
Ln(Hospital Expenditures)			
Absolute Generosity	19.85	1531.00	***
	(651.10)	(445.38)	
Log(h')	-18.483	15.117	
	(12.100)	(10.140)	
N	1876	1738	
R^2 Logit	0.072	0.201	
R^2 LogLin	0.104	0.206	
Ln(Prescription Drug Expenditures)			
Absolute Generosity	88.97	192.27	***
	(60.72)	(37.75)	
Log(h')	0.993	0.446	
	(1.155)	(1.093)	
N	1876	1726	
R^2 Logit	0.206	0.317	
R^2 LogLin	0.177	0.139	

Note: Regressions weighted to account for non-random missing observations in HC-IC Linked File sample.
*** - Significantly different from zero at the 99% confidence level ** - 95% confidence level * - 90% confidence level

Table 2.7: Two-Part Regression of Total Expenditures on Absolute Generosity and Log Threshold Spending Level, By Employee Contribution Scheme and h' Level

	Marginal Effects					
	Same Employee Contributions		Different Employee Contributions		More Generous Has Higher Premium ($0 < h' < max$)	
HMO Enrollees						
Absolute Generosity	579.592		-37.817		70.240	
	(400.068)		(309.188)		(629.713)	
Log(h')	2.707		6.556		2.027	
	(6.823)		(6.427)		(38.281)	
N	1080		745		518	
R^2 Logit	0.171		0.373		N/A	
R^2 LogLinear	0.170		0.209		0.213	
Non-HMO Enrollees						
Absolute Generosity	1595.861	**	596.516	***	463.467	*
	(781.468)		(183.753)		(277.813)	
Log(h')	-10.884		2.230		3.998	
	(19.896)		(4.164)		(28.663)	
N	462		1210		749	
R^2 Logit	N/A		0.344		0.419	
R^2 LogLinear	0.358		0.229		0.292	

Note: Regressions weighted to account for non-random missing observations in HC-IC Linked File sample.
*** - Significantly different from zero at the 99% confidence level ** - 95% confidence level * - 90% confidence level

Table 2.8: Poisson Regression of Visits on Absolute Generosity and Log Threshold Spending Level

	Total Visits		Physician Visits		Hospital Nights	
HMO Enrollees						
Absolute Generosity	7.44	***	7.27	***	0.16	
	(2.61)		(2.73)		(0.88)	
Log(h')	0.041		0.020		0.020	
	(0.052)		(0.042)		(0.019)	
N	1877		1877		1877	
Non-HMO Enrollees						
Absolute Generosity	16.21	***	13.73	***	4.00	***
	(2.26)		(2.18)		(0.93)	
Log(h')	0.066		0.046		0.007	
	(0.065)		(0.059)		(0.022)	
N	1738		1738		1738	

Note: Entry in each cell is marginal effect from Poisson regression; bootstrapped standard errors in parentheses. Regressions weighted to account for non-random missing observations in HC-IC Linked File sample.
*** - Significantly different from zero at the 99% confidence level ** - 95% confidence level * - 90% confidence level

Table 2.9: OLS Regressions of Numbers of Family Members with an Adverse Health Outcome

	Number of Family Members with					
	Chronic Conditions	Physical Fair/Poor	Mental Fair/Poor	Any Limitation	Age 0 to 2	Age 50+
HMO Enrollees						
Absolute Generosity	-0.180	-0.142	0.047	0.138	0.030	0.190
	(0.194)	(0.123)	(0.046)	(0.162)	(0.075)	(0.165)
Log(h')	0.0038	0.0033 **	0.0009	0.0021	-0.0009	0.0006
	(0.0024)	(0.0015)	(0.0007)	(0.0021)	(0.0013)	(0.0032)
N	1810	1810	1810	1810	1810	1810
R^2	0.165	0.102	0.205	0.153	0.112	0.195
Non-HMO Enrollees						
Absolute Generosity	0.411 ***	0.182 ***	0.010	0.220 ***	0.028	0.050
	(0.112)	(0.065)	(0.042)	(0.083)	(0.048)	(0.169)
Log(h')	0.0010	0.0037 **	0.0006	0.0047 **	0.0010	0.0030
	(0.0026)	(0.0018)	(0.0006)	(0.0019)	(0.0009)	(0.0032)
N	1667	1667	1667	1667	1667	1667
R^2	0.124	0.139	0.055	0.103	0.062	0.188
All Enrollees						
HMO (0/1)	-0.061	-0.012	0.006	0.041	0.029 *	-0.071
	(0.048)	(0.025)	(0.011)	(0.034)	(0.016)	(0.051)
N	3477	3477	3477	3477	3477	3477
R^2	0.091	0.093	0.075	0.080	0.076	0.145

Note: "All Enrollees" regression includes absolute generosity and its square, log(h'), and firm and plan characteristics. All regressions weighted to account for non-random missing observations in HC-IC Linked File sample.
*** - Significantly different from zero at the 99% confidence level ** - 95% confidence level * - 90% confidence level

Table 2.10: Two-Part Regression of Total Expenditures on Absolute Generosity and Log Threshold Spending Level, Offered Exactly Two Plans

	Logit		Log Linear		Marginal Effect	
HMO Enrollees						
Absolute Generosity	2.758	*	-0.458		755.72	***
	(1.426)		(0.561)		(218.48)	
Log(h')	0.032		-0.0303	**	8.642	**
	(0.037)		(0.0138)		(4.221)	
N	399		369			
R^2	0.225		0.291			
Non-HMO Enrollees						
Absolute Generosity	1.652		1.658	***	186.79	
	(1.283)		(0.603)		(143.96)	
Log(h')	-0.060		0.0311	*	-6.801	*
	(0.047)		(0.0160)		(3.637)	
N	414		373			
R^2	0.531		0.244			

Note: Regressions weighted to account for non-random missing observations in HC-IC Linked File sample.
*** - Significantly different from zero at the 99% confidence level ** - 95% confidence level * - 90% confidence level

Table 2.11: Two-Part Regression of Total Expenditures in the Middle Eight Months on Absolute Generosity and Log Threshold Spending Level

	Logit		Log Linear		Marginal Effect	
HMO Enrollees						
Absolute Generosity	0.406		0.479		39.31	
	(1.113)		(0.410)		(93.71)	
Log(h')	0.002		-0.0097		0.242	
	(0.027)		(0.0082)		(2.426)	
N	1325		1086			
R^2	0.146		0.125			
Non-HMO Enrollees						
Absolute Generosity	2.123	**	1.263	***	231.74	*
	(1.047)		(0.435)		(137.86)	
Log(h')	0.027		0.0178	**	2.970	
	(0.026)		(0.0084)		(3.085)	
N	1322		1055			
R^2	0.239		0.295			

Note: Regressions weighted to account for non-random missing observations in HC-IC Linked File sample.
*** - Significantly different from zero at the 99% confidence level ** - 95% confidence level * - 90% confidence level

Table 2.12: Two-Part Regression of Total Expenditures on Absolute Generosity and Log Threshold Spending Level, by Coverage and Marital Status

	Marginal Effects					
	Single Coverage			Family Coverage		
	Unmarried		Married	Unmarried	Married	
HMO Enrollees						
Absolute Generosity	800.337		84.261	-433.585	2795.770	***
	(804.028)		(623.377)	(929.763)	(673.015)	
Log(h')	-2.526		4.123	-7.585	7.797	
	(10.819)		(12.836)	(17.435)	(7.729)	
N	445		283	184	814	
	0.244		N/A	N/A	0.652	
R^2	0.223		0.269	0.325	0.123	
Non-HMO Enrollees						
Absolute Generosity	1240.244	**	676.847	822.619	1853.478	***
	(605.350)		(1265.150)	(4614.680)	(501.334)	
Log(h')	-7.664		-0.570	-18.208	11.882	
	(14.164)		(21.630)	(49.359)	(9.057)	
N	376		208	80	1001	
	0.363		0.404	N/A	0.579	
R^2	0.660		0.458	0.849	0.175	

Note: Regressions weighted to account for non-random missing observations in HC-IC Linked File sample.
*** - Significantly different from zero at the 99% confidence level ** - 95% confidence level * - 90% confidence level

Table 2.13: Two-Part Regression of Total Expenditures on Absolute Generosity and Log Threshold Spending Level, by Whether the Predicted Plan was Selected

	Marginal Effects			
	Predicted Plan		Not Predicted Plan	
HMO Enrollees				
Absolute Generosity	239.010		178.401	
	(401.861)		(312.134)	
Log(h')	-7.960		13.543	**
	(9.154)		(5.444)	
N	928		840	
	N/A		0.231	
R^2	0.157		0.146	
Non-HMO Enrollees				
Absolute Generosity	610.266	***	344.289	
	(186.365)		(291.500)	
Log(h')	-4.239		0.821	
	(4.669)		(7.527)	
N	975		687	
	0.388		0.400	
R^2	0.236		0.338	

Note: Regressions weighted to account for non-random missing observations in HC-IC Linked File sample.
*** - Significantly different from zero at the 99% confidence level ** - 95% confidence level * - 90% confidence level

CHAPTER III

Decomposing the Growing Disparity in Health Insurance Coverage between Hispanics and Non-Hispanic Whites

3.1 Introduction

Between 1987 and 2008, the number of Americans without health insurance grew from 31.0 million to 46.3 million, or from 12.9 percent of the population to 15.4 percent (DeNavas-Walt et al, 2009).[1] This expansion in the ranks of the uninsured, despite public coverage expansions and a macroeconomy that was generally booming during this time, is a concern for policymakers, as lower coverage rates could result in more exposure to medical and financial risk and less access to necessary health care. The growing uninsured rate has therefore received a great deal of attention both in the mainstream media and by academic journals in the fields of medicine, public health, and economics.

Less recognized has been the fact that for many population subgroups, the uninsured rate has remained constant, or actually declined, over this time. In 1983, 19 percent of children under age 18 were uninsured; in 2007, this rate had fallen to 17 percent, aided by expansions to Medicaid and SCHIP.[2] The number of elderly (over 65) uninsured remains negligible, due to pervasive Medicare coverage. The uninsured rate fell even among some groups of nonelderly adults, who are largely

[1]These figures are from the Current Population Survey.
[2]These figures are from the Survey of Income and Program Participation.

ineligible for public coverage. Though the uninsured rate for Blacks remains almost ten percentage points higher than Whites, the proportion of Blacks without health insurance fell slightly, mirroring the small decline in uninsured Whites. Asians and Native Americans are also slightly less likely to be uninsured than they were in the 1980's.

For Hispanics, though, the health insurance problem has only grown worse. Already sixteen percentage points higher than non-Hispanic Whites in the early 1980's, the uninsured rate for Hispanics increased to 38.7 percent in 2007, a gap of nearly 25 percentage points with non-Hispanic Whites. Whereas the uninsured rates for Whites and Blacks have fluctuated, with increases in the early 1990's and early 2000's sandwiched by declines during the late 1980's, late 1990's, and mid-2000's, the Hispanic uninsured rate has grown almost monotonically throughout this period. While much of the growth in the number of uninsured Hispanics has been concentrated in the non-native population, U.S.-born Hispanics are also more likely to be uninsured compared to twenty-five years ago.

In a previous paper (Rutledge and McLaughlin, 2008), we analyze the trends in the uninsured rate across racial and ethnic groups, and within the Hispanic group, in more detail. We find that the overall uninsured rate would have been essentially flat over the last two decades if the Hispanic noncitizen proportion of the population had remained constant, and actually would have declined if the proportion Hispanic (without regard to citizenship) had remained unchanged. We also suggest that changes in the composition of the Hispanic population, particularly the growth in the likelihood of having less than a high school education relative to other ethnic groups, have contributed to the growing gap between Hispanics and non-Hispanics.

This paper extends our previous work on uninsured rate trends by formally analyzing how differences in citizenship status, educational attainment, public coverage eligibility, income, and labor market outcomes between Hispanics and non-Hispanic Whites have contributed to the divergence in uninsured rates between the two groups. We examine not only how observable differences contribute to point-in-time coverage gaps, but also how these characteristics have contributed to the *growth* in the gap over two decades.

We use two methods to decompose the contribution of observable variables to the diverging uninsured rates: 1) the estimation of a regression-adjusted trend line, from the coefficients of the interaction of the Hispanic indicator with year dummies, and determining how this trend line changes when adding each variable sequentially, and 2) a Blinder-Oaxaca-style decomposition of the trend from a base period to the end of our sample. We find that, consistent across our decomposition methods, differences in citizenship and educational attainment between Hispanics and non-Hispanic Whites account for each about one third of the growth in the uninsured rate gap, and a similar amount of the annual point-in-time gaps, more than any other of the characteristics we observe. Still, in each of our decomposition methods, more than half of the divergence in uninsured rates from 1984 to 2007 remains unexplained.

We also analyze the trends in the rate of coverage separately by insurance source. The gap in public coverage has been reliably countercyclical, but after adjusting for personal characteristics, Hispanic and white non-Hispanic public coverage rates are essentially equal and constant over time. Adjusting for observable differences between the groups, Hispanics and non-Hispanic Whites were also just about as likely to have private coverage through their own employer. There was, however, a sharp decline for Hispanics in coverage through another family member, the one insurance source category where a significant portion of the gap with non-Hispanic

Whites remains unexplained even after controlling for observable characteristics.

Our results indicate that the high, and growing, uninsured rate among Hispanics cannot be attributed to lower educational attainment, citizenship rates, and income only. In 2007, 9 million Hispanics between the ages of 18 and 55 were uninsured, according to our sample. If, instead, the uninsured rate gap had remained constant at its 1984 level, 6.6 million nonelderly Hispanic adults would have been uninsured. Based on our estimates, only 0.7 to 1.4 million of this 2.4 million difference can be explained by differences in observable characteristics, leaving an unaccounted-for increase of more than one million uninsured Hispanic adults since 1984.

Because much of the increase in the uninsured rate for Hispanics relative to non-Hispanic Whites remains unexplained, we offer some potential reasons for the widening gap, among them job composition, social networks, risk tolerance, and health quality. Government, advocacy groups, and private individuals must address these issues, in addition to improving labor market and educational outcomes, in order to assure equitable access to health care and protection from high medical expenditures to all Americans, regardless of ethnicity.

3.2 Previous Literature

The concentration of the growth in the uninsured within the growing Hispanic population is well-established in the literature. Carrasquillo, Himmelstein, et al (1999) analyzed trends in the uninsured rate from 1989 to 1996 and found that Hispanics account for 36.4 percent of the increase in the number of uninsured during that time. Rutledge and McLaughlin (2008) calculate that the growth in the Hispanic share of the population accounts for 31 percent of the increase in the number of uninsured adults.

Though we focus on how the gap in uninsured rates between Hispanics and non-Hispanics has changed over time, we begin by decomposing how personal characteristics contribute to the gap within each year. Monheit and Vistnes (2000), using a similar decomposition method to ours, find that wages, family income, and education can account for most of the gap between Hispanic and white men in 1987 and 1996, though a substantial portion remains unexplained.

Two other studies suggest separate analysis and decomposition by insurance source. Waidmann, Garrett, and Hadley (2000) decompose how demography, human capital, labor and health care market changes, and state policies affect offers and takeup of employer sponsored insurance (ESI), and find that most of the gap is explained by English language proficiency and education. Buchmueller, LoSasso, Lurie, and Dolfin (2007) sequentially add human capital and employer variables (similar to our Method 1) to regressions of ESI eligibility, offers, and takeup, and find that only the gap between natives and non-citizens in offer rates remains substantially unexplained.

Other studies make clear that the Hispanic population is not uniform, suggesting separate analysis by country of origin. In descriptive work, Shah and Carrasquillo (2006) examine trends over a twelve-year period among Latino subgroups compared to non-Hispanic Whites, and find that Mexicans were hardest hit by Medicaid cuts in the 1990's and employer coverage declines in the early 2000's. Berk, Albers, and Schur (1996) find similar results between 1977 and 1992. Fronstin, Goldberg and Robins (1987) use Blinder-Oaxaca decomposition to compare Mexicans to Puerto Ricans and Cubans, both of whom have lower uninsured rates, and find that most of the difference can be explained by wages, age, education, industry, and firm size.

Besides county of origin, Hispanics differ in other ways, including nativity, gender, and employment, that influence their coverage probability. Hamilton, Hummer, You, and Padilla (2006) separate Mexican-Americans by generation; while second-generation Mexican-Americans have much lower coverage rates than Blacks, third-generation Mexican-Americans actually have higher rates than Blacks and are similar to Whites. Alegria et al. (2005) find gender differences for Latinos, with females far more likely to have public coverage and males more likely to have private coverage, and that only immigrants within their first five years in the U.S. have substantially different coverage rates than U.S. natives. Schur and Feldman (2000) look at firm size and industry, and find that even within a finely-defined industry, Hispanics have lower coverage rates. While we are unable to directly control for generation, we do include gender, taking into account that public insurance programs may affect women differently from men, and industry.

Immigration is also an important issue with any public policy concerning Hispanics. Camarota and Edwards (2000) attribute much of the growth in the overall uninsured rate to immigrants, due mostly to lower educational attainment and higher poverty rates. Borjas (2003) found no significant change in the uninsured rate among non-citizens following welfare reform in the mid-1990's, as reductions in public coverage eligibility were canceled out by increased labor force participation and a greater rate of employer-sponsored coverage. Kandilov (2007), though, finds no difference in private coverage rates, and higher uninsured rates overall, among immigrants in states with more restrictive Medicaid rules post-reform. Our study takes into account not only how citizenship impacts the likelihood of insurance coverage, but how citizenship can differentially affect Hispanics and non-Hispanics.

Our paper analyzes not only the point-in-time uninsured rate difference between Hispanics and non-Hispanic Whites, the most common approach in the existing literature, but also how differences in personal characteristics between these groups have resulted in diverging uninsured rates. No other paper uses a dataset as rich, or with as long a sample window. In addition to the sequential regression method used previously (Buchmueller, et al., 2007), we use a more formal and heretofore underutilized method, a Blinder-Oaxaca decomposition of the growth in the uninsured rate gap between two different points in time.[3] We also decompose trends separately by insurance source and Hispanic subgroup.

3.3 Data

The Survey of Income and Program Participation (SIPP), is a panel survey of U.S. individuals about income sources, welfare, household and family structure, jobs and work history, and health insurance. Each individual in a household is interviewed every four months (regarding each intervening month) for two to four years. New panels began in each year from 1984 through 1993, and then again in 1996, 2001, and 2004; questions are comparable across panels, though over time some answer options changed (such as educational attainment). The samples sizes and length of the panels vary (Table 3.1). Most of the panels track respondents for about three years, or 8 to 9 interview waves, though the 1996 and 2004 panels lasted for four years.

We pool each completed panel[4] into one dataset of 9.3 person-months, resulting in a continuous cross-section of respondents for all but seven months (March through September 2000, between the end of the 1996 panel and the beginning of the 2001 panel) of the interval between June 1983 and December 2007.

[3]Three exceptions are Van Hook, Brown and Kwanda (2004) for trends in poverty among the children of immigrants, Fairlie and Sundstrom (1999) for the racial unemployment gap, and Ashraf (1996) for the gender wage gap.

[4]The 1989 panel was discontinued after three waves, and lacks information on citizenship, so it was excluded from our sample.

We focus on individuals age 18 to 55, since the difference in the uninsured rate between Hispanics and white non-Hispanics is most pronounced for this age group, and public insurance programs are less likely to cover working age adults. The uninsured rate among children declined slightly for both Hispanics and non-Hispanic Whites over this time (Figure 3.1), and stayed relatively constant for those over age 55 and over (Figure 3.2).[5]

Our primary dependent variable is an indicator variable for whether the individual reported health insurance coverage, irrespective of source, in the surveyed month.[6] We consider individuals with missing values for the insurance questions to be out-of-sample; in our time-series figures, we correct the SIPP-provided person-month weights to account for potential non-random attrition (see Appendix), though all regressions are unweighted.[7]

We represent the individual's insurance coverage choice as a linear model with only this one outcome, insured or uninsured. While we could run multinomial regressions where the possible coverage outcomes are private, public, or none, public coverage is an appropriate part of the choice set of only a select few, since Medicaid is typically available only to expectant mothers, the elderly, and the disabled. Unfortunately, we do not know with certainty who is eligible for public coverage.[8] Our proxy for public

[5]Because of pervasive Medicare coverage, the uninsured rate for non-Hispanic Whites over the age of 65 is negligible throughout. Like their younger counterparts, elderly Hispanics are more likely to be without health insurance (3 to 4 percent each year), but the uninsured rate for this group has remained relatively stable throughout the sample. The uninsured rate gap between Hispanics and non-Hispanic Whites between the ages of 55 and 64 has also remained stable since 1983.

[6]SIPP asks about insurance status for each month individually; while there could be seam bias, where the retrospective months are less reliable than the interview month, our results are no different using only the interview months. The Current Population Survey (CPS) uninsured rates are more often quoted in the press, but the CPS asks whether the individual ever had health insurance during the previous year; this is a very different question than the SIPP, as insurance status often changes during the calendar year. Researchers have found that the CPS measure of insurance status actually follows the point-in-time status around the interview month more closely than the ever-in-year status implied by the question. In addition, the CPS did not ask about citizenship status until 1993, so using the CPS would result in a shorter sample window. We repeat our analysis for the CPS for 1993 through 2007 and get similar results.

[7]The coefficients in weighted regressions, using either the uncorrected or corrected weights, are nearly identical to the unweighted regressions.

[8]If we had a clear indicator of public coverage eligibility, we would a multinomial regression for just those who are eligible, while leaving the remainder in a more traditional linear or nonlinear single-choice model. Even if we knew the eligibility rules in the individual's state, however, we would likely have a large number of individuals who appear

coverage eligibility is an indicator variable equal to unity when the individual is below the (time-varying) federal poverty line. We also interact this variable separately with our female and Hispanic indicators; female because poor women are more likely to be eligible than poor men, Hispanic because we hypothesize that Hispanics may have different takeup responses to public coverage eligibility than non-Hispanics.[9]

In addition, SIPP provides data on the source of insurance. Though the exact options on the questionnaire change between panels, we are able to construct a consistent indicator for whether an individual has coverage through a public program (grouping Medicare, Medicaid, SCHIP, and other federal and state programs together) or through a private insurer. We further divide private coverage into coverage obtained through one's own employer, coverage obtained in one's own name from some source other than own's employer (including the individual market), and coverage obtained in someone else's name (irrespective of who that person is, or where that person obtains coverage). We separately analyze trends by each insurance source category, and separately decompose how the same right-hand side variables contribute to the Hispanic-White gap in each insurance source.

3.4 Trends in the Uninsured Rate

The uninsured rate among adults aged 18 to 55 in the U.S. increased slightly between 1983 and 2007 (Figure 3.3). The uninsured rate increased from 18 percent

to be ineligible but have public coverage. The other kind of error, where individuals who appear to be eligible do not have public coverage, are considered "conditionally covered" in the health insurance literature. While hospitals will assist them in acquiring their rightful public assistance when they present for emergency care, they miss out on regular preventive and non-emergency care. Because of this difference in what kind of care is covered, there is a debate over whether these individuals should be considered insured or uninsured. The definition of "covered by health insurance" that is consistent with the SIPP questionnaire would include non-emergency care, so we consider the "conditionally covered" to be uninsured.

[9]While the percent of public coverage recipient above 100 percent of the federal poverty line increased slightly during our sample window, the vast majority of public coverage recipients had family incomes below the poverty line. Throughout our sample period, the proportion of the population under the poverty line was greater than the proportion with public coverage, indicating a large number of "conditionally covered" (see note 8). Also, the percent uninsured closely matches the difference between our eligibility proxy and the percent with public coverage, suggesting that the "conditionally covered" did not have any health insurance.

in the mid 1980's to almost 20 percent in 1993 and 1994, before falling during the economic boom of the late 1990's. After 1998, the proportion uninsured increased again, but fell slightly after 2003, resulting in an overall 25-year upward trend of 0.04 percentage points per year, which is statistically significant at the 90 percent confidence level.

The overall uninsured rate increased despite statistically significant downward trends among non-Hispanic Whites and non-Hispanic Blacks, particularly since the early 1990s. Both groups follow the pattern of the overall rate, but the larger decreases in the 1990's led to a decline in the uninsured rate for Whites of 0.1 percentage points per year, and 0.14 percentage points for Blacks. The uninsured rate among Asians and Native Americans (not shown) also had a statistically significant downward trend, of 0.14 and 0.60 percentage points per year, respectively.

Alone among the racial and ethnic groups investigated, the uninsured rate for Hispanics trended upward, and by amounts that dwarf the improvement in the other groups. The Hispanic uninsured rate was 27.7 percent in 1984, more than twelve percentage points greater than the uninsured rate for non-Hispanic Whites. Over the next 25 years, the proportion of Hispanics without health insurance increased by 0.4 percentage points per year, to more than 40 percent in the early 2000's. The gap with non-Hispanic Whites more than doubled, increasing to 22 percentage points in 2007. At the beginning of our sample, a nonelderly Hispanic adult was nearly twice as likely to be uninsured as her non-Hispanic White peer; by the end of our sample, she is almost three times as likely.

Most, but not all, of the growth in the Hispanic uninsured rate was among non-citizens (Figure 3.4). In 1984, the uninsured rate among Hispanic non-citizens was 37.4 percent. By 2004, their uninsured rate was more than 60 percent, before falling slightly to 53.7 percent in 2007. For the full 25-year period, the uninsured rate grew by 0.7 percentage points each year, statistically significant at the 99 percent level. Since 1996, non-citizens of Hispanic origin are more likely to be uninsured than insured.

While the level of insurance coverage is higher and the rate of growth smaller than for those born elsewhere, native-born Hispanics also saw an increase in the likelihood of being uninsured. The uninsured rate for native-born Hispanics increased from 24.5 percent in 1983 to 30.5 percent in 2002. Though the uninsured rate for this group fell sharply thereafter, to 20.3 percent in 2007, the proportion of native Hispanics who are uninsured remains far above native Asians (11.0 percent) and Whites (11.5 percent), and comparable to Blacks (21.0 percent). The insurance coverage gap for Hispanics, then, cannot be written off as just an immigration problem.

Still, over this period, the health insurance status of all immigrants got worse. Non-Hispanic non-citizens also saw an increase in their uninsured rate, from 15.6 percent in 1984 to 26.3 percent in 2007, a statistically significant upward trend of 0.32 percentage points per year, and as high as 30 percent in 1996 and 2004. Naturalized Hispanics (not shown) also saw a slight but significant increase, from 26.0 percent in 1983 to 32.9 percent in 2007, a slower rate of increase than Hispanic natives, while naturalized non-Hispanics had a slower, insignificant increase.[10]

Besides citizenship, previous studies have indicated that education, income, and age explain significant portions of both the growth rate of the uninsured population and the Hispanic-non-Hispanic coverage gap.

Figures 3.4(a) (Hispanics) and 3.4(b) (non-Hispanics) display the sharp increase in the uninsured rate among those with less than a high school diploma. This increase

[10]The uninsured rate among Hispanics without immigration information increased steadily, from around 30 percent in the mid 1980's to more than 45 percent in the 2000's.

is especially large for Hispanics, nearly 0.8 percentage points per year, but is also significant for non-Hispanics. It is among college graduates, however, where the Hispanic gap is most stark. The uninsured rate for non-Hispanic college graduates declined from 10.4 percent to 8.2 percent from 1983 to 2007. For Hispanic college graduates, though, the uninsured rate actually increased, from 14.2 percent in 1984 to more than 19.8 percent in 2007. Also, while the uninsured rate for non-Hispanics with just a high school diploma increased slightly, and for those with some college experience was nearly exactly flat over this time, the Hispanic uninsured rate for these groups significantly increased.

A similar story is told in Figures 3.5(a) and 3.5(b). Even the lowest income non-Hispanics are less likely to be uninsured in 2007 than they were in the early 1980's, but the uninsured rate for even the highest income Hispanics has grown significantly. For each category of family income, normalized to the percent of the federal poverty line, the uninsured rate among Hispanics was already higher than non-Hispanics, and that gap has steadily grown over the last 25 years.

Within age groups, the gap between Hispanics and non-Hispanics has also increased (Figures 3.6(a) and 3.6(b)). The non-Hispanic uninsured rate has changed only within a tight band, with the youngest group at the highest level. The Hispanic age groups, though, have each increased, though generally at a slower rate than we see in the education and income categories.

These graphs seem to indicate that differences between Hispanics and non-Hispanics in citizenship, education, and income, and to a lesser extent age,[11] have a great deal of influence on the divergence of the uninsured rate between the two groups. We will focus on these variables, along with our measure of public coverage eligibility (which

[11] When the age categorical variables are added separately to Method 1, the Hispanic-year coefficients do not change much. Also, in the trend decomposition in Method 2, the effect of age is much smaller than education, citizenship, income, and public coverage. As a result, we include age only in the X variables.

is represented by the top line in Figures 3.5(a) and 3.5(b)), in our analysis.

But these variables are likely to be closely related, and their influence is hard to differentiate; greater educational attainment will generally lead to higher income, and the age-income profile is well documented. Though identifying the trends in these variables and how they relate to the uninsured rate is interesting and important, we will not be able to perceive of the relative contribution of each variable to insurance status separate from the others by merely graphing the means. Our two decomposition methods allow us to better separate the effect of each variable on the trends in the uninsured rate.

Previous studies have indicated that the Hispanic population is far from homogeneous, and that we must also consider country of origin within Hispanic ethnicity, as the uninsured rate is higher for individuals of Mexican heritage than for those from other areas in Latin America. Indeed, in each year in our sample, Mexicans had a higher uninsured rate than Puerto Ricans, Cubans, and those from other Latin countries, with a higher growth rate over the two-decade period through 2003.[12] The uninsured rate is highest for Mexican non-citizens, though Hispanic non-citizens from countries other than Mexico narrowed that gap slightly (Figure 3.8), particularly among Cubans. The greater growth rate for Hispanics of Mexican heritage is due to the faster increase among Mexican-Americans, while there was no significant trend among non-Mexican Hispanics born in the U.S. Because of the very different rates of growth seen in Figure 3.8, we decompose the contribution of each variable to the growth in the uninsured rate gap for Mexicans separately from other Hispanics, and also the difference between the two groups.

[12]SIPP has consistent Hispanic heritage indicators only for Mexicans, Puerto Ricans, and Cubans, with all other Hispanics grouped together. Only in the 1996 and 2001 panels did SIPP further differentiate Hispanics. Also, the question that differentiates Hispanics by country of origin was eliminated from the 2004 panel, so the time series ends in 2003.

We may also be concerned with the source of insurance and how the coverage type has changed over time for Hispanics relative to non-Hispanics, particularly if we feel that employer-sponsored coverage is a “better” (or at least more stable) source than public coverage or non-group private plans. From the mid-1980’s through the mid-1990’s, the public coverage rate increased nearly across the board, with the largest growth rate among Blacks (Figure 3.9). Since welfare reform in 1996, however, only Whites have maintained their public coverage level, while a smaller percentage of Hispanics are covered by Medicare, Medicaid, or a state program. The increase in the public coverage gap was especially pronounced for Hispanic non-citizens; the proportion of Hispanic non-citizens’ with public coverage rate fell by three percentage points with passage of the Personal Responsibility and Work Opportunity Reconciliation Act in 1996, which limited Medicaid coverage for new immigrants (Ku and Bruen, 1999), but White and Black non-citizens actually saw a small increase in public coverage.

Private coverage rates were relatively constant for most groups, but fell from more than 60 percent to less than 50 percent among Hispanics. There was a slight increase in the proportion with employer-sponsored insurance among Whites and Blacks, but among Hispanics, there was a substantial decrease from an already much lower percentage (Figure 3.10). All groups saw small, (but statistically insignificant, declines in private insurance from a source other than one’s employer. Finally, while the proportion of Whites receiving insurance through someone else’s policy declined slightly (and this percentage among Blacks actually increased, albeit insignificantly), the proportion among Hispanics fell significantly, from more than 20 percent to less than 15 percent.

3.5 Methods

At its most basic, we run the following linear regression model, for individual i in month t:

$$Unins_{it} = \beta_0 + \beta_1 Hisp_i + \gamma' X_{it} + \sum_{y=1983}^{2002} [\zeta_y I(year_{it} = y) + \delta_y Hisp_i I(year_{it} = y)] \quad (3.1)$$

where $I(year_{it} = y)$ is an indicator function that equals unity when the person-month of observation occurs in year y. Thus, our "adjusted" measure of the Hispanic-White is the uninsured rate in year y, accounting for observable variables, is $\beta_1 + \delta_y$.

Our X_{it} vector in each regression includes indicator variables for whether the individual is black, Asian, or Native American,[13] female, married, has children of one's own, and a categorical variable for one's age in 3-to-5-year increments. We also control for the seasonally-adjusted unemployment rate in her state during that month, as a control for local economic conditions.[14] Finally, we include indicator variables for the interview wave, to control for the increased probability of non-random attrition in later waves. Sample means are reported in Appendix Table 3.2.

We choose a linear model for simplicity, but it is important to note that our results hold for a probit model. Using a linear model allows for easier interpretation of coefficients and marginal effects, and with our multiple methods of decomposition, we should attempt to keep the interpretation as simple, and as comparable between methods, as possible. While a Blinder-Oaxaca decomposition from a nonlinear re-

[13]In our regressions, we exclude non-Hispanic Blacks, Asians, and Native Americans from the sample. Our goal is to explain the difference in the trend in the uninsured rate between Hispanics and the majority group, white non-Hispanics, where the gap is most stark; we thus obtain no additional explanatory power by the inclusion of individuals who are in these other groups. The coefficients and standard errors in regressions including all Blacks, Asians, and Native Americans are nearly identical. We do include race indicators in the regressions to account for potential differences between white and black Hispanics (or Asian Hispanics, or Native American Hispanics, though these interactions with Hispanic are much less common), though the results excluding these variables are not substantially different.

[14]Cawley and Simon (2005) find that state unemployment rate is positively correlated with insurance coverage, and that this correlation survives even after controlling for employment transitions.

gression is possible (Fairlie 2005), the calculation of the trend decomposition becomes unwieldy.[15]

We use two different methods to determine how differences in particular variables (or categories of variables) between Hispanics and non-Hispanic Whites may contribute to the growing gap in the probability of being uninsured between the two groups: (1) the statistical significance of the trend line in the uninsured rate gap, regression-adjusted for each variable sequentially, and (2) pairwise Blinder-Oaxaca-style decomposition of the annual gap, and the change in the gap between starting (e.g., 1984) and ending (e.g., 2007) years.[16]

3.5.1 Method 1 - Regression-Adjusted Trend Line

In the first method, we see how $\beta_1 + \delta_y$, the adjusted gap, changes when we add variables sequentially to the regression. After adding each group of related variables, we then determine whether the slope of the trend in the annual uninsured rate gap is significantly different from zero.

We begin with the unadjusted trend line, the difference in the (unweighted)[17] uninsured rate between Hispanics and non-Hispanic Whites, averaged over each month in the specific calendar year.

We then add categories of variables one-by-one. If the variables were independent, then the order they are added to our regression would be immaterial, but this is likely not the case with education, income, and citizenship, among other variables. We must then determine an order for their addition. We choose to add variables

[15] Of the 9.3 million person-month observations in our OLS regression, the predicted value of the dependent variable is below zero for 985,000 of them (10.6 percent), but 75 percent of those are between zero and -0.04. Only 86 observations have a predicted value of greater than one.

[16] The results of a decomposition of each variable's contribution to the change in the gap's trend in the style of Bound and Freeman (1992) are similar to both of the included methods.

[17] As the uninsured are more likely to leave the sample, and perhaps more so if they are also Hispanic, our unweighted estimates may actually underestimate the gap. Our results should thus be considered a lower bound on Hispanic-White coverage differences.

in rough order of when they were endowed to the individuals in our sample. We start with the variables that are largely out of the individual's control: gender, race, age, and family structure. Next, individuals are either born or choose to live in the United States and, if the latter, whether to become citizens. Then, they decide what level of education to acquire. Next, they (and their families) settle on a certain living standard, which either meets or fails to meet the poverty threshold calculated by the U.S. government. Finally, the individual and his or her spouse choose how much to work, in what industry, and for what pay.

Our first adjusted trend line, therefore, includes controls for Hispanic, the variables in X_{it}, and the year and Hispanic-year dummy variables.

Second, we add citizenship variables. Each SIPP respondent is asked their citizenship status, country of birth, and year entering the U.S. (if non-native) once during their time in the panel, in a special "topical module" separate from the usual ("core") questions.[18] The topical module including the immigration and citizenship questions was during the second interview wave for all panels but the 1984 (wave 8) and 1985 (wave 4) panels. Because these questions are asked just once, and early in the life of the panel, there are many individuals who are never asked about their citizenship status. Also, not surprisingly, there are also many individuals who have no response for these questions. We may worry that the late-arriving and/or non-responding individuals may not be randomly distributed with respect to the other variables (especially if those with missing information are more likely to be undocumented immigrants), so we include "no answer" as one possible citizenship status, along with naturalized citizen, noncitizen, and native-born (the omitted condition). We also interact citizenship status with Hispanic origin. In addition, we include

[18] The one exception is the 2004 panel, where respondents were asked citizenship and immigration questions both in the core and the second topical module. Between the two sources, no individual in this panel has missing citizenship information.

a categorical variable for how long the individual has been in the U.S.: less than ten years, ten to twenty years, or twenty-plus years, with native-born (or missing information) as the omitted condition.[19]

Next, we add a categorical variable for one's education level: less than high school, high school degree only, or some college, with bachelors degree or higher as the omitted condition.[20]

We next add the proxy variable for public coverage eligibility, an indicator variable for whether family income is below the federal poverty line for that year. We include interactions between this proxy variable and Hispanic ethnicity, and between the proxy and the female indicator. In addition, we include interactions with each of the three citizenship indicators, and with both citizenship and Hispanic ethnicity.

Finally, we add variables that account for labor supply and income. We add a categorical variable for the respondent's family income as a percent of the federal poverty level. We include controls for whether the individual is working full time, part time, self employed, or unemployed (the omitted condition), and a broad categorical variable for her industry, based on the job for which she earns the highest hourly wage (imputed or directly reported).[21] In addition, we include indicator variables for spouse's work status, on the theory that a spouse working, especially full time, should increase one's access to health insurance.

As we add more variables, our "adjusted gap" shrinks toward zero in each year. We measure the amount of the gap that remains unexplained in each year, and

[19]We have experimented with adding each of these variables and interactions separately. After adding citizenship status, the interactions and time in U.S. variables have little effect on the Hispanic coefficients; most are significant, though, so we include them.

[20]Interactions between educational attainment and citizenship status are mostly insignificant and changed the other coefficients negligibly.

[21]We can also include an indicator variable for union membership. If we do, though, we lose all observations in 1983 and most of 1984, because union status was not part of the SIPP in the 1984 panel. As the results including a union dummy are not substantially different, we have opted to restore the full sample and drop the union dummy from our regressions.

its significance, as the difference between the sum $\beta_1 + \delta_y$ (for year y) and zero. More importantly, as a measure of how much the upward trend in the gap remains unexplained, we run a linear trend through the unexplained portion of the gap in each year and determine its significance. There are also some obvious trend breaks in the data, so we also include results for piecemeal linear trends.

3.5.2 Method 2 - Pairwise Blinder-Oaxaca Change Decompositions

As mentioned above, conclusions made about the analysis of Method 1 are somewhat limited, because the order that the variables are added to the regression controls the amount of influence we detect for each set of variables. For instance, if income has a large effect on the Hispanic-year coefficients, but much of that effect is sopped up by controlling for variables that are added earlier, such as education or public coverage eligibility, we may incorrectly conclude that income is not an important explanatory variable.

A more thorough method, and one independent of the order of addition for variables, is the Blinder-Oaxaca decomposition. In this paper, we perform two different types of Blinder-Oaxaca decomposition. First, we decompose the contribution of particular variables to the annual uninsured rate gap between Hispanics and non-Hispanic Whites. Then, we decompose the contribution of each variable to the change in the gap between the base year and each given year.

Adapting the setups from Blinder (1973) and Oaxaca (1973), we start with two mutually exclusive and exhaustive groups, non-Hispanic Whites (denoted W) and Hispanics (H). For a particular outcome y (in our primary estimates, the indicator variable for uninsured), we have two parallel models of the effect of variables X on that outcome y:

$$y_H = \beta_H X_H + \epsilon_H,$$
$$y_W = \beta_W X_W + \epsilon_W, \tag{3.2}$$

where the error terms ϵ_H and ϵ_W are mean zero. We expect $\hat{y}_H > \hat{y}_W$, where $\hat{y}_J = \hat{\beta}_J \overline{X}_J$, the fitted value for y for group $J \in \{H, W\}$. The difference between the two fitted outcomes (suppressing the hat notation for simplicity) is:

$$y_H - y_W = \beta_H \overline{X}_H - \beta_W \overline{X}_W. \tag{3.3}$$

Traditional Blinder-Oaxaca decomposition then adds and subtracts either $\beta_W \overline{X}_H$ or $\beta_H \overline{X}_W$. In the former case, after grouping terms, the following equation results:

$$y_H - y_W = \beta_W(\overline{X}_H - \overline{X}_W) + (\beta_H - \beta_W)\overline{X}_H. \tag{3.4}$$

which decomposes the difference between y in populations H and W into the portion that can be explained by differences in the mean of the variables X in the two groups (the first part, commonly called the "explained" portion) and the portion owing to differences in the coefficients between the two groups for the same values of X (the latter part, or the "unexplained" portion).[22]

A similar process can be used to decompose the change in the uninsured rate gap between Hispanics and white non-Hispanics between any two points in time (Altonji

[22] Alternatively, if we added and subtracted $\beta_W \overline{X}_H$ from equation (5), our decomposition equation would be

$$y_H - y_W = \beta_H(\overline{X}_H - \overline{X}_W) + (\beta_H - \beta_W)\overline{X}_W. \tag{3.5}$$

Oaxaca and Ransom (1994) demonstrate that these two methods bound the actual Hispanic effect on the probability of being uninsured. Jann (2005) puts the Blinder-Oaxaca decomposition in general form:

$$y_H - y_W = \beta^\star(\overline{X}_H - \overline{X}_W) + [(\beta_H - \beta^\star)\overline{X}_H + (\beta^\star - \beta_W)\overline{X}_W], \tag{3.6}$$

where $\beta^\star$ is the "benchmark coefficient." In the case outlined in the text, $\beta^\star = \beta_W$; in the footnote case, $\beta^\star = \beta_H$. We present results from only the decomposition that uses the white non-Hispanic coefficient as the benchmark ($\beta^\star = \beta_W$), since most of our sample is white non-Hispanic (between 84 and 94 percent) and therefore the population-weighted average coefficients (Reimers, 1983, and Cotton, 1988) and the pooled coefficients (Neumark, 1988) are very similar to β_W.

and Blank, 1999). For members of group $J \in \{H, W\}$, the change in y between periods t and t' can be written as

$$\begin{aligned}
\Delta y_J \equiv y_{Jt'} - y_{Jt} &= \beta_{Jt'}\overline{X}_{Jt'} - \beta_{Jt}\overline{X}_{Jt} \\
&= \beta_{Jt'}\overline{X}_{Jt'} - \beta_{Jt}\overline{X}_{Jt} + \beta_{Jt'}\overline{X}_{Jt} - \beta_{Jt'}\overline{X}_{Jt} \\
&= \beta_{Jt'}(\overline{X}_{Jt'} - \overline{X}_{Jt}) + (\beta_{Jt'} - \beta_{Jt})\overline{X}_{Jt},
\end{aligned} \tag{3.7}$$

where the first line uses the definition $\Delta z_J = z_{Jt'} - z_{Jt}$, for some vector z.

Then, we can decompose the change in the gap between H and W for outcome y as

$$\begin{aligned}
\Delta y_H - \Delta y_W = \beta_{Ht'}(\overline{X}_{Ht'} - \overline{X}_{Ht}) + (\beta_{Ht'} - \beta_{Ht})\overline{X}_{Ht} \\
- \beta_{Wt'}(\overline{X}_{Wt'} - \overline{X}_{Wt}) - (\beta_{Wt'} - \beta_{Wt})\overline{X}_{Wt}.
\end{aligned} \tag{3.8}$$

We use the white non-Hispanic coefficients $\beta_{Wt'}$ and β_{Wt} as benchmarks for their respective years, so the decomposition between base year t and ending year t' is

$$\begin{aligned}
\Delta y_H - \Delta y_W = \beta_{Wt}[(\overline{X}_{Ht'} - \overline{X}_{Wt'}) - (\overline{X}_{Ht} - \overline{X}_{Wt})] \\
+ (\beta_{Wt'} - \beta_{Wt})(\overline{X}_{Ht'} - \overline{X}_{Wt'}) \\
+ [(\beta_{Ht'} - \beta_{Wt'}) - (\beta_{Ht} - \beta_{Wt})]\overline{X}_{Ht} \\
+ (\beta_{Ht'} - \beta_{Wt'})(\overline{X}_{Ht'} - \overline{X}_{Ht}).
\end{aligned} \tag{3.9}$$

Altonji and Blank (1999) identify the first line as the effect of relative changes over time in the observed characteristics of the two groups, and the second line as the effect of changes over time in the white non-Hispanic coefficient, holding differences in the observed characteristics fixed; these two parts make up the "explained" portion of the change decomposition. The third line is the effect of changes over time in the relative coefficients between the two groups, and the fourth line "captures the fact that changes over time in the characteristics of [the two groups] alter the consequences of differences in group coefficients" (p. 3226); these two lines make up the "unexplained" portion of the change decomposition.[23] We isplay only the total explained and unexplained portions of equation (9), rather than separately present the results for each of the four portions.

In our main analysis, t is 1984, the first full year of the sample and the year where the uninsured rate gap between Hispanics and non-Hispanic Whites is smallest, while t' is the last year of the sample, 2007. It is clear from the graphs that there is a break in the trend between the 1993 and 1996 SIPP panels, probably due to subtle differences in surveying or weighting techniques or variable definitions. Also, the gap has a clear peak in 2001, decreasing slightly in the last years of the panel. Therefore, we will also decompose the change in the gap between 1984 and 1995, between 1996 and 2001, and between 2002 and 2007.

One disadvantage of this method is that we are unable to ascertain the amount of the annual gap, or the change in the annual gap, that can be explained by variables that are perfectly collinear with Hispanic. In particular, we can no longer include the interaction between Hispanic and the public coverage eligibility proxy, or Hispanic and citizenship status. While the coefficients on these variables (in the last regression of Method 1) are significantly different from zero, they are of relatively small magnitude, and their exclusion leaves the key coefficients largely unchanged.

3.6 Results

3.6.1 Regression Adjusted Trend Line

The first column of Table 3.3 is the actual (unweighted, because the regressions are all unweighted) gap in uninsured rates between Hispanics and non-Hispanic Whites,

[23] Fairlie and Sundstrom (1999) have a mathematically equivalent change decomposition, but have a broader definition of the "explained" portion that would include parts of the third line of equation (3.9).

averaged over the months in each year. The actual gap grew from a low of 12.4 in 1984 to a high of 28.7 in 2001, decreasing in the remaining years of the sample. Fitting a linear trend to the numbers in this column (bottom panel of Table 3.3), we find a slope of 0.505, and we can reject the null of a zero slope at 99 percent confidence level.

The remaining columns of Table 3.3 present the adjusted Hispanic-White gap, the sum of the coefficient on the Hispanic intercept term plus each Hispanic-year coefficient ($\beta_1 + \delta_y$). In the second column, we add the X variables: gender, race and ethnicity, family structure, age categories, and the state unemployment rate. On average, the annual gaps shrink by six percent, or about 0.3 percentage points. Most annual gaps are at least a little smaller, with the 2004 gap shrinking by the greatest percent (4.1 percent, or 1.1 percentage points). The slope of the fitted linear trend line is flatter at 0.476 percentage points per year, a decrease of 5.7 percent in the slope from the unadjusted gap, but is still significantly different from zero.

Next, we add citizenship variables, including indicators for naturalized citizen and non-citizen, an indicator for missing citizenship information, interactions of each of those three statuses with the Hispanic dummy, and a categorical variable for how long the individual has spent in the United States. On average, the annual gaps fall by an additional ten percentage points, or about half of the actual gap. The reduction is largest for the early years; the gap shrunk by at least 50 percent in nine out of the first thirteen years, and on average by about 40 percent in the 2000's. The trend line is 26.2 percent flatter than the unadjusted trend, but is still significantly different from zero.

In the fourth column, we add categorical variables for educational attainment. The annual gaps are an additional 3.5 percentage points smaller on average, and the average adjusted gap is now a third of the unadjusted gap for that year. Still, this reduction is substantially smaller for the more recent years in the sample, and the linear trend has a statistically significant slope of 0.341 percentage points per year.

We next add our proxy variable for public coverage eligibility, plus interactions between the proxy and Hispanic, female, each citizenship status, and both Hispanic and citizenship status. Accounting for public coverage eligibility actually increases the gap slightly in most years, particularly in the years after welfare reform, suggesting that Hispanics would be even less likely to have insurance coverage than non-Hispanic Whites were it not for the increased likelihood of being eligible for Medicaid. This leads to a steeper slope in the trend line, though still a 29 percent slower average annual increase in the gap than the unadjusted figures.

Finally, we add income and labor supply variables, including indicators for whether one is full time, part time, or self employed, the same employment status categories for one's spouse, and five industry categories. Controlling for these variables reduces each of the gaps, undoing the reversal caused by adding public coverage eligibility. Again, though, the reduction in the gap is largest in the early years; in fact, the 1984 gap disappears. Still, 30 to 40 percent of the gap in each year of the 2000's remains unexplained. As a result, the trend line is slightly steeper than in column 4, with the gap growing at a statistically significant 0.351 percentage points per year, only 30.4 percent less steep than the unadjusted trend. Despite significant contributions by citizenship and education, about 70 percent of the divergence in the uninsured rate between Hispanics and non-Hispanic Whites remains unexplained.

Figure 3.11 graphs these adjusted gaps. It is clear from the graph that citizenship and education have the largest effects, while income also reduces the annual gaps, but most of its effect worked through variables added previously. Drawing one trend line through all of the points in each series, we see that we can explain some of the growth of the gap, but that a significant upward trend remains even in the all-inclusive series.

In Figure 3.11, there appears to be a break at 1996, possibly due to subtle differences between earlier panels and the 1996 panel (though the definitions for most of the variables we use appear to be largely unchanged), or because of the growth of the economy during this period. It may be more appropriate to draw separate trend lines, one from 1983 to 1995, another from 1996 to the peak of the uninsured rate gap in 2001, and another from 2002 to the end of the sample. Table 3.4 gives linear trend coefficients through Figure 3.11 for these time periods. Between 1983 and 1995 (second panel), the unadjusted gap has a positive slope of 0.37 percentage points per year, significantly different from zero, but becomes essentially flat once we control for citizenship. The gap grew the most between 1996 and 2001 (third panel), and adding personal characteristics actually makes the positive trend line steeper. In the later period, between 2002 and 2007 (bottom panel), the uninsured rates for the two groups started to converge; the adjusted trend lines suggest that this convergence should have been even faster, based on relative improvements in Hispanics' citizenship status and education. Except for the last few years, decomposing the 25-year trend into smaller portions strengthens the finding that the growing (and then ameliorating, somewhat) gap in uninsured rates between Hispanics and non-Hispanic Whites cannot be explained by observable differences between the groups.

The other coefficients are largely in the expected direction, and because of our large sample size (9.3 million person-months, with standard errors clustered by nearly 360,000 individuals), almost all statistically significant. Females and married individuals are less likely to be uninsured, while those with children are more likely to be uninsured. The uninsured rate is decreasing in age (between ages 21 and 55), education, and family income, and increasing in the state unemployment rate. Non-citizens are most likely to be uninsured, more so if they are also Hispanic, and while naturalized citizens are more likely to be uninsured than native-born citizens, there is no difference between Hispanic and non-Hispanic naturalized citizens. Foreign-born individuals that have been in the U.S. for more than ten years are less likely to be uninsured than even the native-born, all else equal. Those with family income less than the poverty level (our proxy for public coverage eligibility) are far more likely to be uninsured, though being female, Hispanic, or a non-Hispanic immigrant reduces this probability. Hispanic immigrants (both non-citizens and naturalized citizens) are particularly vulnerable to being uninsured. Finally, full time workers, or those with a working spouse, are least likely to be uninsured, while the self-employed are actually worse off than the unemployed, though a self-employed spouse is better than an unemployed spouse. Coefficients are unchanged, for the most part, when adding each new set of controls, except for those variables that are directly correlated, such as citizenship variables interacted with the public coverage eligibility proxy, or the married dummy with spouse's work status.

3.6.2 Blinder-Oaxaca Decompositions

The Blinder-Oaxaca decomposition for the Hispanic-White uninsured rate gap in each year yields similar results to the adjusted trend line (Appendix Table 3.6). Except for citizenship in 1984, each variable contributes some positive amount to "explaining" the actual gap in all years. Differences in educational attainment consistently explain between ten and 20 percent of the actual gap, growing slightly more important in later years. Income and labor market differences seem to be important from 1990 through 1995 and in the last years of the sample, but otherwise do not contribute much. Citizenship becomes much more important starting in 1990, but remains less influential than differences between the two groups in education level. Differential public coverage eligibility is crucial from 1983 to 1989, explaining up to half of the actual gap, but gradually fades in importance. Still, in most years, the largest portion of the actual gap is “unexplained,” growing to more than half of the actual gap in the 2000’s.

The results in Table 3.5, which formalize the linear trend analysis in the previous section, reach a similar conclusion: much of the divergence in uninsured rates between Hispanics and non-Hispanic Whites over 25 years remains unexplained by differences in personal characteristics between the groups. Of the 12.3 percentage point increase from 1984 to 2007 (top line), differences between Hispanics and non-Hispanic Whites in citizenship account for 6.4 percentage points (50.7 percent). Citizenship accounts for almost half of this amount, with income somewhat less important. Much as we saw (though to a lesser extent) in Method 1, Hispanic-White differences in public coverage eligibility actually work in the wrong direction, suggesting that Hispanics’ greater likelihood of living below the poverty line should actually make the increase in the gap greater than it is. Again, a large portion of the decomposition in Table 3.5, 5.2 percentage points (42.4 percent of the actual gap), is “unexplained.”

The other lines in Table 3.5 decompose the change in the uninsured rate gap between varying starting and ending years. As in Table 3.4, the decomposition results for the change between 1996 and 2001 indicate that differences in personal characteristics between Hispanics and non-Hispanic Whites should result in a larger increase in the uninsured rate gap, and the gap’s decrease between 2002 and 2007 should have been larger considering changes in the citizenship, education, and age compositions of the Hispanic population. Also, the largest portion of the increase in the gap between 1984 and 1995, and between the gap’s lowest (1984) and highest

(2001) points, is unexplained. Other than the last six years of the sample, then, observable differences between Hispanics and non-Hispanic Whites do not explain the largest portion of the divergence in these groups' uninsured rates.

3.6.3 Decompositions by Insurance Source

Figure 3.12 repeats the analysis from Method 1, with an indicator for public coverage, rather than uninsured, as the dependent variable. Adding our X variables reduces the average annual public coverage gap by 16 percent. Accounting for citizenship, the adjusted gap is actually larger than the unadjusted gap in all years. The education variables induce about a 30 percent reduction in the average gap and, not surprisingly, accounting for public coverage eligibility makes the gap nearly disappear. Adding variables does not change the shape of the public coverage gap's long-term time series; the unadjusted gap and all of the adjusted gaps each initially decline, increase during the early 1990's, fall through 2000, and then rise again sharply at the end of our sample. The public coverage gap between Hispanics and Whites is almost fully explained by Hispanics' greater probability of having family income below the poverty line, and unlike the uninsured rate gap, there does not seem to be a concern for the widening of the public coverage gap.

If the public coverage gap can be almost completely explained away, but the uninsured rate gap persists even after controlling for observable characteristics, then the private coverage adjusted gap must also be growing, as we see in Figure 3.13. The unadjusted gap is negative, meaning the private coverage rate for Hispanics is lower than the rate for non-Hispanic Whites, and getting progressively more negative as our sample advances. Adding citizenship variables reduces the private coverage gap by about a third, education by about 20 percent, and public coverage eligibility (which could have a crowdout effect on private coverage), and income each reduce the gap by about 10 percent. Still, about a quarter of the average annual gap remains unexplained, and the adjusted trend indicates an increasing divergence for the full sample, 1983 through 1995, and 1996 through 2001, though the gap shrinks after 2002.

The persistence of the private coverage gap between Hispanics and non-Hispanic Whites, even after accounting for observable characteristics, does not appear to be because of differences in rates of coverage by employer-sponsored insurance rates. Figure 3.14 shows that citizenship explains about half of the average annual gap in coverage from one's own employer, while the difference in educational attainment covers nearly the rest. While the own employer coverage gap is increasing at a small but significant rate, 0.17 percentage points per year, nearly all of this growth is between 1996 and 2001, while Hispanics' adjusted employer coverage rate is nearly equal to non-Hispanic Whites' in 2007.

The growing private coverage gap also does not seem to be due to a divergence in non-group coverage, as seen in Figure 3.15. The gap in own private coverage through a source other than one's employer is never larger than 3.5 percentage points, and accounting for observable characteristics reduces the gap to approximately one percentage point. There is no significant widening of the gap; in fact, unlike most of the other graphs, the trend after 1996 is toward zero, reflecting a tightening of the gap.

The source of the widening adjusted gap in the uninsured rate seems to be the decreased likelihood of private coverage through someone else, usually a spouse or parent. While this rate has remained constant at around 28 percent for non-Hispanic Whites, the proportion of Hispanics covered by someone else fell from 21 percent to 14 percent (Figure 3.16). Differences between the groups in citizenship and education each explain only about 14 percent of the average annual gap, public coverage eligibility another 11 percent, and income about 24 percent, but more than 40 percent the average annual gap remains unexplained. In each time period, the adjusted trend line, including all variables, is no different than the trend in the unadjusted gap; over the full sample period, there is a significant increase of 0.2 percentage points per year in gap in coverage through a family member, with the largest increase between 1996 and 2001.

3.6.4 Decompositions by Hispanic Subgroup

Figure 3.8, and previous studies, suggest that the uninsured rate has been much different for Mexicans living in the U.S. compared to other Hispanic subgroups. In Appendix Table 3.7, we present the results of separate decompositions for the growth in the uninsured rate gap between each Hispanic subgroup and non-Hispanic Whites.

The largest subgroup, and the one with consistently the highest uninsured rate, is Mexicans, including both Mexican immigrants and U.S.-born Mexican-Americans. Between 1984 and 2003, the uninsured rate among Mexicans increased from 33.7 percent to 43.5 percent, an increase in the gap with non-Hispanic Whites of 14.7 percentage points. As we found in Table 3.5, differences in education and citizenship account for the largest portion of this gap, but 9.4 percentage points, or nearly two-thirds, of the Mexican-White gap remains unexplained. Similarly, the gap between Puerto Ricans and Whites, and other Hispanics (those who are neither Mexican, Puerto Rican, nor Cuban) and Whites, are largely left unexplained even after controlling for observables. The Cuban-White gap has actually shrunk slightly, despite differences in citizenship that suggest the gap should have widened.

We also repeat the analysis from Fronstin, Goldberg, and Robins (1997), decomposing the differences within the Hispanic group (Appendix Table 3.8). The uninsured rate among Mexicans increased much faster than the uninsured rate for Puerto Ricans or Cubans, but the widening of these intra-Hispanic gaps remains largely unexplained after controlling for observables. Again, education and citizenship differences between Mexicans and these two groups seem to be most important, especially considering nearly all individuals of Puerto Rican heritage are citizens. The gap between Mexicans and the remaining Hispanics did not widen nearly as much, and observables account for more than the actual gap.

3.7 Discussion

Irrespective of the decomposition methodology, a large portion of the growth in the uninsured rate for Hispanics relative to non-Hispanic Whites remains unexplained, even after accounting for differences in income, education, citizenship, and other factors available in our data. To get an idea for the magnitude of this unexplained portion, we can do some simple accounting.

The uninsured rate for white non-Hispanic nonelderly adults fell from 19.6 percent in 1984 to 14.0 percent in 2007. The uninsured rate for Hispanics increased from 32.0 percent in 1984 to as much as 42.2 percent in 2001, before falling to 38.7 percent in 2007; the gap, therefore, increased from 12.4 to 28.7 percentage points, and then back down to 24.7 percentage points, during that time. Of that 12.3 percentage point increase in the gap over the full sample, we can explain 3.9 percentage points under the first decomposition method,[24] and 7.1 percentage points under the more statistically rigorous method.[25]

If the gap had remained constant at 1984 proportions, the Hispanic uninsured rate in 2007 would have been 26.4 percent, or 6.6 million uninsured out of 25.1 million

[24]For Method 1, we multiply the trend coefficient (0.351) times 24 years, so the regression-predicted change in the gap for 2003 is 8.4 percentage points.

[25]For Method 2, a change of 5.2 percentage points remains unexplained between 1984 and 2007, as can be seen in the last row of Table 3.5.

Hispanic nonelderly adults. Instead, there were 9.0 million nonelderly Hispanic adults without health insurance in the average month in 2007, an increase of almost 2.4 million over the counterfactual. Accounting for income, education, citizenship, and other variables, we can explain 749,000 (Method 1) to 1.37 million (Method 2) of the difference between the actual and the counterfactual. In other words, between 1.0 and 1.6 million of the increase in the number of uninsured Hispanic nonelderly adults remains unexplained by observable differences between Hispanics and non-Hispanic Whites.

Another counterfactual can help us estimate how the growth of the Hispanic population in the U.S. has led to an increase in the number of uninsured. The non-Hispanic White population grew by 4.4 percent between 1984 and 2007; in contrast, the Hispanic population more than tripled. If the Hispanic population had grown by only 4.4 percent, but experienced the same increase in the uninsured rate as they actually did over that time (to 38.7 percent uninsured), there would have been 3.0 million uninsured Hispanics in 2003, 6.0 million fewer than the actual figure. Based on our regression estimates, we can only explain between 1.9 million (Method 1) and 3.5 million (Method 2), leaving an extra 2.5 to 4.1 million uninsured Hispanics for whom we cannot account from interethnic differences in observable characteristics.

Though our data cannot account for the extra several million uninsured Hispanics, there are other potential explanations. We suggest a number of possible factors that could have increased the price of insurance, changed preferences for health insurance, or decreased the (relative) income available to purchase insurance for Hispanics during this time.

Price of Insurance

Our estimation by source of insurance indicate that most of the unexplained increase in the Hispanic uninsured rate is from fewer Hispanics acquiring coverage through another person. There is also some evidence that employer coverage is harder to come by, though we can explain much of the Hispanic-White difference in employer coverage by controlling for observable characteristics. Perhaps employers are still offering coverage to employees at roughly the same rate, but are attempting to control costs by being more restrictive about coverage for family members. Unfortunately, SIPP does not consistently ask whether an individual receives an offer of coverage from her employer, and whether that offer will include not just her but also her family members; we are able to determine whether she accepted the offer, but SIPP does not ask whether an offer was rejected.

If jobs that are considered "high-quality" are also the ones who are more likely to offer health insurance for the whole family as part of the compensation package, then perhaps Hispanics have grown more likely to take lower-quality jobs. Whether because of labor market tightness or changing returns to relative skill level, Hispanics could be increasingly likely to be stuck in jobs that offer lower total compensation, less certainty over hours or eventual job tenure, or less workforce bargaining power, all of which would reduce the probability of comprehensive insurance offers. While we find little change in the industrial composition of the Hispanic workforce in our data, it could be that Hispanics more often find themselves in positions that are less generous with insurance offers within these broad industrial categories.

Also worth considering is the percentage of Hispanics who are self-employed or work in small businesses owned by others, since both groups are likely to face higher and more volatile health insurance plan prices. While SIPP does not have a consistent measure of the size of an employee's firm throughout the years, we do know from our data that the proportion of Hispanics for whom self-employment is their primary job has decreased slightly, from 4.8 percent in 1983 to 3.5 in 2007. The self-employment rate for white non-Hispanics has fallen by a higher percentage, from 9.7 percent in 1983 to 6.3 in 2007. Still, Hispanics are, and have been throughout our sample, much less likely to be self-employed, and this explanation is probably not overly relevant.

It's also possible that Hispanics have increased job volatility. Since many jobs, even "high quality" positions, do not offer health insurance until after a probationary period (often six months or one year), people who change jobs frequently will often experience long spells without health insurance while they wait for eligibility. If, over the time covered in our sample, job tenure declined and job volatility increased for Hispanics, then this could explain part of the growing disparity; we will explore this possibility in future research on insurance transitions.

As with most large survey datasets, we should note that there is no information in the SIPP about an immigrant's legal status. An immigrant lacking proper documentation is likely to find jobs that are of lower quality, as she has little leverage in the job market. Very few of these jobs will offer health insurance for her, and even fewer will extend coverage to her spouse or children. If the proportion of the immigrant population without documentation grew larger from the mid-1980's to the 2000's, then the immigrant population would be less likely to acquire affordable private coverage, in ways that are not captured by our citizenship variables.[26]

Another problem with the SIPP data is the lack of a primary language variable (other than in the 2001 and 2004 panels). If language or thickness of one's accent is

[26]This possible difference in the immigrant population relies on the assumption that SIPP is able to interview the same proportion of undocumented immigrants that it currently does, and that a larger undocumented population will result in more interviews with undocumented immigrants. On the contrary, if currently no additional undocumented immigrants respond to SIPP interview requests, then there would be no measured difference in the immigrant population.

a significant barrier to acquiring health insurance, whether through finding a quality job or discovering eligibility for a public program, many foreign-born (and even some U.S.-born) Hispanics would be more likely to go without coverage. We probably control for most of the language impact by including citizenship variables and their interaction with Hispanic ethnicity and poverty, but within these groups we could still have some with language difficulty having trouble getting insurance, with others with English proficiency being more successful.

For non-English speakers, social networks could prove to be very important when attempting to access health care and health insurance. Gresenz, Rogowski, and Escarce (2007) find that Mexican immigrants are more likely to be insured when they live in a neighborhood with a high concentration of Spanish-speakers and other Hispanics, though this effect disappears for U.S.-born Mexican-Americans, suggesting that those who are more acculturated are less reliant on networks to find access to affordable health care. Also, Borjas and Hilton (1996) find that immigrants are more likely to be covered by Medicaid if earlier immigrants from that same country also had high levels of Medicaid coverage. Our finding of a widening gap in coverage through another person fits into the social network model of these two papers. Still, since network effects are less important for native-born Hispanics, these spillovers are unlikely to explain much of the growing gap with other groups.

Preferences for Risk, Health Care, and Insurance

The lack of insurance among Hispanics could be a matter of personal preference as it relates to risk, in ways that we are not capturing in the variables we have. Barsky et al. (1997) find that in hypothetical questions on a survey, Hispanics are the second-most (after Asians) risk tolerant group, and that those with higher risk tolerance are less likely to have health insurance coverage. The survey responses are consistent with increased prevalence of risky behaviors found by other researchers, including lower saving rates (Engen et al. 1999), more credit card usage (Medina and Chau 1998), lower seatbelt usage rates (Campos-Outcalt et al. 2003), and lower vaccination rates (Herrera, Zhao and Klevens 2001).

If native-born Hispanics are healthier, it may make sense for more of them to choose to go without health insurance. But Hendee (1991) finds that Hispanics have a higher risk of diabetes, hypertension, tuberculosis, HIV, alcoholism, cirrhosis, some cancers, and violent deaths. On the other hand, Hispanics may be less likely to seek treatment in an emergency department (Taylor, Larsen, and Correa-de-Araujo 2006), use prescription drugs (Weinick et al. 2004), or have a usual source of care (Weinick, Zurekas, and Cohen 2000), so even if they are less healthy, access to and financial coverage of medical care may be less important. Future research should consider whether Hispanic health care utilization has changed over this time; the decreased likelihood of insurance relative to other groups makes more sense if Hispanics are using less care than they were in the early 1980's, not just that they're less likely to use care at any particular point in time.

Also, Hispanics may value insurance less as a part of their total compensation package, whether it be from an employer or part of a welfare plan. Non-citizens who are remitting income to their families in other countries may prefer to receive compensation in the form of wages rather than fringe benefits, particularly if their time in the U.S. labor market is short or seasonal. This is less likely for citizens, though we should also consider whether Hispanics are more likely to be a secondary wage earner, or have the possibility of being covered by another family member's plan.

Income Available for Insurance

Another possibility is that many Hispanics' lower income makes them unable to afford insurance, even if the price is reasonable and their preference is for coverage. If Hispanics have fallen further behind non-Hispanic Whites in income over the sample period, then affordability could explain some of the unexplained health insurance disparity. In our data, though, Hispanics have actually gained on white non-Hispanics in family income relative to the federal poverty level. Hispanics' increased likelihood of being in the poorest part of the distribution has shrunk somewhat, and Hispanics are also more likely to be in the 300-400 percent "middle class" range, relative to how the distribution has changed for non-Hispanic Whites. On the other hand, Hispanics are even less likely than Whites to be in the highest income group relative to where they were in the 1980's. Still, the most constrained group seems to have shrunk somewhat, so growing income disparity is likely not the culprit.

Income could be insufficient to purchase insurance if long term or frequent illness cuts into earning power. A person with tuberculosis or cirrhosis would probably be unable to work, or would have frequent interruptions in their work schedule which would make them less attractive to potential employers (not to mention more expensive to insure under a small business employer sponsored plan). The Hendee paper cited above makes the case that Hispanics are more likely to have the type of long term health conditions that could prevent a stable income source.

3.8 Conclusions

Hispanics were already more likely to be uninsured in the early 1980's. Over the course of the next 25 years, they fell even further behind, not only compared to non-Hispanic Whites, but also relative to non-Hispanic Blacks, Asians, and Native

Americans. Our research suggests that, though differences in education and citizenship between Hispanics and non-Hispanic Whites can account for some of the divergence in health insurance coverage, a substantial portion of the growth in the uninsured rate gap remains unexplained, particularly among individuals of Mexican heritage. In fact, we estimate this portion to represent more than one million extra uninsured Hispanics in 2007. Looking further into how health insurance is delivered, we find that most of the unexplained increase was due to lower rates of coverage through a family member, while gaps in public programs and employer-sponsored coverage largely disappear when controlling for observable characteristics.

While we are confident in our findings, we should stress that our paper has some limitations. The biggest downside to using the SIPP is that English language proficiency is not included, except in the 2001 and 2004 panels. While we expect that we have controlled for most of the effect that language would have on the likelihood of being uninsured with our citizenship status variables, it is likely that English proficiency varies within the Hispanic non-citizen group, and immigrants with better understanding of English will be more likely to find a “quality” job that offers insurance, or better understand their public coverage eligibility.

We are also reliant upon the data being accurate and complete, with only random attrition, though our results are no different when re-weighting observations to control for non-random attrition (see Appendix). Of particular concern is citizenship status; we control for the possibility that those who do not answer the citizenship question could be different than those who do, but it is possible that many non-citizens, especially those without proper documentation, may claim to be citizens, even native-born. We are also unable to tell which immigrants are legal and which are not, which probably has a substantial impact on the likelihood of having health insurance and access to affordable health care.

Our preliminary research indicates that, in addition to the increased likelihood of being uninsured at any particular time relative to other groups, insured Hispanics have also grown more likely to lose coverage, while uninsured Hispanics are less likely to gain coverage, though the divergence is not nearly as monotonic as the point-in-time coverage analyzed in this paper. While some medical care can wait until the individual has insurance, frequent transitions may subject policyholders to restrictions on pre-existing conditions, and leave those who have dropped coverage vulnerable to medical and financial risk from poorly timed illnesses or injuries. In the near future, we would like to apply the analysis from this paper to insurance transitions, to determine if the variables we found to be important sources of coverage disparities, such as citizenship and education, also account for the disparity in coverage lapses.

3.9 Appendix: Inverse Probability Weighting for Time Series Graphs

As with most panel surveys, the SIPP experiences nonrandom attrition from the sample. The nonrandom nature of this attrition is a concern for us because evidence suggests that those who are most likely to be uninsured are also the most likely to attrit from the sample. To account for nonrandom attrition, we reweight each observation by the inverse probability of its being present in the sample. Following Robins, Rotnitzky, and Zhao (1995) and Wooldridge (2000), we first determine the probability of individual i being present in the sample at time t using a probit regression, where the covariates are the time $t-1$ (or time-invariant) characteristics and the SIPP interview wave:

$$Prob(present)_{it} = \Phi(Z'_{i,t-1}\gamma),$$

where Φ is the standard normal cumulative distribution function and Z is a vector of personal characteristics. Note that because, by definition, we do not know the value for any time-varying characteristic when the individual is not present, Z includes the value of each characteristic at time $t-1$ (unless, of course, it is time-invariant), with the exception of the SIPP interview wave. In addition to the wave variable, we include dummy variables for being uninsured, unemployed, Hispanic, black, Asian, Native American, female, and married in Z, and use the SIPP weight (lagged as necessary) when running the probit regression.

After fitting a selection probability, π, to each observation, we then find the probability of being present in time t, conditional on being present in each previous period from entry time ($t = 1$) to time $t-1$:

$$p_{it} = \pi_{i1} * \pi_{i2} * \pi_{i3} * \ldots * \pi_{i,t-1} * \pi_{it}$$

The new bias-corrected weight is then the SIPP-created weight times $1/p_{it}$.

Finally, we rescale all weighted totals so that they match the total number of nonelderly adults in the United States for that month, according to the Census figures.

We use the inverse probability weights in our graphs of the time series of insurance status for different groups, but not in our regressions. Regression results using the weights are not substantially different than the unweighted results, despite SIPP's oversampling of lower income communities. These results are available from the authors upon request. We should also note that we control for the interview wave of the person-month observation in our regressions, in order to pick up some of the potential effect of nonrandom attrition, which is increasingly likely in later waves.

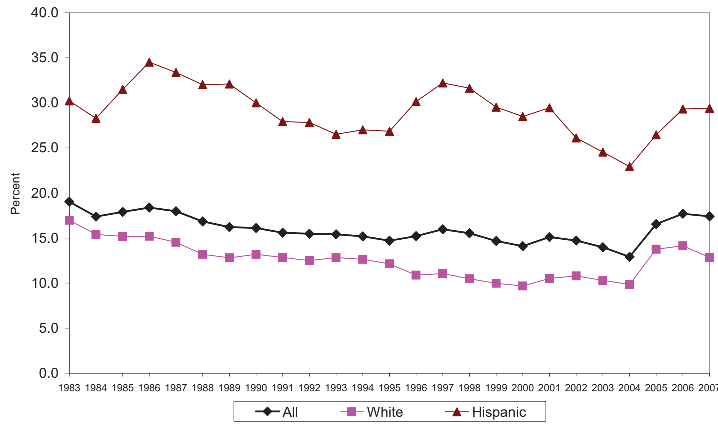


Figure 3.1: Uninsured Rate By Ethnicity, Age 17 and Under

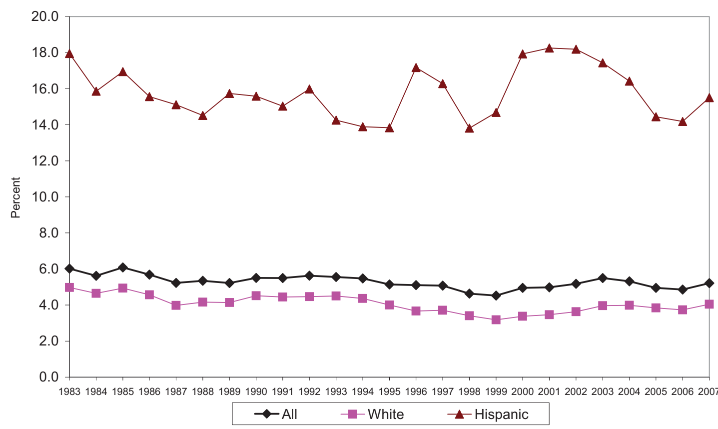


Figure 3.2: Uninsured Rate By Ethnicity, Age 55 and Over

3.10 Figures and Tables

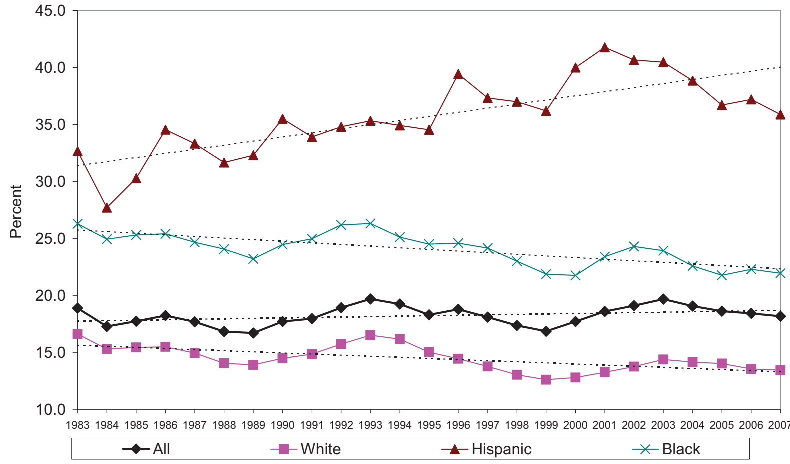


Figure 3.3: Uninsured Rate by Racial and Ethnic Group, Age 18 to 55

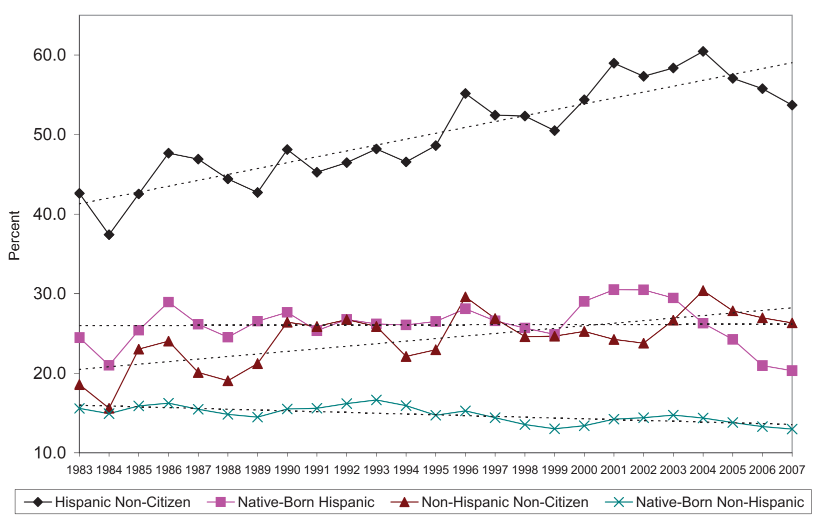
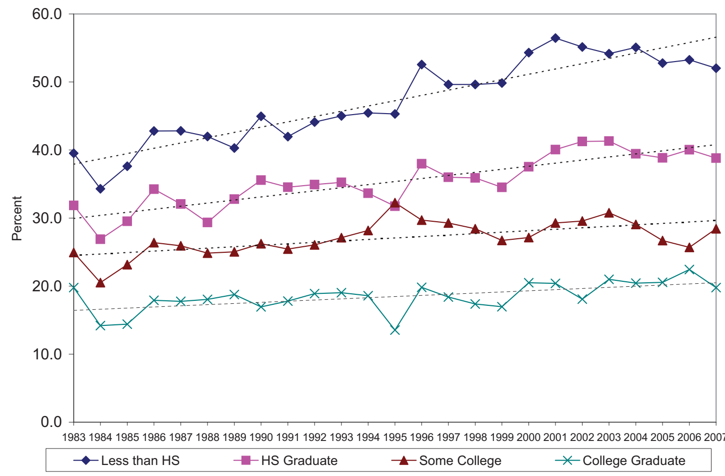


Figure 3.4: Uninsured Rate by Hispanic Origin and Citizenship, Age 18 to 55

(a) Hispanics



(b) Non-Hispanics

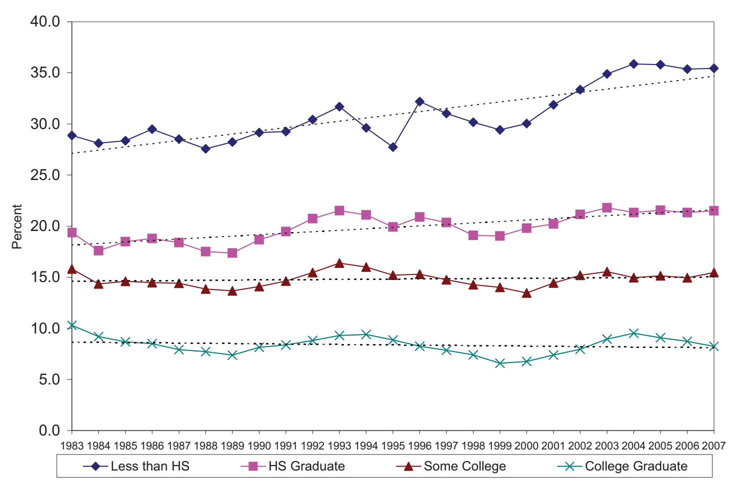
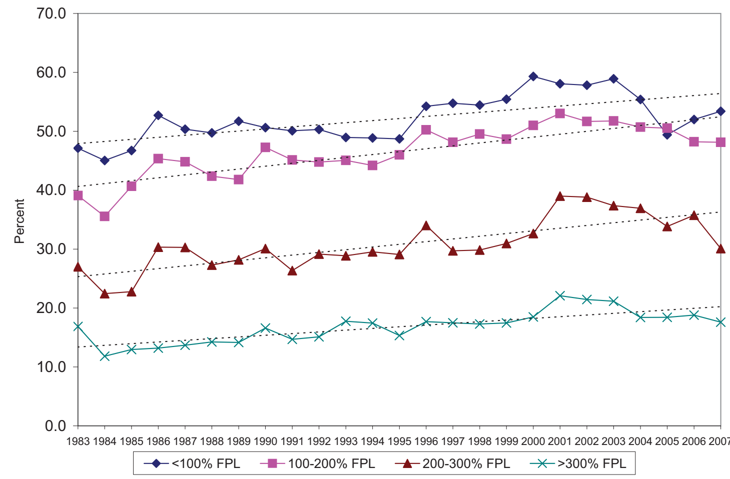


Figure 3.5: Uninsured Rate by Educational Attainment, Age 18 to 55

(a) Hispanics



(b) Non-Hispanics

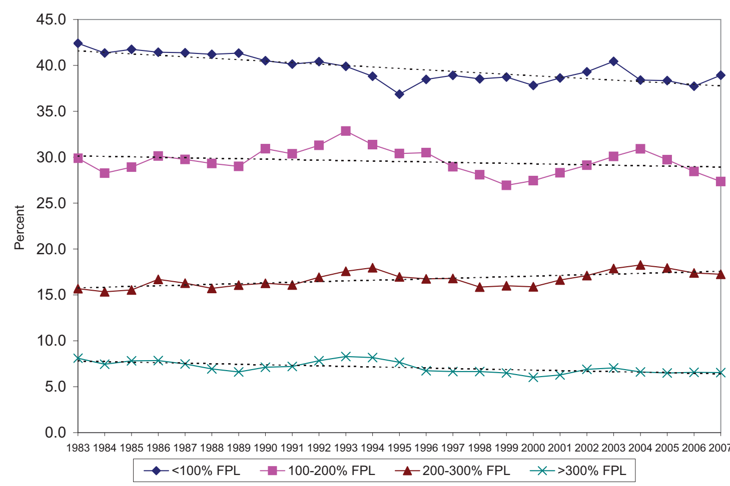
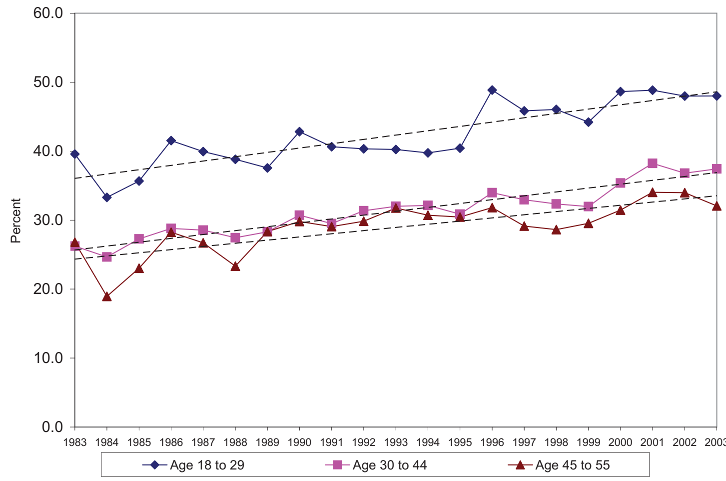


Figure 3.6: Uninsured Rate by Family Income as a Percent of the Federal Poverty Level, Age 18 to 55

(a) Hispanics



(b) Non-Hispanics

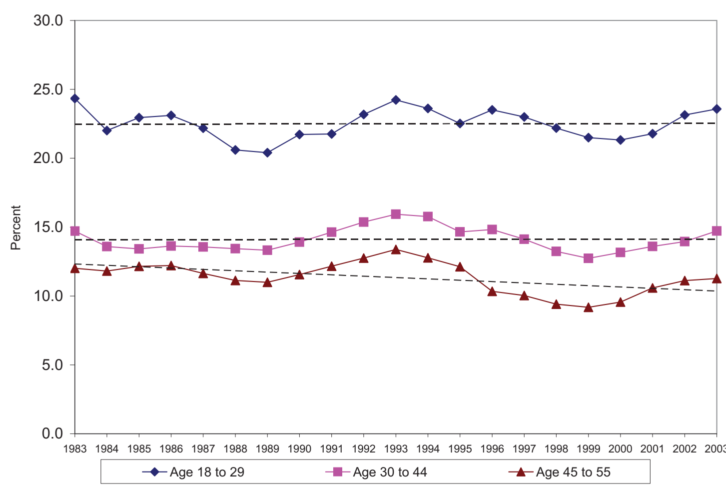


Figure 3.7: Uninsured Rate by Age, Age 18 to 55

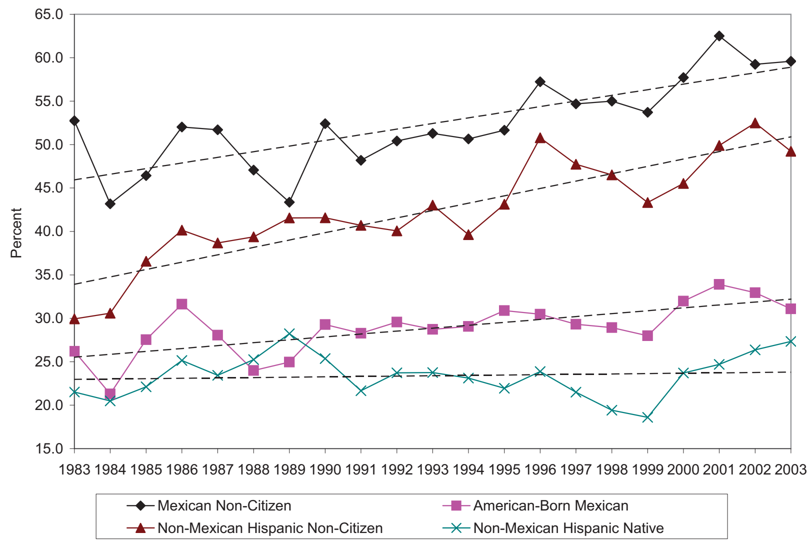


Figure 3.8: Uninsured Rate by Hispanic Subgroup and Citizenship, Age 18 to 55

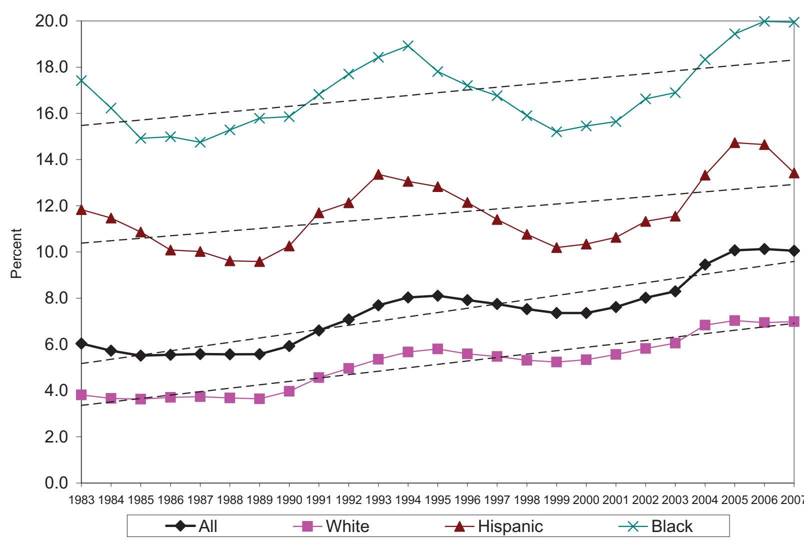


Figure 3.9: Public Coverage Rate By Racial and Ethnic Group, Age 18 to 55

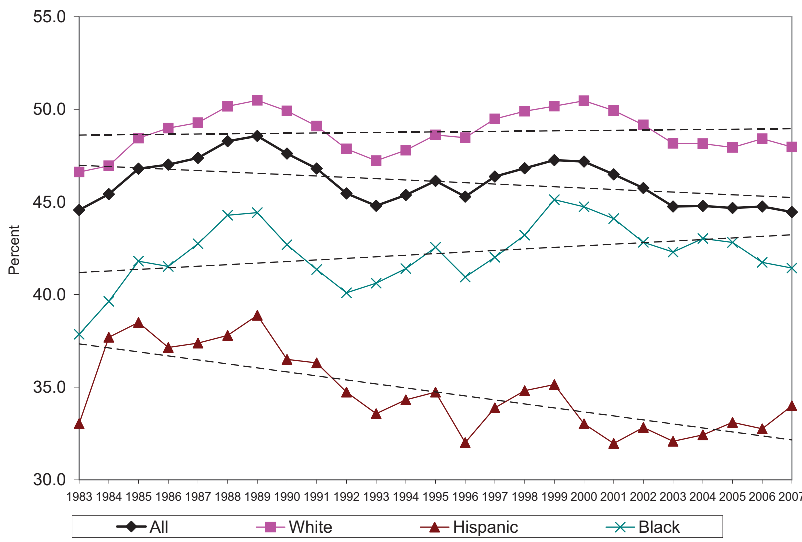


Figure 3.10: Employer-Sponsored Coverage in Own's Name Rate By Racial and Ethnic Group, Age 18 to 55

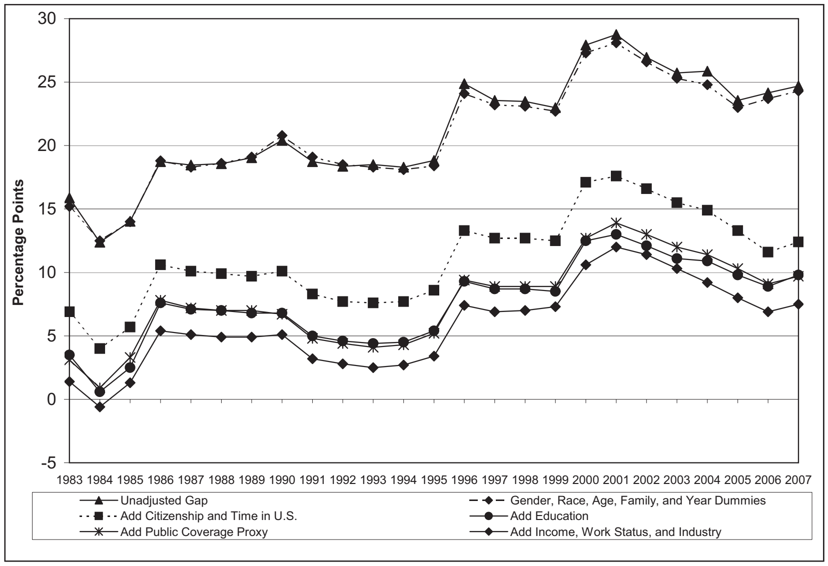


Figure 3.11: Hispanic - White Non-Hispanic Gap in Uninsured Rate

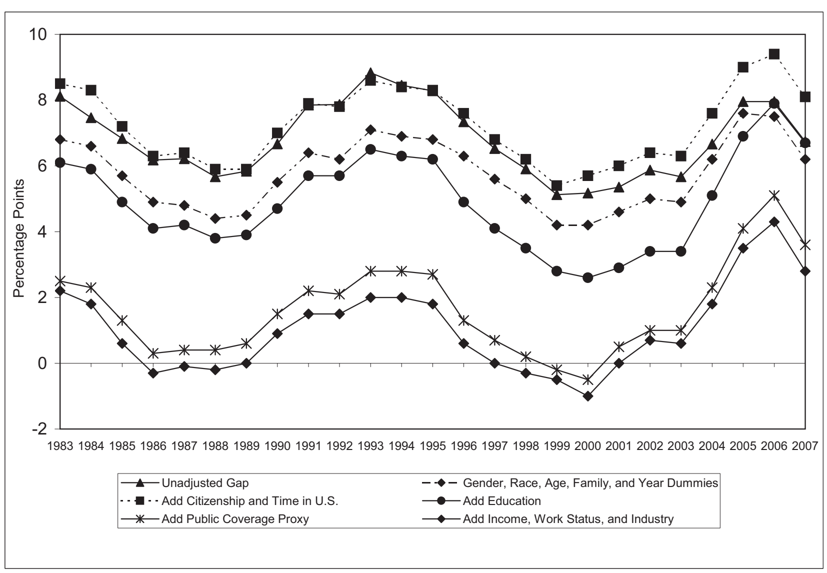


Figure 3.12: Hispanic - White Non-Hispanic Gap in Public Coverage Rate

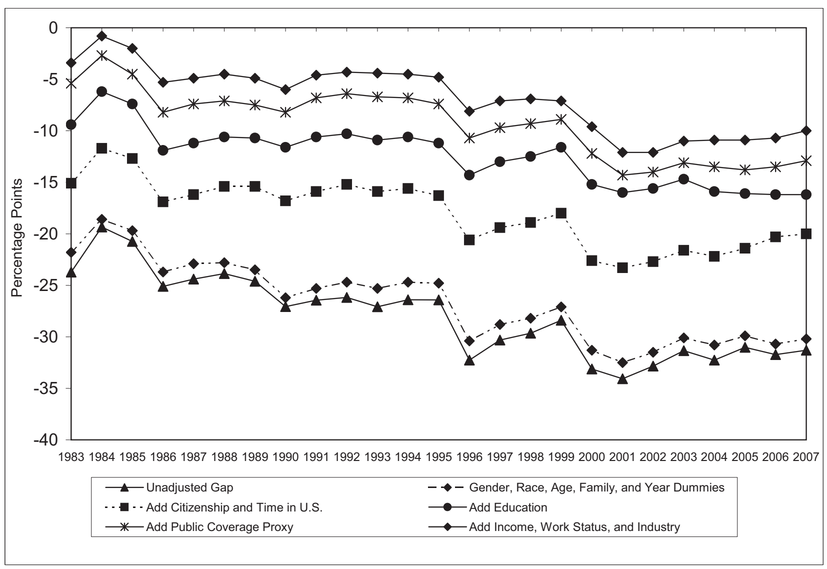


Figure 3.13: Hispanic - White Non-Hispanic Gap in Private Coverage Rate

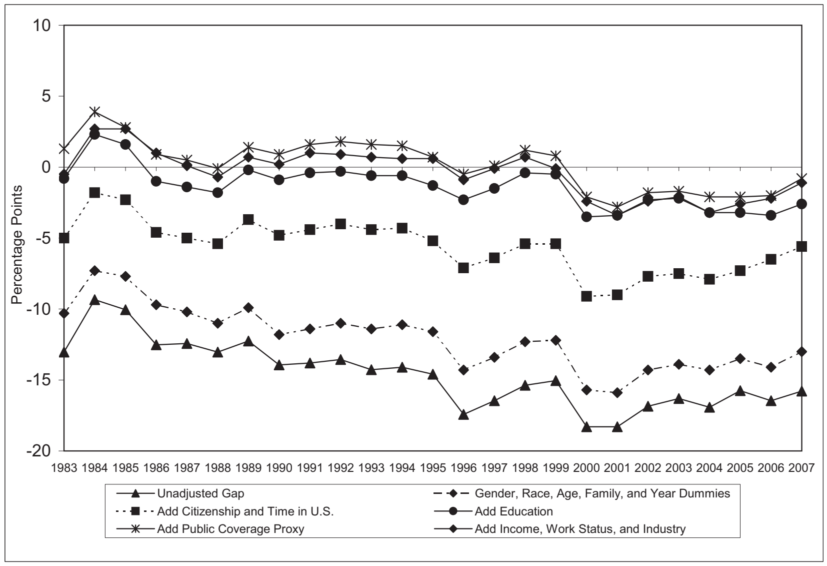


Figure 3.14: Hispanic - White Non-Hispanic Gap in Rate of Own Employer-Sponsored Coverage

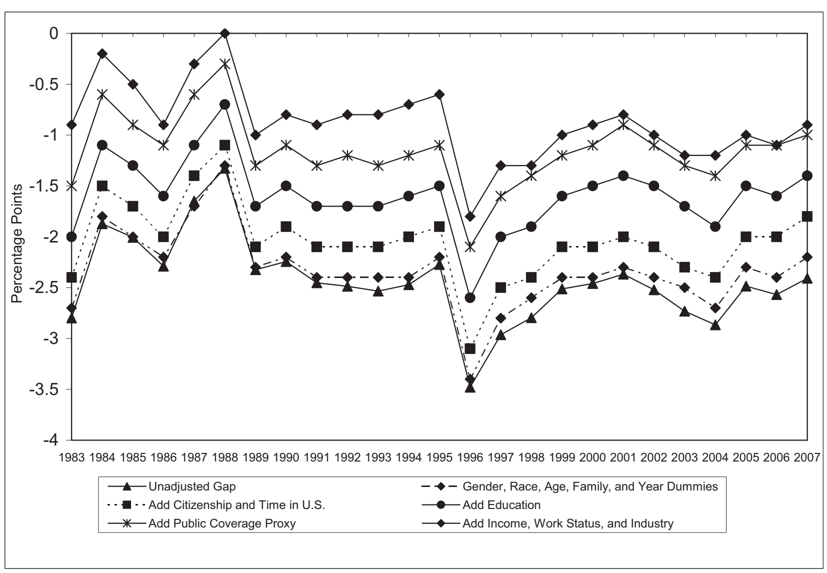


Figure 3.15: Hispanic - White Non-Hispanic Gap in Rate of Own Private non-ESI Coverage

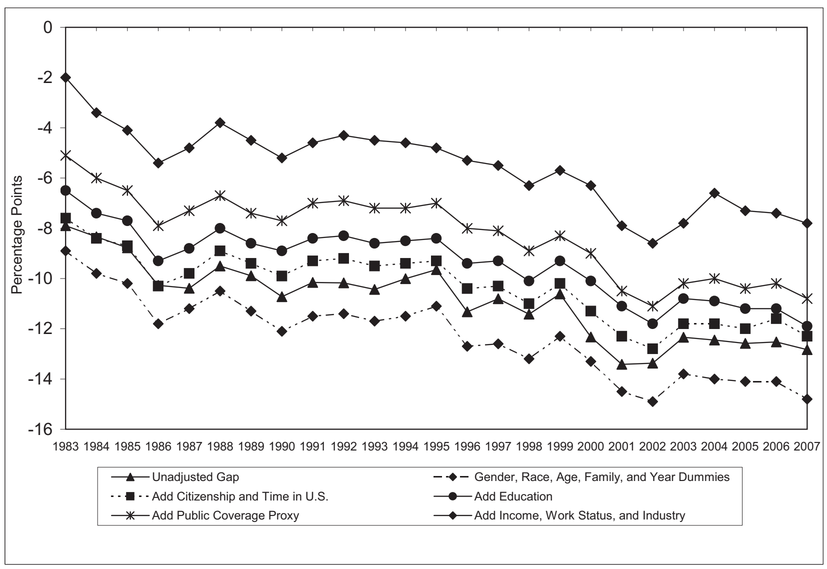


Figure 3.16: Hispanic - White Non-Hispanic Gap in Rate of Private Coverage From Someone Else

Table 3.1: Survey of Income and Program Participation (SIPP) Panels

Panel	Waves	Sample Size	Beginning Month	Ending Month
1984	9	23,283	June 1983	July 1986
1985	8	16,011	October 1984	July 1987
1986	7	13,292	October 1985	March 1988
1987	7	13,681	October 1986	April 1989
1988	6	13,793	October 1987	December 1989
1990	8	25,595	October 1989	August 1992
1991	8	16,358	October 1990	August 1993
1992	9	22,438	October 1991	December 1994
1993	9	22,694	October 1992	December 1995
1996	12	41,378	December 1995	February 2000
2001	9	39,551	October 2000	December 2003
2004	12	150,282	October 2003	December 2007

Note: Sample size reflects number of nonelderly adults in the first month of Wave 1 of that panel.

Table 3.2: Sample Means

	All		Hispanic		White Non-Hispanic	
	Mean	Std Dev	Mean	Std Dev	Mean	Std Dev
Uninsured	0.180	(0.384)	0.369	(0.483)	0.154	(0.361)
Hispanic	0.119	(0.324)	1.000	(0.000)	0.000	(0.000)
Female	0.512	(0.500)	0.522	(0.500)	0.511	(0.500)
Black	0.008	(0.090)	0.069	(0.254)	0.000	(0.000)
Asian	0.001	(0.035)	0.010	(0.100)	0.000	(0.000)
Native American	0.002	(0.047)	0.019	(0.136)	0.000	(0.000)
Married	0.211	(0.408)	0.302	(0.459)	0.198	(0.399)
Own Children Present (0/1)	0.662	(0.473)	0.757	(0.429)	0.648	(0.477)
Age						
18-20	0.078	(0.268)	0.099	(0.299)	0.075	(0.263)
21-23	0.076	(0.265)	0.094	(0.292)	0.074	(0.262)
24-26	0.080	(0.271)	0.097	(0.296)	0.078	(0.267)
27-30	0.113	(0.316)	0.133	(0.340)	0.110	(0.313)
31-35	0.145	(0.353)	0.159	(0.365)	0.144	(0.351)
36-40	0.146	(0.353)	0.140	(0.346)	0.147	(0.354)
41-45	0.138	(0.345)	0.117	(0.321)	0.141	(0.348)
46-50	0.121	(0.326)	0.091	(0.288)	0.125	(0.330)
51-55	0.103	(0.304)	0.070	(0.256)	0.108	(0.310)
State Unemployment Rate	5.782	(1.640)	6.152	(1.481)	5.732	(1.654)
Citizenship						
N/A	0.153	(0.360)	0.140	(0.347)	0.154	(0.361)
Non-Citizen	0.054	(0.227)	0.323	(0.468)	0.018	(0.132)
Naturalized citizen	0.029	(0.168)	0.108	(0.311)	0.018	(0.133)
Citizenship Interacted with Hispanic						
N/A	0.017	(0.128)	0.140	(0.347)	0.000	(0.000)
Hispanic Non-Citizen	0.039	(0.193)	0.323	(0.468)	0.000	(0.000)
Hispanic Naturalized	0.013	(0.113)	0.108	(0.311)	0.000	(0.000)
Time in U.S.						
Less than 10 Years	0.032	(0.177)	0.179	(0.384)	0.012	(0.110)
10 to 20 Years	0.022	(0.147)	0.126	(0.331)	0.008	(0.089)
More than 20 Years	0.019	(0.138)	0.076	(0.265)	0.012	(0.108)
Less than HS	0.128	(0.334)	0.368	(0.482)	0.095	(0.293)
HS Graduate	0.337	(0.473)	0.321	(0.467)	0.339	(0.473)
Some College	0.230	(0.421)	0.174	(0.379)	0.237	(0.425)
College	0.306	(0.461)	0.137	(0.344)	0.329	(0.470)
Public Coverage Eligibility Proxy						
Under 100% FPL	0.116	(0.321)	0.228	(0.420)	0.101	(0.302)
Female Under 100% FPL	0.067	(0.251)	0.138	(0.345)	0.058	(0.233)
Hispanic Under 100% FPL	0.027	(0.163)	0.228	(0.420)	0.000	(0.000)
Citizenship Interacted with Eligibility Proxy						
N/A	0.025	(0.156)	0.037	(0.188)	0.023	(0.151)
Non-Citizen	0.013	(0.112)	0.087	(0.282)	0.002	(0.050)
Naturalized citizen	0.004	(0.061)	0.020	(0.138)	0.002	(0.040)
Citizenship Interacted with Hispanic and Eligibility Proxy						

continued on next page

continued from previous page

	All		Hispanic		White Non-Hispanic	
	Mean	Std Dev	Mean	Std Dev	Mean	Std Dev
N/A	0.004	(0.066)	0.037	(0.188)	0.000	(0.000)
Non-Citizen	0.010	(0.102)	0.087	(0.282)	0.000	(0.000)
Naturalized citizen	0.002	(0.048)	0.020	(0.138)	0.000	(0.000)
Family Income as % of Poverty Line						
Less than 50	0.059	(0.236)	0.098	(0.297)	0.054	(0.226)
50-100	0.057	(0.232)	0.131	(0.337)	0.047	(0.212)
100-150	0.075	(0.263)	0.150	(0.357)	0.065	(0.246)
150-200	0.088	(0.283)	0.139	(0.346)	0.081	(0.273)
200-300	0.185	(0.388)	0.203	(0.403)	0.182	(0.386)
300-400	0.163	(0.369)	0.121	(0.327)	0.168	(0.374)
More than 400	0.374	(0.484)	0.158	(0.365)	0.403	(0.491)
Part Time	0.127	(0.333)	0.115	(0.318)	0.128	(0.334)
Full Time	0.581	(0.493)	0.536	(0.499)	0.587	(0.492)
Self Employed	0.082	(0.274)	0.048	(0.214)	0.086	(0.281)
Unemployed	0.606	(0.489)	0.567	(0.495)	0.611	(0.487)
Industry						
N/A	0.231	(0.422)	0.307	(0.461)	0.221	(0.415)
Manufacturing	0.244	(0.429)	0.240	(0.427)	0.244	(0.430)
Utilities	0.047	(0.212)	0.037	(0.190)	0.048	(0.215)
Sales	0.144	(0.351)	0.136	(0.343)	0.145	(0.352)
Service	0.295	(0.456)	0.250	(0.433)	0.301	(0.459)
Government	0.039	(0.193)	0.030	(0.170)	0.040	(0.196)
Spouse Work Status						
Part Time	0.064	(0.245)	0.049	(0.216)	0.066	(0.249)
Full Time	0.355	(0.479)	0.310	(0.463)	0.361	(0.480)
Self Employed	0.062	(0.241)	0.034	(0.182)	0.066	(0.247)
Unemployed	0.117	(0.321)	0.151	(0.358)	0.112	(0.316)
Mexican Origin	0.071	(0.257)	0.614	(0.487)	0.000	(0.000)
Public Insurance	0.463	(0.499)	0.333	(0.471)	0.481	(0.500)
Private Insurance	0.055	(0.228)	0.033	(0.178)	0.058	(0.234)
Own Employer-Sponsored Insurance	0.260	(0.439)	0.162	(0.369)	0.273	(0.446)
Own Non-Employer Private Insurance	0.768	(0.422)	0.517	(0.500)	0.802	(0.399)
Covered by Family Member	0.062	(0.241)	0.124	(0.330)	0.053	(0.225)
Number of Observations	9,268,530		1,107,374		8,161,156	
Number of Person Clusters	359,033		44,435		314,598	

Table 3.3: Predicted Hispanic - White Non-Hispanic Gaps in Uninsured Rate (Method 1)

	Actual Gap (Percentage Points)	Predicted Gap (Percentage Points)				
		Add Basic Controls	Add Citizenship and Time in U.S.	Add Education	Add Public Coverage Eligibility Proxy and Interactions	Add Income, Work Status, and Industry
1983	15.9	15.2	6.9	3.5	3.1	1.4
1984	12.4	12.5	4.0	0.6	0.9	-0.6
1985	14.0	14.0	5.7	2.5	3.3	1.3
1986	18.7	18.8	10.6	7.6	7.8	5.4
1987	18.5	18.3	10.1	7.1	7.2	5.1
1988	18.6	18.6	9.9	7.0	7.0	4.9
1989	19.0	19.1	9.7	6.8	7.0	4.9
1990	20.4	20.8	10.1	6.8	6.7	5.1
1991	18.7	19.1	8.3	5.0	4.8	3.2
1992	18.4	18.5	7.7	4.6	4.4	2.8
1993	18.5	18.3	7.6	4.4	4.1	2.5
1994	18.3	18.1	7.7	4.5	4.3	2.7
1995	18.8	18.4	8.6	5.4	5.2	3.4
1996	24.9	24.1	13.3	9.3	9.4	7.4
1997	23.6	23.2	12.7	8.7	8.9	6.9
1998	23.5	23.1	12.7	8.7	8.9	7.0
1999	23.0	22.7	12.5	8.5	8.9	7.3
2000	27.9	27.3	17.1	12.5	12.7	10.6
2001	28.7	28.1	17.6	13.0	13.9	12.0
2002	26.9	26.6	16.6	12.1	13.0	11.4
2003	25.7	25.3	15.5	11.1	12.0	10.3
2004	25.9	24.8	14.9	10.9	11.4	9.2
2005	23.6	23.0	13.3	9.8	10.3	8.0
2006	24.2	23.7	11.6	8.9	9.1	6.9
2007	24.7	24.3	12.4	9.8	9.7	7.5
Linear Trend, 1983-2007						
Coefficient	0.505	0.476	0.373	0.341	0.360	0.351
t-statistic	8.0	7.8	5.6	6.0	5.8	5.9
Percent of Trend in Actual Gap Explained		5.7%	26.2%	32.6%	28.7%	30.4%

Table 3.4: Linear Time Trends in the Uninsured Rate Gap between Hispanics and Non-Hispanic Whites

	Actual Gap (Percentage Points)	Predicted Gap (Percentage Points)				
		Add Basic Controls	Add Citizenship and Time in U.S.	Add Education	Add Public Coverage Eligibility Proxy and Interactions	Add Income, Work Status, and Industry
1983-2007						
Coefficient	0.505	0.476	0.373	0.341	0.360	0.351
t-statistic	8.0	7.8	5.6	6.0	5.8	5.9
Percent of Trend in Actual Gap Explained		5.7%	26.2%	32.6%	28.7%	30.4%
1983-1995						
Coefficient	0.365	0.370	0.133	0.138	0.096	0.120
t-statistic	2.7	2.6	0.9	0.9	0.6	0.9
Percent of Trend in Actual Gap Explained		-1.4%	63.5%	62.2%	73.6%	67.0%
1996-2001						
Coefficient	0.914	0.911	0.986	0.849	0.969	0.983
t-statistic	1.9	2.1	2.5	2.4	2.7	3.1
Percent of Trend in Actual Gap Explained		0.2%	-7.9%	7.1%	-6.0%	-7.6%
2002-2007						
Coefficient	-0.522	-0.517	-0.980	-0.549	-0.751	-0.883
t-statistic	-2.5	-2.4	-6.1	-4.0	-6.2	-6.2
Percent of Trend in Actual Gap Explained		0.9%	-87.9%	-5.2%	-44.1%	-69.2%

Table 3.5: Blinder-Oaxaca Decompositions for the Change in the Hispanic-White Non-Hispanic Uninsured Rate Gap

Change between	Actual Change in Gap	Percentage Point Change						Percent of Actual Change in Gap					
		Inc	Educ	Cit	Pub	Oth	Unexp	Inc	Educ	Cit	Pub	Oth	Unexp
1984 and 2007	12.3	1.9	3.4	6.3	-4.7	0.3	5.2	15.0	27.3	50.7	-37.9	2.4	42.4
1984 and 1995	6.4	1.8	1.1	1.9	-3.5	1.1	4.0	28.5	17.6	30.0	-54.9	16.4	62.4
1984 and 2001	16.4	1.9	2.9	3.3	-5.0	0.9	12.4	11.7	17.5	20.1	-30.4	5.5	75.6
1996 and 2001	3.9	0.4	0.5	-0.2	-1.9	-0.1	5.3	10.6	12.6	-6.3	-50.2	-3.2	136.4
2002 and 2007	-2.2	-0.1	-0.1	3.2	0.6	-0.5	-5.3	4.5	4.9	-141.0	-25.1	20.5	236.2

Note: Benchmark coefficient is from the White Non-Hispanic regression.
Key:
Inc includes family income as a percent of poverty, work status (own and spouse), and industry of employment.
Educ includes the three categories of educational attainment.
Cit includes citizenship status and time lived in the U.S.
Pub includes the public coverage eligibility proxy and its interaction with gender and citizenship status.
Oth includes gender, marital status, presence of children, wave of interview, age, and state unemployment rate.
Unexp is the total of the "unexplained" portions of the decomposition, including the constant term.

Table 3.6: Blinder-Oaxaca Decompositions for Annual Hispanic-White Non-Hispanic Uninsured Rate Gap

	Actual	Percentage Points						Percent of Actual Gap					
	Gap	Inc	Educ	Cit	Pub	Oth	Unexp	Inc	Educ	Cit	Pub	Oth	Unexp
1983	15.9	1.3	2.3	0.1	5.6	1.0	5.6	8.5	14.3	0.5	35.4	6.2	35.1
1984	12.4	0.3	1.7	-0.1	7.1	0.1	3.3	2.7	13.7	-1.2	57.3	0.4	27.0
1985	14.0	0.9	1.7	0.5	6.9	0.0	4.0	6.4	11.9	3.3	49.5	0.1	28.8
1986	18.7	0.9	2.0	1.2	7.2	0.2	7.2	5.0	10.6	6.3	38.2	1.3	38.5
1987	18.5	0.6	1.9	0.9	7.6	0.2	7.4	3.3	10.2	4.9	40.9	0.8	39.9
1988	18.6	0.8	1.8	0.2	7.7	0.2	7.8	4.5	9.5	1.2	41.5	1.2	42.0
1989	19.0	1.1	2.0	1.1	7.1	0.3	7.4	5.5	10.4	5.7	37.5	1.8	38.9
1990	20.4	2.9	3.1	2.9	3.5	0.3	7.7	14.3	15.3	14.1	17.2	1.3	37.8
1991	18.7	2.5	2.8	2.0	4.1	0.0	7.3	13.4	14.9	10.6	22.1	0.1	38.9
1992	18.4	2.4	3.0	1.5	3.5	0.1	7.8	12.9	16.3	8.1	19.3	0.7	42.7
1993	18.5	2.5	3.1	1.5	3.8	0.4	7.1	13.8	16.8	8.2	20.6	2.3	38.4
1994	18.3	2.6	2.8	1.2	2.7	1.0	8.1	14.0	15.1	6.3	14.7	5.6	44.3
1995	18.8	2.2	2.8	1.8	3.6	1.1	7.4	11.6	15.1	9.5	18.9	5.9	39.1
1996	24.9	1.9	4.1	3.4	4.0	1.1	10.4	7.4	16.4	13.6	16.3	4.3	42.0
1997	23.6	2.0	3.6	2.3	4.1	1.0	10.6	8.6	15.3	9.7	17.3	4.2	44.9
1998	23.5	1.7	3.8	2.3	3.7	1.4	10.6	7.4	16.0	9.8	15.7	5.8	45.3
1999	23.0	1.5	3.8	2.5	3.2	1.1	10.9	6.6	16.6	10.9	14.0	4.7	47.3
2000	27.9	2.3	4.2	1.5	3.3	1.3	15.3	8.2	15.2	5.4	11.8	4.6	54.9
2001	28.7	2.3	4.6	3.1	2.1	1.0	15.7	7.9	15.9	10.9	7.3	3.3	54.7
2002	26.9	2.3	5.2	2.9	1.9	0.8	13.9	8.5	19.2	10.9	6.9	3.0	51.5
2003	25.7	2.4	4.9	3.5	2.1	0.6	12.2	9.5	18.9	13.8	8.3	2.3	47.2
2004	25.9	2.8	4.6	4.8	3.3	0.8	9.6	10.8	17.8	18.5	12.6	3.0	37.2
2005	23.6	2.9	4.2	4.1	2.8	0.6	8.9	12.2	18.0	17.4	12.0	2.6	37.9
2006	24.2	2.7	4.1	5.6	2.5	0.5	8.7	11.3	16.8	23.4	10.2	2.3	36.1
2007	24.7	2.2	5.1	6.1	2.4	0.3	8.6	8.9	20.5	24.7	9.8	1.4	34.7
Average	20.7	1.9	3.3	2.3	4.2	0.6	8.9	9.3	16.0	11.0	20.5	3.0	43.2

Note: Benchmark coefficient is from the White Non-Hispanic regression.
Key:
Inc includes family income as a percent of poverty, work status (own and spouse), and industry of employment.
Educ includes the three categories of educational attainment.
Cit includes citizenship status and time lived in the U.S.
Pub includes the public coverage eligibility proxy and its interaction with gender and citizenship status.
Oth includes gender, marital status, presence of children, wave of interview, age, and state unemployment rate.
Unexp is the total of the "unexplained" portions of the decomposition, including the constant term.

Table 3.7: Blinder-Oaxaca Decompositions for the Change in Gap with Non-Hispanic Whites

	Actual Change in the Gap, 1984-2003	Percentage Point Change						Percent of Actual Change in Gap					
		Inc	Educ	Cit	Pub	Oth	Unexp	Inc	Educ	Cit	Pub	Oth	Unexp
Mexican	14.7	2.4	3.7	3.2	-4.5	0.4	9.4	16.5	25.4	21.7	-30.3	3.0	63.6
Puerto Rican	3.9	2.0	1.4	1.1	-11.4	0.6	10.2	51.8	36.8	27.1	-294.5	15.2	263.6
Cuban	-0.9	0.4	0.6	4.3	-4.7	0.0	-1.5	-42.4	-70.9	-476.1	522.0	2.9	164.4
Other Hispanic	11.6	1.2	2.3	3.6	-4.2	0.4	8.4	10.5	19.5	30.6	-36.3	3.2	72.5

Note: Benchmark coefficient is from the White Non-Hispanic regression.
"Other Hispanic" refers to Hispanics that are neither Mexican, Puerto Rican, nor Cuban.

Table 3.8: Blinder-Oaxaca Decompositions for the Change in Gap with Mexicans

	Actual Change in the Gap, 1984-2003	Percentage Point Change						Percent of Actual Change in Gap					
		Inc	Educ	Cit	Pub	Oth	Unexp	Inc	Educ	Cit	Pub	Oth	Unexp
Puerto Rican	-10.9	-1.6	-0.2	-3.7	1.0	1.1	-7.4	-14.8	-1.4	-34.1	8.8	10.1	-68.5
Cuban	-15.6	-3.9	3.8	4.1	-0.9	-3.9	-14.9	-24.8	24.5	26.5	-6.0	-24.9	-95.2
Other Hispanic	-3.1	-2.4	-1.7	-0.2	-0.5	0.8	1.0	-78.4	-55.4	-7.7	-15.9	24.3	33.1

Note: Benchmark coefficient is from the Mexican regression.
“Other Hispanic” refers to Hispanics that are neither Mexican, Puerto Rican, nor Cuban.
Key:
Inc includes family income as a percent of poverty, work status (own and spouse), and industry of employment.
Educ includes the three categories of educational attainment.
Cit includes citizenship status and time lived in the U.S.
Pub includes the public coverage eligibility proxy and its interaction with gender and citizenship status.
Oth includes gender, marital status, presence of children, wave of interview, age, and state unemployment rate.
Unexp is the total of the ”unexplained” portions of the decomposition, including the constant term.

CHAPTER IV

Another Moment to Consider: Health Plan Choice and the Distribution of Out-Of-Pocket Medical Spending

DISCLAIMER: The research in this paper was conducted while the author was a Special Sworn Status researcher of the U.S. Census Bureau at the Michigan Census Research Data Center. Any opinions and conclusions expressed herein are those of the author and do not necessarily represent the views of the U.S. Census Bureau. All results have been reviewed to ensure that no confidential information is disclosed.

4.1 Introduction

More than seventy percent of insured Americans obtain their health insurance coverage from an employer (DeNavas-Walt, Proctor, and Smith, 2009). According to the Kaiser Family Foundation's Employer Health Benefits Survey (2005), 63 percent of covered workers are offered two or more plans. The decision to choose one health plan over potentially multiple other options has interesting implications for economists, including the willingness to pay for insurance against medical expenditures, the perceived value of fringe benefits, whether adverse selection effects the employer insurance market, and whether plans are chosen rationally and without error.

Health insurance is an especially interesting product, because unlike most consumer goods, the "price" of insurance coverage is paid over time and is contingent on how much it is used. The upfront price, or premium, is analogous to the price of traditional consumer goods, though it is complicated by the fact that in the U.S. the premium is subsidized, often heavily, by employers. The portion of the cost of care paid by the consumer during the plan year, meanwhile, is both part of the price of coverage (especially to the extent that care cannot be delayed or is sought at regular intervals) and perhaps the most important measure of the quality of the plan. Moreover, this part of the price may depend on the consumer's prior use of care, which is difficult to predict ex-ante; paradoxically, insurance coverage, which is meant to reduce financial uncertainty, has a price which is itself uncertain.

In this paper, I examine how each element of the price of coverage influences how consumers choose between the health insurance plans offered by their employers, using a national sample of employer plan offers that should provide a more general conclusion about demand for insurance than existing single-firm studies. In addition to the employee's contribution toward the premium, I derive a measure of potential out-of-pocket spending under each health plan, allowing me to consider how both the mean and the variance of medical costs over the next year affect the choice of plans. I also estimate how other elements of the plan, including supplemental coverage, plan characteristics, and plan popularity at one's establishment, influence the choice of plans. To my knowledge, this is the first paper to measure consumers' responsiveness to each aspect of the price of coverage, including the upfront premium as well as the level and variance of plan generosity.

I find that employees are less likely to choose a plan the higher is the expected level of out-of-pocket spending. The variance of the within-year "price" of coverage, however, does not influence employees' choice of plans. I find weak evidence that employees are responsive to their share of the premium, but no evidence that the types of insurance coverage beyond basic medical care that are included or optional in the plan influence their choice. Holding plan parameters constant, an employee is substantially more likely to choose a plan the larger is its market share among her coworkers, which suggests the presence of unobserved differences in plan quality and/or that employees choose the default option rather than negotiate the complex, and often small, differences between plans.

While there has been a rich literature on the choice of plans, most previous studies use data from a single employer or a small sample of firms, and nearly all consider only the level, but not the variance, of plan generosity. This paper fills some of the gaps in the literature outlined by Scanlon, Chernew, and Lave (1997), as I use data from a recent national sample of employers and employees with detailed information on the plans and personal characteristics, and take into account the important differences in the modern insurance market between health maintenance organization (HMO) plans and more traditional fee-for-service (FFS) or preferred provider organization (PPO) plans that offer less restrictive access to providers at a higher cost. Unlike most of the existing literature, I also examine the influence, or lack thereof, of non-standard coverage types and covered medical services that are either missing in most data sources or are differenced away in a one-firm sample.

The paper proceeds as follows. In section 4.2, I discuss the existing literature on plan choice and the distribution of plan generosity and this paper's contribution. I outline a theoretical framework that yields predictions on the role of each plan characteristic in the consumer's utility function in section 4.3. In section 4.4, I introduce the Medical Expenditure Panel Survey data. I address the empirical specification and discuss unconditional means in section 4.5. In section 4.6, I present the estimation results, and in section 4.7, I conclude with a discussion of other factors that may contribute to plan choice.

4.2 Related Literature

Many researchers over the last several decades have examined the question of how consumers pick among their choice set of health insurance plans. Most of these studies, though, focus on the elasticity of plan choice with respect to premiums. Scanlon, Chernew, and Lave (1997) review the literature and conclude that health insurance enrollees' response to price is reasonably well understood and consistently negative and significant across studies, but much less is known about elements other than price, and how price and these other elements interact with personal characteristics. They also are hopeful that future studies would use newer data that includes information on not just the chosen plan but also rejected alternatives, and will be not be specific to one market, geographic area, or type of consumer.

Many of the older studies reviewed by Scanlon, et al., and Feldman, et al. (1989) use data from a single employer, which calls into question the identification of premium differences at firms that likely charge the same price to every employee for any particular plan, and whether the findings are broadly applicable. Another potential source of bias is the estimation strategy used in most of these studies. Short and Taylor (1989) use a national sample of firms, but estimate two unrelated logit regressions, whether to select an HMO or FFS plan and whether, conditional on choosing some FFS plan, to select the high- or low-premium option, even though the probabilities associated with each of these decisions are likely related. Barringer and Mitchell (1994) estimate a conditional logit model on a large multi-state firm that charges different premiums by state for the same plans. These studies find that employees are especially responsive to the premium and coinsurance of non-HMO plans, and that consumers are more likely to select a plan that has additional features, including dental or mental health coverage.

Feldman, Finch, Dowd, and Cassou (1989) suggest that the independence of irrelevant alternatives (IIA) assumption necessary for the conditional logit approach is easily violated in a health insurance market with close substitutes. The sharp contrast between health maintenance organization (HMO) plans, that feature lower premiums and cost sharing but restrict access to providers and specialists, and more traditional fee-for-service (FFS) plans means that when a new HMO plan is offered, it is likely to draw enrollees mostly from the other HMO plan. They instead estimate a nested logit model, which relaxes the IIA assumption by grouping strong substitutes into a nest (HMO vs. non-HMO), estimating the probability of choosing any plan in the nest, and then estimating the conditional probability of selecting a particular plan within that nest.[1] Using data from a small group of Minneapolis-area firms, they find that consumers have a significant negative elasticity with respect to the premium, but each individual element of the coinsurance schedule is either insignificant or significant in the wrong direction, suggesting that consumers are not as responsive to cost sharing as theory would predict. They also find that older people (but not those with chronic conditions, surprisingly) are more likely to choose the non-HMO nest, and that a premium increase for one HMO plan will increase the market share for both the other HMO plans in the same nest and plans in the other nest.

Most recent papers have adopted the nested logit empirical framework. Abraham, Vogt, and Gaynor (2006) estimate a nested logit model using a subset of the same

[1] Though Feldman, et al. (1989) reject the equivalence of the conditional logit model and the nested logit model, Royalty and Solomon (1999) do not. The latter paper reports only conditional logit results, which requires that the IIA assumption holds.

data I use in this paper. They find that employees are sensitive to employee contributions to the premium for both HMO and non-HMO plans. They also interact personal characteristics with the employee premium, finding that married employees and those who file the full IRS-1040 form are less price-sensitive, and federal government workers have higher elasticities (in absolute value). Although the approach and data are strikingly similar to mine, there are some important methodological differences. They include only two incomplete measures of cost sharing under the plan, deductibles and the outpatient coinsurance rate; the effect of the deductible on plan choice is small and insignificant for both HMOs and non-HMOs, and the coinsurance rate is statistically insignificant for HMOs and significant in the wrong direction for non-HMOs. They ignore other elements of the plan besides premium, deductible, coinsurance, and plan type (HMO vs. non-HMO). Finally, they join with nearly the entire stock of the literature in neglecting higher moments of the out-of-pocket spending distribution and their effect on plan choice.

Only a handful of papers consider aspects of the medical cost distribution other than the mean. Marquis and Holmer (1986, 1996), as part of the RAND Health Insurance Experiment, use self-reported expected total medical expenditures and coinsurance information of one's randomly-assigned plan to simulate the distribution of one's expected out-of-pocket expenditure. They find that a model of asymmetric utility over monetary gains and losses, rather than eventual net income, is a better fit for consumers' behavior. Ellis (1989) uses a nonparametric method to determine which plan offered to employees of a single financial services firm is optimal at various points in the distribution. Fang, Keane and Silverman (2007) include the variance of expected expenditures, calculated as part of an effort to predict missing expenditure data in the Health and Retirement Study, as a potential source of advantageous selection.

My paper contributes to the literature on health plan choice in several ways. Satisfying the Scanlon, et al. (1997), critique, I use data from a recent national sample of employers that includes detailed information on each plan they offer their employees, reweighted to be nationally representative, from a period where HMO and non-HMO plans were both prominent in the marketplace. I calculate the distribution of a family's potential out-of-pocket expenditure in a way that completely summarizes how the deductible, differences in coinsurance by care type, and out-of-pocket and plan-paid maximums interact, using the actual expenditure profiles of others in the national sample. This calculation allows me to consider how both the mean and the variance of the distribution influence one's choice of health insurance plans, which most previous work ignores. In addition to the employee contribution and the two moments of the out-of-pocket expenditure distribution, I also consider other plan elements like supplemental coverage and covered services that are not usually available in most data sources.

4.3 Theoretical Framework

In the choice of health plans, the consumer likely has utility over many factors related to health care delivery and non-financial elements of the plans. For exposition, though, I first model the consumer's utility over a single dimension, the money remaining to spend on all other goods after accounting for necessary health care spending ("disposable income," $\tilde{y}$).

The consumer earns a fixed after-tax annual salary y, and is offered J health plans by her employer, with plan $j \in J$ requiring an annual employee contribution toward the premium of p_j.[2] For a given level of health spending h, the plan will reimburse

[2] I do not model the employer's decision over which plans to offer; see Moran, Chernew, and Hirth (2001) and

the employee's care providers for $g_j(h)$ on the dollar, requiring $(1 - g_j(h))h$ of the employee. Thus, for fixed h, plan j will leave the employee with $\tilde{y} = y - p_j - (1 - g_j(h))h$ to spend on all other goods.[3]

I model the consumer's utility as a quadratic function of disposable income, $u(\tilde{y}) = b\tilde{y} + c\tilde{y}^2$, though this specification could indicate either that the utility function is indeed quadratic, or represent a second-order approximation of the true utility function. Hanoch and Levy (1970) derive that this utility function must have $b \geq 0$ because the marginal utility $u'(\tilde{y}) = b + 2c\tilde{y}$ must be positive for all $\tilde{y}$ and $\tilde{y} = 0$ is a valid value, and $c < 0$ because $-u''/u' > 0$ for all $\tilde{y}$.[4] The consumer will select the plan that maximizes her expected utility, $Eu(\tilde{y})$. In $\tilde{y}$, y and p are both known with certainty as part of the contractual arrangement with the employer, and $g(\cdot)$ is known for any particular value of h. But h itself is subject to uncertainty about the health of her and her family. Because of the often complex coinsurance schedule $1 - g(h)$, it is both notationally and empirically convenient to consider the entire expression $(1 - g(h))h$, or her total out-of-pocket spending, to be uncertain, and distributed $H(\mu, \sigma^2)$. Her expected utility while covered by plan j is therefore

$$\begin{aligned} Eu_j(\tilde{y}) &= bE(\tilde{y}) + cE(\tilde{y}^2) \qquad (4.1) \\ &= bE[y - p_j - (1 - g_j(h))h] + cE[(y - p_j - (1 - g_j(h))h)^2] \\ &= by - bp_j - bE[(1 - g_j(h)h] + cy^2 + cp_j^2 - 2cyp_j \\ &- 2cyE[(1 - g_j(h))h] - 2cp_jE[(1 - g_j(h))h] + cE[((1 - g_j(h))h)^2] \\ &= by - bp_j - b\mu + cy^2 + cp_j^2 - 2cyp_j - 2cy\mu - 2cp_j\mu + c\mu^2 + c\sigma^2, \end{aligned}$$

Bundorf (2002). I also assume that employees sort into jobs exogenously with respect to the composition of insurance offers, a conventional, albeit potentially problematic, assumption in the literature on employees' plan choice.

[3]It is likely that the more generous is plan j, i.e. the larger is g_j, the higher h will be due to moral hazard, which will reduce $\tilde{y}$ even further for more generous plans. Furthermore, this moral hazard effect will likely result in a higher premium for plan j to cover the increased payments to care providers. The model will therefore understate the tradeoff of all other goods required for more generous health insurance.

[4]This framework requires that $\tilde{y}$ have an upper bound which keeps marginal utility positive.

where the last two terms derive from the definition of variance.[5]

Recalling that $c < 0$, equation (4.1) suggests that the consumer's expected utility from selecting a particular health plan depends on her after-tax income (with an expected positive effect) and its square (negative), the premium she pays and its square (both negative), the mean expected out-of-pocket spending (in levels) under that plan (negative), the pairwise interaction of these three terms (all positive), and the second moment of the out-of-pocket spending distribution (the sum of the distribution's variance and its squared mean, with an expected negative sign). Intuitively, larger and more uncertain out-of-pocket medical costs are unattractive, a higher premium is undesirable and increasingly so as it gets larger, and the marginal utility of income is increasing but at a decreasing rate.[6]

Of course, utility is also likely to be a function of other characteristics of the plan, such as whether the consumer is required to get a referral to see a specialist or whether the plan includes coverage for mental health or long term care, as well as characteristics of the consumer, including demographic variables and the health care needs of her and her covered family members. Let X represent the plan characteristics, which include p, p^2, μ, μ^2, σ^2, and their interactions both with each other and with y. Let Z represent the characteristics of the consumer, which includes y and its square. Most of the existing literature (e.g., Barringer and Mitchell, 1994) estimates a conditional logit model, where the expected utility function for consumer i,

$$Eu_{ij} = X_{ij}\beta + Z_i\alpha + \epsilon_{ij}$$

includes an independently and identically distributed error term with an extreme-

[5]By definition, $\sigma^2 = E[(x-\mu)^2] = E[x^2 - 2x\mu + \mu^2] = E[x^2] - 2\mu E[x] + \mu^2 = E[x^2] - 2\mu^2 + \mu^2 = E[x^2] - \mu^2$, so $E[x^2] = \sigma^2 + \mu^2$.

[6]Hanoch and Levy (1970), in the context of evaluating two portfolios of financial assets, discuss the conditions under which one portfolio strictly dominates another in a quadratic utility model. In a pairwise comparison where the means and variances of both portfolios' distributions are known, portfolio A dominates portfolio B if $(\Delta\mu^2 - \Delta\sigma^2) > 0$, where $\Delta X = X_A - X_B$ for each parameter X. To make this applicable to health plans, one would have to assume that the premium for each plan is proportional to its desirability.

value distribution.

In the model of plan choice, however, the error terms are likely to not be independent in one important variable: how much the plan restricts the consumer's choice of providers. Feldman, Finch, Dowd, and Cassou (1989) give the example of a firm that offers two plans, one traditional fee-for-service plan and another that is part of an HMO, each enrolling half of the firm's employees (so the odds ratio of selecting the HMO plan to the FFS plan is 1:1). If the firm begins to offer another HMO plan, the independence of irrelevant alternatives (IIA) assumption that is inherent in the conditional logit model requires that the odds ratio between the two plans, 1:1, remains unchanged, so that the third plan picks off enrollees equally from the two existing plans (e.g., each plan enrollees one-third of the employees). More likely, because they are less concerned about restrictions on provider choice, most of the enrollees in the new HMO plan will come from the existing HMO plan, which is its closest substitute. This reduces the odds ratio, violating independence.

Feldman, et al., instead suggest the nested logit model. The choice set of J plans can be split into two nests: HMO plans on the one hand, and plans that offer less restrictive access to providers like FFS plans or preferred provider organizations (PPOs) on the other. For example, a firm may initially offer just two HMO plans and two non-HMO plans. Within the HMO nest, a third HMO plan offered on top of two existing ones may pull enrollees disproportionately from the plan that is its closest substitute, violating IIA. Some of the new HMO plan's enrollees may also come from the non-HMO plans in the choice set, but because HMO and non-HMO plans are different enough with respect to provider choice, both non-HMO plans should lose enrollees to the new plan proportionally, so the IIA assumption is maintained across nests.[7]

The consumer selects the plan that yields the highest expected utility. Plan j among the J_k plans in nest k will yield the highest utility with probability P_{jk}, which equals the probability of choosing plan j conditional on choosing any plan in nest k, $P_{j|k}$, times the probability of choosing any plan in nest k, P_k. The nested logit model therefore estimates how the variables in X and Z, respectively, affect these two probabilities. Per McFadden (1978), the former probability can be written as

$$P_{j|k} = \frac{e^{X_{jk}\beta/\tau}}{\sum_{p\in J_k} e^{X_{pk}\beta/\tau}} = \frac{e^{X_{jk}\beta/(1-\sigma)}}{e^{I_k}},$$

where $I_k = ln(\sum_{p\in J_K} e^{X_{pk}\beta/\tau})$ is called the "inclusive value."

In this paper, there are two possible nests, HMO plans or non-HMO plans, so $k \in \{HMO, NHMO\}$. The probability of selecting an HMO plan, for example, is

$$P_{HMO} = \frac{e^{Z_{HMO}\alpha+\tau I_{HMO}}}{e^{Z_{HMO}\alpha+\tau I_{HMO}} + e^{Z_{NHMO}\alpha+\tau I_{NHMO}}}.$$

The smaller the coefficient τ, the more important are the characteristics of the nest, or plan type. If $\tau = 1$, then consumers make no distinction between HMOs and non-HMOs and are equally likely to substitute across plan type, so IIA holds. An estimate of $\tau < 1$, therefore, serves as a specification test for the nested logit model compared to the conditional logit model.

4.4 Data

The Medical Expenditure Panel Survey (MEPS) is a series of surveys conducted by the Agency for Healthcare Research and Quality. The Household Component

[7] One commonly misunderstood implication of the nested logit model is that the consumer need not first decide whether to select an HMO plan or not, and then choose from the options within that category. On the contrary, Hensher, Rose, and Greene (2005, p. 482) make clear that the nest structure of this model does not imply such a temporal ordering. More realistically, the consumer considers each plan simultaneously, but her choice is influenced by, among other things, the degree to which the plan restricts access to providers.

(HC) includes information about demographics, labor market outcomes, self-reported health status, and insurance coverage by month and source from a household survey conducted five times over a two-year period. In addition, households provide information about the source and frequency of health care they have received, which MEPS uses to survey medical providers about which conditions were treated, the treatments that were performed, drugs prescribed, and the amounts and sources of payment for care. The information from the Household and Medical Provider Components are publicly available and nationally representative (when weighted), with a new two-year panel starting each year since 1996.

MEPS also interviews employers about firm and establishment characteristics and the plans they offer to their employees. The Insurance Component (IC) includes information on up to four plans for private employers and up to 36 plans for public employers, including employee and employer contributions to premiums, coinsurance/copayments, out-of-pocket and plan-paid maximums, included and optional coverage types, covered services, and plan requirements (such as whether there is a waiting period for a pre-existing condition or if one needs a referral to see a specialist) for each offered plan.[8]

During five of the first six years of the MEPS survey, 1996 through 1999 and 2001, the IC sample purposefully included employers of individuals in the Household Component, thus allowing for a link between demographic and health variables from the household survey and details about offered plans' premiums, coinsurance schedule, and nonfinancial elements for approximately 8000 employees. MEPS warns that the link rate is low and not nationally representative; in my sample, employees at large firms (100 employees or more) are particularly over-represented, so all results are re-

[8] I include up to seven plans per employee: the union of the selected plan, the next three most popular plans overall and the next three most popular plans of the type (HMO or non-HMO) ultimately selected by the employee. The weighted average, excluding those with a single offer, was 3.2 plans per employee.

weighted to reflect the differences between the (weighted) nationally representative sample in the full HC and the smaller sample from the HC-IC Linked File.

I focus only on the intensive margin of insurance coverage; the decision whether to take up any employer-sponsored health insurance plan is left for further research. My sample excludes, therefore, all employees who were offered only one plan and thus were not choosing between plans, and anyone who refused employer coverage and either remains uninsured or received coverage though an individual with an employer who was not part of the HC-IC Linked File. I observe generally one employee per employer, and plans or insurers are not linked across employers, so I have no way of controlling for firm-, insurer- or plan-specific fixed effects.

Because the choice of health plans is often a family decision, I estimate each regression separately for unmarried individuals and married employees without a working spouse. I exclude individuals with a working spouse for two reasons. First, for most dual-earner families (about 92 percent), I observe only one of the two employed spouses' employer insurance offers; if both spouses were offered insurance, the choice set is incomplete. Second, as Feldman et al. (1989) point out, the IIA assumption may be violated for dual earner families if plans in one spouse's choice set are close substitutes for the other's offers. To properly account for the family's full choice set, I would need more information than the data can provide.

4.5 Empirical Implementation

The coefficients for the plan (X) and individual (Z) characteristics and the inclusive value (I) in the nested logit model, β, α, and τ, respectively, are estimated by maximum likelihood, separately by coverage status (single coverage or family coverage excluding those with working spouses).

Two variables among the plan characteristics (X) are of particular interest: the mean and variance of the distribution of out-of-pocket spending in one's plan. If the consumer, when choosing among her plan offers, knew with certainty how much she would spend in the coming year, only her projected out-of-pocket spending in each plan would affect her decision. While the consumer likely has some information about her imminent health needs, much remains uncertain, and hence the desire for health insurance. Ideally, the researcher would know the shape of the distribution of her expectations about her future spending, but this is a difficult question to include on a survey. Instead, I use data on medical spending by others in the MEPS sample to trace out the distribution of potential out-of-pocket spending under each plan.[9]

The MEPS Insurance Component includes detailed information on the cost-sharing elements of each offered plan. Thus, for any given pattern of spending during the year, I can calculate how much the individual would have paid out-of-pocket, taking into account whether and when the deductible (if applicable) was satisfied, differences in coinsurance or co-payments for spending on physician care, inpatient or outpatient hospital visits, and prescription drugs, and whether an out-of-pocket or plan-paid maximum is reached. The MEPS Household Component sample includes thousands of these patterns of spending; therefore, following Chapter II, I calculate how much any individual in the MEPS would have spent had he or she had any particular plan from the HC-IC Linked File.[10] I then calculate the mean and variance of the resulting distribution of possible out-of-pocket spending outcomes.[11]

[9]Another approach would be to use the distribution of out-of-pocket spending by other enrollees in the same plan, but MEPS lacks an identifier variable that would link plans across employers, so as far as I know, I observe only one enrollee per plan. Also, nearly all of the households in the Linked File appear during the first of their two years in the MEPS-HC, so I cannot use the previous year's expenditures (as does Ellis, 1989) to predict the current year's spending.

[10]The calculated mean and variance is similar when I restrict the distribution to individuals in the Linked File, rather than the full MEPS household sample.

[11]In Chapter II, I calculated the out-of-pocket and plan-paid rates, rather than the levels used here, but the process is the same.

I refine the sample of MEPS-HC observations that are included in the out-of-pocket calculation in two ways. First, if the employee has only individual coverage, I restrict the sample to other people with single coverage; likewise, employees opting for family coverage are compared to other families.[12] Second, an employee likely has some sense of the shape of the distribution based on whether she, or at least one of her family members, has a chronic condition that requires regular medical care. Therefore, if there is at least one chronic condition in her family, I included only individuals or families with a chronic condition in the distribution calculation, while limiting the sample to those without chronic conditions if no one in her family has one. There are thus four risk groups, in ascending order by mean out-of-pocket spending: single coverage without a chronic condition, family coverage without a chronic condition, single with a chronic condition, family with a chronic condition.

Table 4.1 shows how elements of the plan's coinsurance schedule contribute to whether the plan is a high- or low-generosity, and high- or low-volatility plan. As expected, plans with low mean out-of-pocket spending, and with low variance on that spending, are less likely to have deductibles, have lower coinsurance and co-payments, and are less likely to have plan-paid annual or lifetime maximums (though, somewhat surprisingly, are also less likely to have annual out-of-pocket limits). There is also little difference in most alternative coverage types, though the higher likelihood of requiring referrals suggests that HMO plans are disproportionately represented in the lower quartiles.[13]

[12] I do not impute whether any individual from the MEPS sample would have chosen single or family coverage. As a result, the uninsured and those with public coverage, who face much different prices and incentives for medical spending, are excluded from the distribution calculation. The exclusion of Medicare recipients likely eliminates many of the highest spending outcomes, but the elderly may not be the best comparison for working-age individuals and their families.

[13] Normally, HMO plans are considered less generous, not more generous. Cutler, McClellan, and Newhouse (2000) find that HMOs are better able to negotiate low prices for care, so both total and out-of-pocket spending will be lower under HMO plans, which will make HMO plans look more generous. Due to this difference in measurement, I feel that the nested logit model is the appropriate choice, as HMO plans are compared only to other HMO plans.

Table 4.1 also suggests the high degree of correlation between the mean and variance of out-of-pocket spending. High deductible plans, for example, are likely to have both high mean out-of-pocket spending, as well as a high variance of out-of-pocket spending. Indeed, the overall correlation between the mean out-of-pocket spending level and its variance is 0.62; the correlation between the difference in mean spending between any plan and the most generous plan offered by a firm and the difference in variance between any plan and the least volatile plan is 0.76. While the correlation between level and volatility is high, the estimation below allows for each moment of the out-of-pocket spending distribution to have its own impact on plan choice, controlling for the other moment.

In addition to the mean and variance of the out-of-pocket spending distribution, X includes the employee's contribution to the premium ("co-premium"). In some specifications, I also include the square of the co-premium, and the pairwise interactions between income, co-premium, and mean out-of-pocket spending level, as suggested by the model in Section 4.3. Approximately 20 percent of employees pay no portion of the premium for a single coverage plan, and 8 percent pay nothing for a family plan (Table 4.2). Six percent pay the same non-zero amount for all plans, and another 19 percent pay the same amount for each HMO plan and the same for each non-HMO plan.[14] For these individuals, there is no tradeoff between more (financial) generosity and a greater upfront price, so it should be a relatively easy decision to choose the most (financially) generous plan, all else equal. I run the estimation separately just for those who pay different premiums for each plan, as this is the group whose response to price, both upfront and during the year, is most relevant.

[14] This includes individuals who are offered exactly one HMO and one non-HMO plan, the most popular combination of plans at 22.8 percent of the sample. The next most common combination is 1 HMO and 2 non-HMO plans (11.7 percent), 2 HMO plans and no non-HMO plans (9.4 percent), and 2 HMO plans and 1 non-HMO plan (9.2 percent).

Table 4.3 displays weighted and unweighted summary statistics for the key variables, including the moments of the out-of-pocket distribution, the employee's share of the premium, and after-tax income. The effect of the weight is most apparent in the reduction in the average number of enrollees in a plan between the first three and last three columns.

Another X variable of interest is the share of enrollees in the plan. The coinsurance schedule for a plan may be confusing, and even if perfectly understood, the consumer may have trouble predicting their probability of becoming ill or injured and evaluating the consequences (Liebman and Zeckhauser, 2008). A simple way to choose a plan would be to ask a coworker, or to choose the plan that the human resources department has set as the default option. In either event, employees would be more likely to select a plan the greater is the share of employees at one's establishment that have also selected that plan, so the predicted effect is positive. Because there may be diminishing returns to the plan's popularity, I include the square of enrollee share as well.

X also includes the benefits in the plan aside from financial coverage for basic care: whether the plan includes (or has as an option for) coverage for dental, mental health, and long-term care; whether the plan requires a referral from the primary care provider to see a specialist; and indicator variables for whether the plan includes a deductible or an out-of-pocket stop loss limit. Table 4.4 presents the probability of selecting a plan given whether that plan and/or one of the employee's alternative offers included a particular service. Most services are either offered in both plans or in neither plan, so there is little variation in nonfinancial plan characteristics between offers off of which to identify the influence of these factors on plan choice.

The variables in Z do not vary between plans, so their primary purpose is to provide an estimate of adverse selection between HMO and non-HMO plans. In the most basic specification, Z includes only after-tax family income, though in subsequent specification Z also includes the square of income, indicator variables for whether the employee is female or married, age and its square, and categorical variables for race and ethnicity and educational attainment. I also include in the full regression measures of the relative health of the individual and her family, with indicators for whether the employee herself only (for single coverage employees; for employees, it is the the number of family members she covers) has a chronic condition, has a disability or limitation (i.e., answers "yes" to any Activities of Daily Living or Instrumental Activities of Daily Living question in the MEPS survey), reports fair or poor physical or mental health (on a 5-point scale), or is age 2 or younger (family coverage only); these variables will measure the degree to which there is adverse selection between HMOs and non-HMOs (Feldman, et al., 1989; see also Chapter II). Finally, Z includes characteristics of the firm, including categorical variables for the number of employees in the establishment and the region of the country, whether the firm is in an urban area, and indicator variables for union status and whether the firm self-insures.[15]

The coefficients on the variables of interest, the first and second moments of the out-of-pocket spending distribution and the premium, are identified off variation in plan offers. Figures 4.0(a), 4.0(b), and 4.0(c) plot the distribution of the differences in the mean out-of-pocket total, variance of the out-of-pocket total, and employee-paid premium, respectively, between the selected plan and each rejected alternative. While the largest portion of the mean out-of-pocket spending level distribution for, especially, singles is at zero, there is a fair amount of dispersion, particularly for

[15] In another specification, I included 11 categorical variables for industry, but none of the coefficients were significant, and the estimation had trouble converging with so many fewer degrees of freedom.

employees with family coverage.

The model predicts that as the mean out-of-pocket spending total increases, the plan becomes less attractive. Figure 4.1(a) displays the probability of selecting a plan given how much greater the out-of-pocket spending level is than the most generous plan in an employee's portfolio. Nearly 45 percent of employees select the most generous plan. As the difference in out-of-pocket burden increases, the plans are chosen with less frequency; greater than 30 percent of plans in the first decile of mean out-of-pocket spending difference are selected, but only 20 percent of the plans in the highest decile are selected.[16] The model also suggests that more uncertainty in out-of-pocket spending is undesirable, so we should observe that as the relative variance of the out-of-pocket spending distribution for a plan increases, the probability of selecting the plan decreases. Figure 4.1(b) suggests that this is weakly the case for single employees, but there is no negative relationship between the difference in out-of-pocket spending variance with the least volatile plan and the probability of selecting a plan for employees in family coverage without a working spouse. Looking at premium alone ignores differences in out-of-pocket burden and other measures of plan quality, but single employees more frequently opt for less expensive plans, while employees in family coverage appear to be more likely to consider other elements besides upfront price (Figure 4.1(c)).

4.6 Results

Table 4.5 presents the results of the nested logit regression for the single coverage subsample. For each observation, I evaluate the derivative of each X (top panel) and Z (bottom panel) variable given that observation's values for the other variables, and

[16] A linear trend line through the points in Figure 4.1(a) has a significantly negative slope for both single employees and employees with family coverage and no working spouse.

then report the mean derivative over the sample. Standard errors (in parentheses) are estimated by bootstrap over 100 iterations of the nested logit model.

The first column of Table 4.5 includes only the basic characteristics of the plan — the first two moments of the out-of-pocket spending distribution, the employee-paid premium, and whether the plan is a fee-for-service plan, plus after-tax family income, which is constant across plans and thus outside the nest. The marginal effects for most of the plan characteristics are in the expected direction. A small increase in the mean out-of-pocket spending under a plan decreases the likelihood it is chosen by 2.6 percent, which is significantly different from zero at the 99 percent confidence level. The effect of the employee contribution toward the premium is also negative and statistically significant; a small increase leads to a 1.2 percent decrease in the probability that the plan is selected. The effect of an increase in the second moment is unexpectedly positive but minuscule and not statistically significant.

In the second column, I include quadratic terms for the co-premium and income, as well as pairwise interactions between the mean out-of-pocket amount, co-premium, and income, as suggested by the theoretical model. As in the first column, a small increase in family income leads to a statistically significant but fairly small (0.1 percent) increase in the probability of choosing an HMO plan. The effect of mean out-of-pocket spending decreases in magnitude but remains statistically significant, so that a small increase in the patient's financial burden would result in a 2.1 percent decrease in the probability the plan is selected. The magnitude of the effect of the second moment remains tiny and statistically insignificant. The negative effect of the employee premium also decreases in absolute value, enough to make it statistically insignificant at conventional levels.

To account for the effect of plan elements aside from price and the financial generosity for basic care types, in specification 3, I add indicator variables for mental health care, long term care, and dental care coverage; whether the plan requires referrals, has a deductible, and has an out-of-pocket stop loss; and the share of employees at one's establishment that select that plan and its square, which will proxy for the *de facto* or *de jure* default option in the choice set. I also include additional personal characteristics outside the nest, including gender, race, education, marital status, geographic variables, union status, firm size, age and its square, and health measures.

In the full specification (column 3), the effect of a small increase in mean out-of-pocket spending still reduces the probability a plan is selected by 1.6 percent, but the magnitude decreases enough that the confidence interval includes zero. The effect of the second moment is negative, but even smaller than before and not close to statistical significance. The co-premium actually has a positive, albeit small and statistically insignificant, effect. Only one plan characteristic has an effect that is significant, either substantively or statistically — a small increase in the plan's market share in the establishment increases the probability that plan is selected by 45 percent, which is statistically different from zero at the 95 confidence level.[17]

The marginal effects in the bottom panel of column 3 provide little evidence of adverse selection between HMO and non-HMO plans. Older people, those with chronic conditions, or employees that report fair or poor physical or mental health are all less likely to choose a (generally less-generous) HMO plan, consistent with adverse selection, but none of the effects are statistically significant. I can only reject the null for the indicator for whether the employee has a limitation, which actually makes one more likely to select an HMO plan; this runs counter to the effect I would

[17]The coefficient on the quadratic term for enrollee share is positive and statistically insignificant, suggesting that the effect of plan's popularity at the establishment does not diminish as it increases.

observe if there was adverse selection.[18]

In Table 4.6, I change the specification slightly — instead of including the second moment, which is shown above to be the sum of the square of the first moment plus the variance, I include these two variables separately. The marginal effect of mean out-of-pocket spending, which now includes its square, is greater in magnitude than in the matching column in Table 4.5, and remains statistically significant at the 90 percent confidence level after the addition of other characteristics of the plan and employee. The variance has the predicted negative effect, but small enough to ignore. The mean derivative for the other coefficients, including the co-premium, are almost exactly the same as in Table 4.5.

One complication with plan choice is that many employers ask for the same employee contribution for any plan they offer, or at least the same employee contribution for any plan of the same type. Table 4.7 repeats the analysis just for the sample that pays a different co-premium for each plan, and the results are largely similar. Without controlling for plan or personal characteristics, the effect of a small increase in mean out-of-pocket level is almost identical in magnitude to estimate with the full sample of single coverage employees, but in the full model, the effect becomes larger and statistically significant for this refined sample. As one might expect, employees appear to be more price sensitive when I eliminate those who face the same premium for each plan, though the marginal effect of co-premium on plan choice is statistically insignificant when I include all plan and personal characteristics. The effect of the market share of a plan is larger; on average, a small increase in plan popularity increases its likelihood that any individual will pick the plan by 60 percent. There

[18] All of the reduction in the marginal effect of mean out-of-pocket spending between specifications 2 and 3 is due to the addition of personal characteristics. Adding these variables without adding the other plan characteristics results in a smaller (in absolute value) marginal effect of the co-premium as well, but it remains negative, while the effect of the second moment is basically unchanged.

is somewhat more evidence of adverse selection between HMO and non-HMO plans, as HMO enrollees are younger and less likely to report fair or poor physical health, but oddly more likely to have a limitation.

Some of the employees in the sample in Table 4.7 pay a larger contribution for a less generous and/or more volatile plan. In Table 4.8, I further parse the sample into just those who pay the lowest premium for the plan with the highest mean out-of-pocket spending and largest out-of-pocket spending variance. The magnitude of the effect of mean generosity on plan choice is larger than in the full sample when I do not control for other plan characteristics, but matches the estimate in Table 4.5 when I include other plan parameters. Most of the other effects match the estimates from the full sample, both in magnitude and significance.

The results for employees in family coverage, excluding those with a working spouse, are quite similar (Table 4.9). Not accounting for other plan characteristics, a small increase in the mean out-of-pocket spending level decreases the likelihood the plan is selected by a statistically-significant 2.1 percent, but this effect approaches zero and is imprecisely measured when controlling for other measures of plan quality. The second moment again has no effect on plan choice when controlling for the first moment. The employee contribution to the premium has a much smaller (and statistically insignificant) effect for employees with family coverage. Plan popularity is an even bigger factor for families; holding plan characteristics constant, a small increase in market share at the establishment increases the probability the plan will be selected by over 80 percent. There is also no evidence of adverse selection between HMO and non-HMO plans.

At the bottom of each table of regression results, I report the coefficient τ on the inclusive value for both HMOs and non-HMOs. None of the τ estimates are significantly less than one; I cannot rule out that the independence of irrelevant alternatives assumption holds, so the conditional logit model, which makes the IIA assumption, may be appropriate.[19] In Table 4.10, I report mean derivatives based on the conditional logit estimates for the full single coverage sample, without (column 1) and with (column 2) interactions between the HMO dummy variable and each basic plan characteristic, to account for the different implications of plan generosity in an HMO plan. The magnitude of the effect of mean out-of-pocket spending level is negative, as expected, but smaller and statistically insignificant than in the nested logit estimates, while the effect of the second moment is again minuscule. The co-premium has a stronger negative effect on plan choice, though I have not controlled for other plan characteristics, which diminished the employee's price sensitivity in previous specifications. An HMO plan is twenty percent more likely to be chosen, an effect which decreases only slightly when I interact the HMO dummy with plan parameters.

4.7 Conclusion

This paper estimates the relative influence of employee-paid premiums and both the level and variance of financial plan generosity in the selection of health plans from among the options offered by one's employer. Essentially, this paper is measuring the price responsiveness of demand for health insurance (and the health care costs which it insures against); where it differs with traditional consumer products is that the price has both upfront and ongoing elements, and that the ongoing portion is subject to uncertainty. In order to control for that ongoing element, I use detailed plan information from an employer survey combined with comprehensive medical

[19] This runs counter to the finding in Feldman, et al. (1989), but is consistent with Royalty and Solomon's (1999) result.

spending data from employees and their providers to derive a unique, flexible measure of potential out-of-pocket spending under any hypothetical plan.

The results suggest that the employees are more concerned with the level of this ongoing element than the upfront price, though their choice is not influenced by any uncertainty in the within-year price. Employees in single coverage, especially, are less likely to choose a plan the greater is the expected value of out-of-pocket spending under that plan, but the potential variance of out-of-pocket spending has a minuscule, imprecisely-estimated effect on the probability a plan is selected. The employee's contribution toward the premium also has a small and, when controlling for the full menu of plan parameters, imprecisely-estimated effect on plan choice. Coverage types beyond basic physician and hospital care, and plan characteristics such as required referrals and deductibles, also have little influence on the choice of plans.

One concern in the estimation of price elasticity is that an observed good is not homogeneous; if a higher price reflects higher unobserved quality, quantity demanded may not be decreasing in price. With health insurance, the most important element of plan quality is the ongoing portion of the plan's price, its potential out-of-pocket cost to the employee; using multiple moments of the distribution of out-of-pocket spending yields a more complete picture of the impact of this quality measure on demand for a plan. My model also includes other ways in which plans differ, including the structure of the plan (HMO, PPO, or FFS), the types of coverage beyond basic care that are included or optional, and whether the plan requires specialist referrals. I also include the plan's market share at the firm, which is in part a proxy measure of quality, as higher quality plans should be more popular, all else equal.

Still, among the disadvantages to using a national sample of employers rather than a single firm or small group of firms is that plan heterogeneity remains, due to missing information and measurement error. While MEPS includes information on whether many potential elements are offered in the plan, other factors which are harder or more sensitive to measure, including network breadth (in particular, whether an employee's desired care provider is in the given plan), waiting times, coverage refusals, enrollee satisfaction, and employees learning about plan quality over time, are missing. Even for the variables that are present, aside from coverage for physician and hospital care, the researcher knows at most the extensive margin of, e.g., mental health coverage. Measurement error is also likely a factor; respondents across firms may have different answers for the same health plan, based on the particular respondent's understanding of the question and his or her personal experience with the plan. At a single firm, many of these unobserved differences across employees and insurers are differenced away. A researcher using a national dataset like MEPS must therefore tradeoff the increased measurement error and missing information for increased external validity.

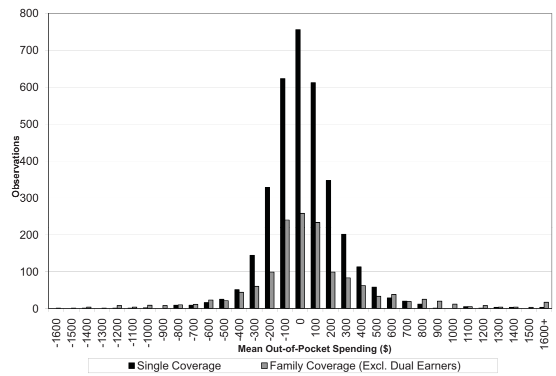
While higher-quality data may yet find that consumers are, in fact, responsive to upfront price and the variance of within-year costs, one result here stands out: the finding that the probability a plan is chosen is substantially increasing in the market share of the plan within the employee's establishment. As mentioned, this may reflect unobserved measures of plan quality. Alternatively, consumers may simply choose their company's explicit default option, if one exists, or simply "follow the herd" and choose the most popular plan. Figures 4.0(a) and 4.0(b) suggest that for many employees, there is little difference between any two plans in generosity. Even establishing that the plans are essentially equal in generosity may be too difficult or time-consuming for some employees. Gibbs, Sangl, and Burrus (1996) cite focus groups of private insurance, Medicare, and Medicaid consumers that report confusion and frustration with the information they receive at the time of enrollment on comparisons between plans, and that consumers would require a great deal more training to interpret the more detailed descriptions in the Health Employer Data Information Set report cards, which were meant to assist in plan choice. Abaluck and Gruber (2009), analyzing the recent Medicare Part D expansion that offered the elderly expanded prescription drug benefits but with a possibly overwhelming number of options, find that the vast majority of enrollees chose a plan that offered less generous benefits at a higher price than some other option available to them. Similarly, I find in Chapter II, using the same MEPS data as this study, that of employees given a choice between one plan and another less generous, more expensive option, nearly half choose the strictly dominated plan.

Employers can counter their employees' confusion by providing more information, particularly about the consequences of each plan's cost sharing elements. Schoenbaum, Spranca, Elliott, Bhattacharya, and Short (2001) find in an experimental setting that providing supplemental cost information to potential health insurance consumers increases the amount of risk that consumers are willing to bear, which they interpret as evidence that under-informed and therefore uncertain consumers overinsure in the absence of clear information. An interesting extension of my paper would consider the choice of plans separately by educational attainment, the number of years the consumer has been insured by their current (or any) employer, and other measures of consumer informedness.

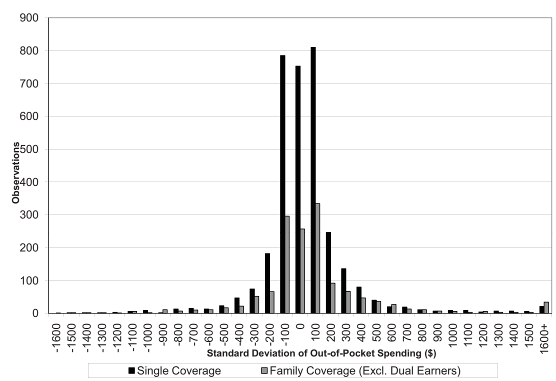
Other authors suggest that the problem is not only under-informed consumers, but under-informed researchers with models that do not properly account for advances in behavioral economics. Ellis (1989) and Marquis and Holmer (1986, 1996), both motivated by the "prospect theory" outlined by Kahneman and Tversky (1979), measure how responsive individuals are to gains and losses relative to a fixed referral point, rather than focusing on overall medical costs predicted by the expected utility model. Liebman and Zeckhauser (2008) suggest that the structure of health plans in today's market, including the expectation that relatively small co-payments will adequately control excess demand for health care and the inability of consumers to recognize that a dollar of employer contribution toward the premium is one less dollar they could have been paid in cash wages, is better predicted by a behavioral model than a more traditional expected utility model. Future models of plan choice should better account for the implications of recent behavioral studies on consumer behavior.

4.8 Figures and Tables

(a) Mean Out-of-Pocket Spending



(b) Standard Deviation of Out-of-Pocket Spending



(c) Employee Contribution

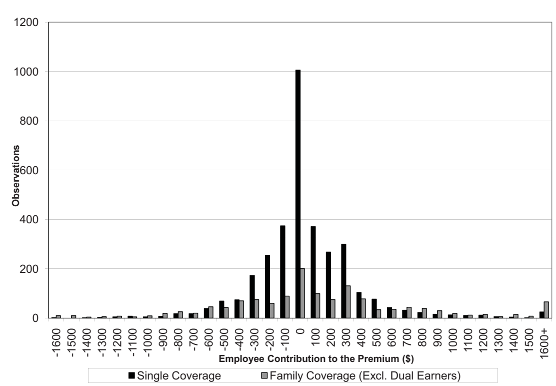
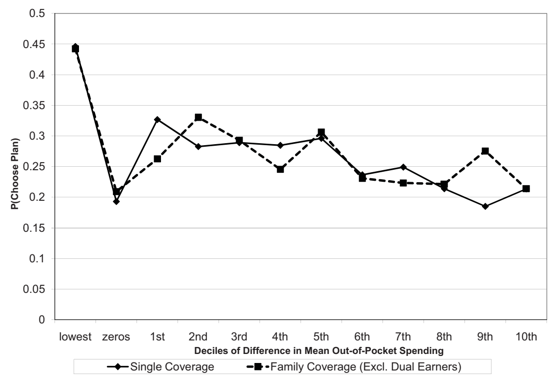
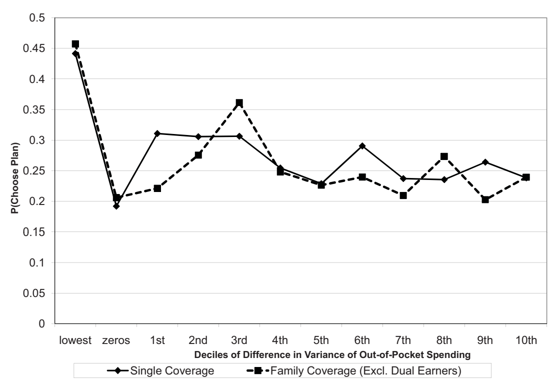


Figure 4.1: Distribution of Differences between Selected and Non-Selected Plans

(a) Mean Out-of-Pocket Spending



(b) Standard Deviation of Out-of-Pocket Spending



(c) Employee Contribution

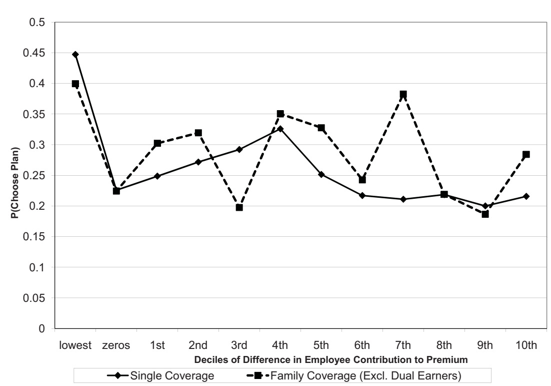


Figure 4.2: Probability of Selecting Plan, By Decile of Difference with Most Generous (Lowest Co-Premium) Plan

Table 4.1: Plan Characteristics by Out-of-Pocket Mean and Variance Quartile

	OOP Mean Quartiles				OOP Variance Quartiles			
	1st	2nd	3rd	4th	1st	2nd	3rd	4th
Any Deductible	0.7	12.3	29.8	48.4	6.9	18.7	26.8	39.0
\$0.01 to \$100	*36.8*	*16.3*	*6.6*	*1.7*	*8.0*	*9.6*	*5.7*	*3.1*
\$100.01 to \$200	*15.8*	*53.4*	*26.8*	*11.8*	*69.7*	*22.2*	*25.2*	*11.9*
\$200.01 to \$500	*5.3*	*24.3*	*37.2*	*41.8*	*14.9*	*30.3*	*40.1*	*43.7*
\$500.01 to \$1000	*0.0*	*0.0*	*2.0*	*35.3*	*0.0*	*0.8*	*10.4*	*37.9*
\$1000.01 to \$2000	*0.0*	*0.0*	*0.2*	*8.4*	*0.0*	*0.0*	*0.3*	*10.4*
\$2000.01 or more	*0.0*	*0.0*	*0.0*	*3.4*	*0.0*	*0.2*	*0.1*	*4.1*
Physician Co-Payment	79.6	89.4	85.5	62.5	80.9	89.9	84.7	61.6
\$0.01 to \$5	*44.3*	*24.9*	*14.4*	*9.9*	*41.9*	*22.9*	*16.6*	*11.6*
\$5.01 to \$10	*41.9*	*48.9*	*55.0*	*46.0*	*46.1*	*51.5*	*51.4*	*41.9*
\$10.01 to \$20	*13.8*	*25.6*	*30.0*	*42.0*	*11.9*	*25.3*	*31.3*	*44.2*
\$20.01 or more	*0.0*	*0.6*	*0.6*	*2.1*	*0.0*	*0.3*	*0.7*	*2.3*
Physician Co-Insurance	0.6	5.1	12.9	35.9	0.6	5.1	13.3	35.5
0.01% to 10%	*93.8*	*54.2*	*27.0*	*21.5*	*58.8*	*62.9*	*32.6*	*18.6*
10.01% to 20%	*6.3*	*41.7*	*65.8*	*67.3*	*41.2*	*35.7*	*62.1*	*69.0*
20.01% to 50%	*0.0*	*4.2*	*6.6*	*10.9*	*0.0*	*0.7*	*5.3*	*11.9*
50.01% or more	*0.0*	*0.0*	*0.5*	*0.4*	*0.0*	*0.7*	*0.0*	*0.5*
Hospital Co-Payment	13.4	25.1	27.8	24.1	10.5	23.6	28.8	27.8
\$0.01 to \$50	*44.7*	*35.2*	*33.5*	*19.3*	*53.6*	*35.5*	*34.5*	*17.6*
\$50.01 to \$100	*28.9*	*24.0*	*27.6*	*24.6*	*29.9*	*25.3*	*25.6*	*25.5*
\$100.01 to \$300	*25.0*	*34.9*	*34.2*	*42.0*	*15.8*	*36.7*	*34.3*	*42.1*
\$300.01 to \$500	*0.5*	*3.9*	*3.2*	*9.6*	*0.0*	*2.3*	*3.4*	*9.9*
\$500.01 or more	*0.8*	*2.0*	*1.5*	*4.5*	*0.7*	*0.3*	*2.2*	*4.8*
Hospital Co-Insurance	3.0	11.7	19.8	48.7	1.0	10.4	20.8	51.2
0.01% to 10%	*86.0*	*67.1*	*40.9*	*31.2*	*82.8*	*80.3*	*44.7*	*30.0*
10.01% to 20%	*14.0*	*31.7*	*55.4*	*60.8*	*17.2*	*19.3*	*52.8*	*61.7*
20.01% to 50%	*0.0*	*1.2*	*3.2*	*7.7*	*0.0*	*0.0*	*2.5*	*7.8*
50.01% or more	*0.0*	*0.0*	*0.5*	*0.4*	*0.0*	*0.3*	*0.0*	*0.5*
Out-of-Pocket Maximum	38.6	52.0	62.0	61.7	40.6	57.7	62.6	53.9
\$0.01 to \$1000	*31.8*	*25.9*	*16.5*	*9.1*	*29.4*	*24.6*	*19.3*	*6.5*
\$1000.01 to \$2000	*37.6*	*42.5*	*45.9*	*32.9*	*40.5*	*41.1*	*46.2*	*30.7*
\$2000.01 to \$3000	*16.1*	*22.4*	*26.9*	*28.3*	*17.4*	*26.4*	*23.5*	*28.2*
\$3000.01 or more	*14.5*	*9.1*	*10.8*	*29.7*	*12.7*	*8.0*	*11.0*	*34.6*
Annual Plan-Paid Max	1.5	2.1	2.0	5.3	1.2	2.0	2.5	5.1
\$0.01 to \$10K	*2.3*	*0.0*	*7.1*	*14.7*	*0.0*	*1.8*	*1.4*	*17.2*
\$10K to \$100K	*9.3*	*5.1*	*3.6*	*2.7*	*0.0*	*10.9*	*4.2*	*2.8*
\$100K to \$200K	*0.0*	*0.0*	*1.8*	*1.3*	*0.0*	*0.0*	*0.0*	*2.1*
\$200K or more	*88.4*	*94.9*	*87.5*	*81.3*	*100.0*	*87.3*	*94.4*	*77.9*
Lifetime Plan-Paid Max	6.1	8.7	11.5	23.0	5.5	8.5	13.2	22.3
\$0.01 to \$1 Mil	*51.7*	*63.3*	*64.3*	*62.9*	*48.4*	*66.5*	*63.4*	*62.7*
\$1 Mil to \$2 Mil	*40.2*	*28.6*	*29.0*	*31.6*	*41.5*	*24.6*	*29.4*	*32.9*
\$2 Mil or more	*8.0*	*8.1*	*6.7*	*5.5*	*10.1*	*8.9*	*7.2*	*4.4*
Require Referral (0/1)	81.1	65.7	49.6	33.8	76.8	59.3	52.9	41.0
Rx Coverage (0/1)	91.0	91.3	89.8	87.1	91.6	92.2	90.5	84.7
Dental Coverage (0/1)	67.5	69.3	69.0	68.0	69.0	72.0	67.1	65.7
LTC Coverage (0/1)	63.1	66.6	65.7	56.0	64.1	68.7	65.3	53.4
Mental Coverage (0/1)	82.9	83.0	83.2	77.9	83.5	84.6	82.1	76.9

Note: In each panel, the first row (non-italicized) is the proportion of plans in the quartile that include the plan parameter. In subsequent rows (italicized), the entry is the proportion of plans with values in the category, conditional on including the plan parameter.

Table 4.2: Employee Premium Sharing Arrangements, By Coverage Type

	Single Coverage			Family Coverage, No Working Spouse		
Employee Premium Arrangement	All	HMO	Non-HMO	All	HMO	Non-HMO
All zero	21.3	17.4	23.6	7.9	10.2	6.4
All same (non-zero)	5.4	6.8	4.6	6.7	4.9	7.9
All HMO same, all Non-HMO same	18.3	14.1	20.9	19.9	11.7	25.4
All HMO same, Non-HMO different	20.7	41.0	8.5	23.9	46.2	8.9
HMO different, all Non-HMO same	17.6	4.9	25.3	26.2	8.0	38.4
All plans different	16.6	16.0	17.0	15.4	18.9	13.0
Total	100.0	100.0	100.0	100.0	100.0	100.0

Table 4.3: Summary Statistics for Unweighted and Weighted Samples

		Unweighted			Weighted		
		All	Single	Family	All	Single	Family
Selected (0/1)	Mean	0.308	0.311	0.308	0.342	0.346	0.344
	(SD)	(0.462)	(0.463)	(0.462)	(0.475)	(0.476)	(0.475)
Employee Contribution ($)	Mean	1012.46	400.98	1545.58	926.69	383.47	1486.92
	(SD)	(1166.02)	(459.51)	(1429.65)	(1125.39)	(493.93)	(1415.78)
Mean OOP Amount ($)	Mean	413.70	276.48	518.18	435.51	302.64	573.33
	(SD)	(457.62)	(253.54)	(559.25)	(477.62)	(314.75)	(663.30)
Std Dev OOP Amount ($)	Mean	458.48	318.51	576.94	498.21	363.06	709.86
	(SD)	(940.66)	(437.74)	(1243.18)	(972.72)	(534.19)	(1622.99)
After-Tax Family Income ($)	Mean	51664	43056	38900	49746	40312	38991
	(SD)	(33960)	(29171)	(31902)	(32815)	(28538)	(32073)
Family Coverage (0/1)	Mean	0.565	0	1	0.528	0	1
	(SD)	(0.496)	(0.000)	(0.000)	(0.499)	(0.000)	(0.000)
HMO (0/1)	Mean	0.591	0.624	0.599	0.585	0.628	0.585
	(SD)	(0.492)	(0.485)	(0.490)	(0.493)	(0.483)	(0.493)
FFS (0/1)	Mean	0.056	0.051	0.063	0.069	0.065	0.077
	(SD)	(0.231)	(0.221)	(0.242)	(0.254)	(0.247)	(0.267)
Enrollee Share	Mean	0.144	0.156	0.136	0.152	0.161	0.154
	(SD)	(0.215)	(0.222)	(0.209)	(0.219)	(0.221)	(0.222)
Any Chronic Conditions*	Mean	0.730	0.452	0.927	0.670	0.416	0.944
	(SD)	(0.811)	(0.498)	(0.932)	(0.808)	(0.493)	(1.012)
Physical Health Fair/Poor*	Mean	0.157	0.105	0.284	0.128	0.080	0.274
	(SD)	(0.424)	(0.307)	(0.597)	(0.390)	(0.272)	(0.604)
Mental Health Fair/Poor*	Mean	0.062	0.040	0.110	0.049	0.031	0.094
	(SD)	(0.267)	(0.196)	(0.379)	(0.242)	(0.175)	(0.372)
Any Limit/ADL/IADL*	Mean	0.286	0.201	0.398	0.259	0.171	0.394
	(SD)	(0.523)	(0.401)	(0.629)	(0.502)	(0.376)	(0.637)
Number of Observations		11328	4890	2129	11095	4731	2108

*Note: Health variables are indicator variables for whether a single coverage employee has each condition, and the number of covered family members with each condition for employees in family coverage.

Table 4.4: Variation in Non-Financial Plan Elements between Selected Plan and Each Non-Selected Offered Plan

	Any Two Plans				Two HMO Plans				Two Non-HMO Plans			
	Both	Selected	Not Selected	Neither	Both	Selected	Not Selected	Neither	Both	Selected	Not Selected	Neither
Prescription Drug Coverage	89.2	2.0	1.4	7.4	89.9	2.0	1.3	6.7	87.3	1.2	0.9	10.5
Dental Coverage	62.1	5.8	7.1	25.0	60.2	5.8	4.9	29.1	67.7	7.0	6.7	18.6
Long Term Care Coverage	64.7	1.4	1.1	32.8	65.9	1.7	1.2	31.2	62.4	0.7	0.5	36.4
Mental Health Coverage	81.9	2.5	1.9	13.8	83.2	2.4	2.1	12.3	77.8	2.0	2.1	18.0
Disability Coverage	0.7	0.3	0.2	72.3	0.3	0.2	0.1	73.0	1.9	0.7	0.4	73.1
Life Insurance Coverage	1.6	0.5	0.4	71.0	1.0	0.5	0.2	71.9	4.1	0.8	0.2	70.9
Mammograms	67.7	2.9	2.0	27.4	72.8	2.4	3.2	21.5	54.9	1.7	0.8	42.6
Chiropractor	33.9	10.7	15.9	39.5	34.0	13.1	10.2	42.8	32.3	8.0	21.7	38.0
Other Non-Physician	36.1	6.7	5.6	25.0	32.4	10.7	2.2	28.2	41.8	4.9	1.0	28.2
Physicals	47.3	10.2	4.0	11.9	57.7	1.6	3.9	10.3	30.4	19.4	7.2	19.1
Pap Smaers	58.0	2.8	1.8	10.8	58.3	2.3	2.5	10.6	57.5	2.8	1.1	14.5
Prenatal Care	53.8	2.2	0.9	16.7	57.3	1.5	1.1	13.8	45.3	4.7	0.9	25.0
Adult Immunizations	46.6	4.3	2.7	19.8	51.6	2.1	3.4	16.6	35.8	6.8	3.8	29.6
Child Immunizations	52.9	2.3	1.2	17.1	55.6	1.8	1.8	14.3	47.9	2.0	0.7	25.4
Well-Infant Care	58.0	3.0	1.7	10.8	59.3	2.0	2.4	9.9	56.9	2.7	0.7	15.6
Well-Child Care	50.3	3.5	1.9	17.7	54.9	1.7	2.1	14.9	40.9	6.9	2.0	26.2
Any Preventive Care	61.4	1.7	0.8	9.5	62.0	1.7	1.2	8.8	59.8	2.1	0.5	13.6
Referral for Specialist Required	38.0	25.0	16.2	20.8	64.7	3.4	28.9	3.0	9.6	14.6	7.4	68.4
Plan Has Refused Coverage	4.0	3.8	5.4	86.8	3.1	4.0	1.8	91.2	8.5	6.1	6.3	79.1
Waiting Period for Pre-Existing	8.4	4.2	6.8	80.6	6.6	4.6	2.5	86.3	18.2	5.4	6.6	69.7
Has Deductible	7.3	14.4	17.2	61.1	1.2	17.4	0.8	80.6	23.5	16.5	16.8	43.3
Has a Lifetime Limit	6.2	4.1	6.1	83.7	4.4	7.0	1.4	87.1	10.6	1.9	2.1	85.4
Has an Annual Plan-Paid Limit	1.5	0.8	1.3	96.4	1.3	1.2	0.5	97.0	1.9	0.7	1.1	96.3
Has an Annual OOP Limit	51.2	11.6	15.1	22.2	41.2	17.4	8.4	33.1	80.0	7.7	6.2	6.1
Number of Pairs	7845				3639				1074			

Table 4.5: Mean Derivatives from Nested Logit Regression of Plan Choice, Single Coverage Only

	(1)		(2)		(3)	
Within Nest						
Mean OOP Amount	-0.026	***	-0.021	**	-0.016	
	(0.010)		(0.010)		(0.012)	
2nd Moment, OOP Amount	0.000026		0.000026		-0.0000013	
	(0.000097)		(0.000099)		(0.0001179)	
Employee Premium*	-0.012	***	-0.0084		0.00031	
	(0.005)		(0.0051)		(0.00716)	
FFS	-0.017		-0.016		-0.013	
	(0.061)		(0.070)		(0.061)	
Enrollee Share*					0.452	**
					(0.187)	
Quadratics and Interactions	N		Y		Y	
Supplemental Coverage	N		N		Y	
Plan Characteristics	N		N		Y	
Nest - Picking HMO						
After-Tax Family Income*	0.0010	***	0.0016	***	-0.00024	
	(0.0004)		(0.0004)		(0.00068)	
Any Chronic Conditions					-0.019	
					(0.023)	
Fair/Poor Physical Health					-0.0032	
					(0.0369)	
Fair/Poor Mental Health					-0.090	
					(0.063)	
Any Limit/ADL/IADL					0.054	**
					(0.025)	
Age*					-0.0012	
					(0.0010)	
Quadratics and Interactions	N		Y		Y	
Demographics	N		N		Y	
Firm Characteristics	N		N		Y	
τ HMO	1.304	*	1.378	**	1.337	
	(0.173)		(0.182)		(0.217)	
τ Non-HMO	0.951		1.111		0.931	
	(0.242)		(0.255)		(0.442)	
Number of plans	4731		4731		4731	
Number of employees	1479		1479		1479	

Note: Specifications 2 and 3 include quadratic terms for variables marked with * and pairwise interactions between Mean OOP Amount, Premium, and Income. All regressions weighted to account for non-random missing observations in HC-IC Linked File sample.
*** - Significantly different from zero at the 99% confidence level ** - 95% confidence level * - 90% confidence level

Table 4.6: Mean Derivatives from Nested Logit Regression of Plan Choice, Single Coverage Only, Splitting The Second Moment

	(1)		(2)		(3)	
Within Nest						
Mean OOP Amount*	-0.034	***	-0.032	***	-0.029	*
	(0.011)		(0.010)		(0.016)	
Variance OOP Amount	-0.000010		-0.000056		-0.000070	
	(0.000205)		(0.000202)		(0.000172)	
Employee Premium*	-0.012	***	-0.0084		0.00004	
	(0.004)		(0.0069)		(0.00723)	
FFS	-0.007		-0.002		-0.005	
	(0.061)		(0.067)		(0.056)	
Enrollee Share*					0.424	**
					(0.185)	
Quadratics and Interactions	N		Y		Y	
Supplemental Coverage	N		N		Y	
Plan Characteristics	N		N		Y	
Nest - Picking HMO						
After-Tax Family Income*	0.0010	***	0.0016	***	-0.00008	
	(0.0004)		(0.0005)		(0.00063)	
Any Chronic Conditions					-0.019	
					(0.027)	
Fair/Poor Physical Health					-0.0025	
					(0.0415)	
Fair/Poor Mental Health					-0.090	
					(0.066)	
Any Limit/ADL/IADL					0.055	**
					(0.028)	
Age*					-0.0012	
					(0.0009)	
Quadratics and Interactions	N		Y		Y	
Demographics	N		N		Y	
Firm Characteristics	N		N		Y	
τ HMO	1.230		1.280		1.277	
	(0.181)		(0.186)		(0.219)	
τ Non-HMO	1.011		1.187		1.013	
	(0.227)		(0.234)		(0.386)	
Number of plans	4731		4731		4731	
Number of employees	1479		1479		1479	

Note: Specifications 2 and 3 include quadratic terms for variables marked with * and pairwise interactions between Mean OOP Amount, Premium, and Income. All regressions weighted to account for non-random missing observations in HC-IC Linked File sample.
*** - Significantly different from zero at the 99% confidence level ** - 95% confidence level * - 90% confidence level

Table 4.7: Mean Derivatives from Nested Logit Regression of Plan Choice, Single Coverage Only, Different Premium for Each Plan

	(1)		(2)		(3)	
Within Nest						
Mean OOP Amount	-0.021		-0.018		-0.037	*
	(0.013)		(0.012)		(0.022)	
2nd Moment, OOP Amount	0.00012		0.00012		0.00047	
	(0.00032)		(0.00023)		(0.00034)	
Employee Premium*	-0.015	**	-0.0138	**	-0.0042	
	(0.006)		(0.0065)		(0.0080)	
FFS	-0.073		-0.079		-0.127	
	(0.064)		(0.071)		(0.124)	
Enrollee Share*					0.597	*
					(0.333)	
Quadratics and Interactions	N		Y		Y	
Supplemental Coverage	N		N		Y	
Plan Characteristics	N		N		Y	
Nest - Picking HMO						
After-Tax Family Income*	-0.000047		0.0004		-0.00144	
	(0.000772)		(0.0010)		(0.00092)	
Any Chronic Conditions					-0.012	
					(0.035)	
Fair/Poor Physical Health					-0.130	**
					(0.059)	
Fair/Poor Mental Health					-0.088	
					(0.103)	
Any Limit/ADL/IADL					0.067	*
					(0.036)	
Age*					-0.0028	*
					(0.0015)	
Quadratics and Interactions	N		Y		Y	
Demographics	N		N		Y	
Firm Characteristics	N		N		Y	
τ HMO	1.325		1.309		1.050	
	(0.278)		(0.309)		(0.287)	
τ Non-HMO	0.868		0.887		2.039	
	(0.293)		(0.361)		(0.740)	
Number of plans	2851		2851		2851	
Number of employees	774		774		774	

Note: Specifications 2 and 3 include quadratic terms for variables marked with * and pairwise interactions between Mean OOP Amount, Premium, and Income. All regressions weighted to account for non-random missing observations in HC-IC Linked File sample.
*** - Significantly different from zero at the 99% confidence level ** - 95% confidence level * - 90% confidence level

Table 4.8: Mean Derivatives from Nested Logit Regression of Plan Choice, Single Coverage Only, Worst Plan Has Lowest Premium

	(1)		(2)		(3)	
Within Nest						
Mean OOP Amount	-0.033	**	-0.021		-0.016	
	(0.015)		(0.016)		(0.029)	
2nd Moment, OOP Amount	-0.000059		-0.00012		-0.00019	
	(0.000340)		(0.00030)		(0.00055)	
Employee Premium*	-0.020	***	-0.0055		-0.0014	
	(0.005)		(0.0091)		(0.0091)	
FFS	-0.054		-0.050		0.075	
	(0.082)		(0.089)		(0.103)	
Enrollee Share*					0.640	**
					(0.271)	
Quadratics and Interactions	N		Y		Y	
Supplemental Coverage	N		N		Y	
Plan Characteristics	N		N		Y	
Nest - Picking HMO						
After-Tax Family Income*	0.0018	***	0.0023	***	0.0012	
	(0.0006)		(0.0006)		(0.0018)	
Any Chronic Conditions					-0.075	
					(0.063)	
Fair/Poor Physical Health					-0.0087	
					(0.0795)	
Fair/Poor Mental Health					0.065	
					(0.100)	
Any Limit/ADL/IADL					0.071	
					(0.066)	
Age*					-0.0034	
					(0.0022)	
Quadratics and Interactions	N		Y		Y	
Demographics	N		N		Y	
Firm Characteristics	N		N		Y	
τ HMO	1.032		1.143		0.846	
	(0.477)		(0.510)		(0.627)	
τ Non-HMO	0.74		0.77		0.863	
	(0.293)		(0.307)		(0.442)	
Number of plans	1004		1004		1004	
Number of employees	369		369		369	

Note: Specifications 2 and 3 include quadratic terms for variables marked with * and pairwise interactions between Mean OOP Amount, Premium, and Income. All regressions weighted to account for non-random missing observations in HC-IC Linked File sample.
*** - Significantly different from zero at the 99% confidence level ** - 95% confidence level * - 90% confidence level

Table 4.9: Mean Derivatives from Nested Logit Regression of Plan Choice, Family Coverage Only with No Working Spouse

	(1)		(2)		(3)	
Within Nest						
Mean OOP Amount	-0.021	**	-0.021	*	-0.006	
	(0.010)		(0.012)		(0.011)	
2nd Moment, OOP Amount	0.000050		0.000058		0.000028	
	(0.000149)		(0.000205)		(0.000116)	
Employee Premium*	-0.001		-0.0006		0.00125	
	(0.002)		(0.0032)		(0.00263)	
FFS	-0.006		-0.005		0.016	
	(0.078)		(0.084)		(0.057)	
Enrollee Share*					0.806	**
					(0.325)	
Quadratics and Interactions	N		Y		Y	
Supplemental Coverage	N		N		Y	
Plan Characteristics	N		N		Y	
Nest - Picking HMO						
After-Tax Family Income*	0.000038		0.000099		-0.00316	***
	(0.000460)		(0.000785)		(0.00093)	
Number with Any Chronic Conditions					0.013	
					(0.023)	
Number with Fair/Poor Physical Health					-0.042	
					(0.033)	
Number with Fair/Poor Mental Health					-0.031	
					(0.042)	
Number with Any Limit/ADL/IADL					0.021	
					(0.028)	
Number Age 2 and Under					0.002	
					(0.040)	
Age*					0.000040	
					(0.001815)	
Quadratics and Interactions	N		Y		Y	
Demographics	N		N		Y	
Firm Characteristics	N		N		Y	
τ HMO	1.122		1.171		1.314	
	(0.212)		(0.216)		(0.326)	
τ Non-HMO	1.138		1.342		1.532	
	(0.302)		(0.344)		(0.758)	
Number of plans	2105		2105		2105	
Number of employees	648		648		648	

Note: Specifications 2 and 3 include quadratic terms for variables marked with * and pairwise interactions between Mean OOP Amount, Premium, and Income. All regressions weighted to account for non-random missing observations in HC-IC Linked File sample.
*** - Significantly different from zero at the 99% confidence level ** - 95% confidence level * - 90% confidence level

Table 4.10: Mean Derivatives from Conditional Logit Regression of Plan Choice, Single Coverage Only

	(1)		(2)	
Mean OOP Amount	-0.0070		-0.0089	
	(0.0098)		(0.0130)	
2nd Moment, OOP Amount	0.0000057		0.00000047	
	(0.0000837)		(0.0002098)	
Employee Premium	-0.0092	**	-0.0100	**
	(0.004)		(0.0039)	
After-Tax Family Income	-0.00067		-0.00063	
	(0.00050)		(0.00046)	
FFS	0.049		0.069	
	(0.066)		(0.071)	
HMO	0.200	***	0.181	***
	(0.030)		(0.052)	
HMO Interactions	N		Y	
Number of plans	4731		4731	
Number of employees	1479		1479	

Note: Specifications 2 and 3 include quadratic terms for variables marked with * and pairwise interactions between Mean OOP Amount, Premium, and Income. All regressions weighted to account for non-random missing observations in HC-IC Linked File sample.
*** - Significantly different from zero at the 99% confidence level ** - 95% confidence level * - 90% confidence level

CHAPTER V

Conclusion

This volume explores the extensive and intensive margins of health insurance coverage, focusing on three different aspects of consumers' interaction with the health care system: the choice of employer-sponsored health plans, how health insurance coverage generosity influences medical spending, and disparities in the uninsured rate across racial and ethnic groups. Because all three chapters use data from national surveys, the results should be more broadly applicable than studies that examine a single firm or a small group of health insurance policyholders.

The results in the first chapter suggest that consumers do respond to the inherent moral hazard incentive in the health insurance system: more generous coverage, which lowers the out-of-pocket price for medical care, encourages more spending. But inefficiency is not an automatic result. Managed care, an innovation of health maintenance organizations that has extended into more traditional health plans, breaks this relationship between generosity and spending to some extent. Moreover, there is little evidence of adverse selection; if consumers with greater health care needs are not opting for more expensive, more generous health plans, then the lack of information sharing in today's employer-sponsored health insurance system, and perhaps tomorrow's health insurance exchange, need not lead to spiraling premiums or inefficient provision of insurance coverage.

In the second chapter, I find that differences in citizenship, education, and labor market outcomes do not fully account for the growth in the disparity in health insurance coverage between Hispanics and non-Hispanic Whites. If universal coverage is the societal goal, policymakers must consider both supply and demand factors that lead to the inequality in coverage rates.

I conclude in the third chapter that, as expected, consumers desire higher levels of coverage against health risks, but do not seem attracted to lower levels of variability in their out-of-pocket spending. Health insurance, and its consequences for how the insured person receives necessary or desired medical care, is complicated, and consumers are notorious for their misunderstanding of future health care risks, so perhaps it is not surprising that consumers do not select plans optimally.

Policymakers took an enormous step toward reforming the health insurance system in March, 2010, especially in extending affordable coverage options to nearly all uninsured American citizens. Further improvements will likely require consumers, with the help of both government and private industry, to educate themselves. A consumer with a better understanding of the elements of health insurance plans, especially managed care and the coinsurance schedule, will likely select the health plan that best protects her and her family from the financial risk of large medical expenditure. Better informed consumers could result in the adverse selection we do not currently observe in the employer system, though this may be countered by a desire by healthier individuals to subsidize the less fortunate members of society, a motive we observe in the progressivity of the tax and transfer system and the popularity of charitable giving. Further, if employees understand the general equilibrium effects of their demand for health insurance and health care — that their demand for more generous coverage results in lower wages for all, and that overspending on medical care raises premiums for all — health insurance could become less a way to prepay for care, and instead true insurance against unexpected medical spending.

BIBLIOGRAPHY

[1] Abaluck, Jason T. and Jonathan Gruber, 2009. "Choice Inconsistencies Among the Elderly: Evidence from Plan Choice in the Medicare Part D Program," NBER working paper 14759, February 2009.

[2] Abbring, Japp H., Pierre-Andre Chiappori, and Jean Pinquet, 2003. "Moral Hazard and Dynamic Insurance Data," *Journal of the European Economic Association*, **1**:767-820.

[3] Abraham, Jean Marie, Wiliam B. Vogt, and Martin S. Gaynor, 2006. "How Do Households Choose Their Employer-Based Health Insurance," *Inquiry* **43**:315-332.

[4] Agency for Healthcare Research and Quality, Center for Cost and Financing Studies, 2003. "Household Component - Insurance Component Linked Data, 1999: Documentation," AHRQ mimeo.

[5] Alegria, Margerita, Zhun Cao, Thomas McGuire, and David Takeuchi, 2005. "Health Insurance for Vulnerable Populations: Understanding Differences across Asian American and Latino Subgroups in the United States," Economic Research Initiative on the Uninsured Working Paper 41.

[6] Altonji, Joseph G., and Rebecca M. Blank, 1999. "Race and Gender in the Labor Market," *Handbook of Labor Economics*, Volume 3 (Amsterdam: Elsevier), 3143-3259.

[7] Arrow, Kenneth J., 1963. "Uncertainty and the Welfare Economics of Medical Care," *American Economic Review*, **53**(5):941-973.

[8] Ashraf, Javed, 1996. "Is Gender Pay Discrimination on the Wane? Evidence from Panel Data, 1968-1989," *Industrial and Labor Relations Review* 49(3): 537-546.

[9] Barringer, Melissa W. and Olivia S. Mitchell, 1994. "Workers' Preferences Among Company-Provided Health Insurance Plans," *Industrial and Labor Relations Review* **48**(1):141-152.

[10] Barsky, Robert B., F. Thomas Juster, Miles S. Kimball, and Matthew D. Shapiro, 1997. "Preference Parameters and Behavioral Heterogeneity: An Experimental Approach in the Health and Retirement Study," *Quarterly Journal of Economics* 112(2):537-579.

[11] Berk, Marc L., Leigh Ann Albers, Claudia L. Schur, 1996. "The Growth in the U.S. Uninsured Population: Trends in Hispanic Subgroups, 1977 to 1992," *American Journal of Public Health* 86: 572-576.

[12] Blinder, Alan S., 1973. "Wage Discrimination: Reduced Form and Structural Estimates," *Journal of Human Resources* 8(4): 436-455.

[13] Borjas, George J., 2003. "Welfare Reform, Labor Supply, and Health Insurance in the Immigrant Population," *Journal of Health Economics* 22: 933-958.

[14] Borjas, George J., and Lynette Hilton, 1996. "Immigration and the Welfare State: Immigrant Participation in Means-Tested Entitlement Programs," *Quarterly Journal of Economics* 111(2): 575-604.

[15] Bound, John, and Richard B. Freeman, 1992. "What Went Wrong? The Erosion of Relative Earnings and Employment Among Young Black Men in the 1980's," *Quarterly Journal of Economics* 107(1): 201-232.

[16] Buchmueller, Thomas C. and Paul J. Feldstein, 1997. "The effect of price on switching among health plans," *Journal of Health Economics*, **16**:231-247.

[17] Buchmueller, Thomas C., Anthony T. LoSasso, Ithai Lurie, and Sarah Dolfin, 2007. "Immigrants and Employer-Sponsored Health Insurance," *Health Services Research* 42(1): 286-310.

[18] Bundorf, M. Kate, 2002. "Employee demand for health insurance and employer health plan choices," *Journal of Health Economics* **21**(1):65-88.

[19] Camarota, Steven A., and James R. Edwards., Jr., 2000. "Without Coverage: Immigration's Impact on the Size and Growth of the Population Lacking Health Insurance," Center for Immigration Studies working paper.

[20] Campos-Outcalt, D., C. Bay, A. Dellapena, and M.K. Cota, 2003. "Motor Vehicle Crash Fatalities by Race/Ethnicity in Arizona, 1990-96," *Injury Prevention* 9(3): 251-256.

[21] Cardon, James H. and Igal Hendel, 2001. "Asymmetric Information in Health Insurance: Evidence from the National Medical Expenditure Survey," *RAND Journal of Economics*, **32**(3):408-427.

[22] Carrasquillo, Olveen, David U. Himmelstein, Steffie Woolhandler, and David H. Bor, 1999. "Going Bare: Trends in Health Insurance Coverage, 1989 to 1996," *American Journal of Public Health* 89: 36-42.

[23] Cawley, John and Tomas Philipson, 1999. "An Empirical Examination of Information Barriers to Trade in Insurance," *American Economic Review*, **89**(4):827-846.

[24] Cawley, J. and K.I. Simon, 2005. "Health Insurance Coverage and the Macroeconomy," *Journal of Health Economics* 24(2): 299-315.

[25] Cebul, Randall, James B. Rebitzer, Lowell J. Taylor, and Mark Votruba, 2008. "Unhealthy Insurance Markets: Search Frictions and the Cost and Quality of Health Insurance," NBER working paper 14455, October 2008.

[26] Chernew, Michael, Kevin Frick, and Catherine G. McLaughlin, 1997. "The demand for health insurance coverage by low-income workers: can reduced premiums achieve full coverage?" *Health Services Research*, **32**(4):453-472.

[27] Chiappori, Pierre-Andre, Franck Durand, and Pierre-Yves Geoffard, 1998. "Moral Hazard and the Demand for Physician Services: First Lessons from a French Natural Experiment," *European Economic Review*, **42**(3-5):499-511.

[28] Chiappori, Pierre-Andre and Bernard Salanie, 2000. "Testing for Asymmetric Information on Insurance Markets," *Journal of Political Economy*, **108**(1):56-78.

[29] Claxton, Gary, Isadora Gil, Benjamin Finder, Jon Gabel, Jeremy Pickreign, Heidi Whitmore, Samantha Hawkins, 2005. *Employer Health Benefits: 2005 Annual Survey.* Menlo Park and Chicago: Henry J. Kaiser Family Foundation and Health Research and Educational Trust.

[30] Cotton, Jeremiah, 1988. "On the Decomposition of Wage Differentials," *Review of Economics and Statistics* 70: 236-243.

[31] Cutler, David M., Bryan Lincoln, and Richard J. Zeckhauser, 2009. "Selection Stories: Understanding Movement Across Health Plans," NBER working paper 15164, July 2009.

[32] Cutler, David M., Mark McClellan, and Joseph P. Newhouse, 2000. "How does managed care do it?" *RAND Journal of Economics*, **31**(3):526-548.

[33] Cutler, David M. and Sarah J. Reber, 1998. "Paying for health insurance: the tradeoff between competition and adverse selection," *Quarterly Journal of Economics*, **113**(2):433-466.

[34] Cutler, David M. and Richard J. Zeckhauser, 2000. "The Anatomy of Health Insurance," *Handbook of Health Economics*, (Amsterdam: Elsevier 2000), 563-643.

[35] DeNavas-Walt, Carmen, Bernadette D. Proctor, and Jessica Smith, 2009. "Income, Poverty, and Health Insurance Coverage in the United States, 2008," U.S. Census Bureau Report P60-236.

[36] Duan, Naihua, Willard G. Manning, Jr., Carl N. Morris, and Joseph P. Newhouse, 1983. "A Comparison of Alternative Models for the Demand for Medical Care," *Journal of Business and Economic Statistics*, **1**(2):115-126.

[37] Einav, Liran, Amy Finkelstein, and Mark R. Cullen, 2008. "Estimating welfare in insurance markets using variation in prices," NBER working paper 14414, October 2008.

[38] Ellis, Randall P., 1989. "Employee Choice of Health Insurance," *The Review of Economics and Statistics*, **71**(2):215-223.

[39] Engen, Eric M., William G. Gale, Cori E. Uccello, Christopher D. Carroll, and David I. Laibson, 1999. "The Adequacy of Household Saving," *Brookings Papers on Economics* 1999(2): 65-187.

[40] Fairlie, Robert W., 2005. "An Extension of the Blinder-Oaxaca Decomposition Technique to Logit and Probit Models," *Journal of Economic and Social Measurement* 30(4): 305-316.

[41] Fairlie, Robert W., and William A. Sundstrom, 1999. "The Emergence, Persistence, and Recent Widening of the Racial Unemployment Gap," *Industrial and Labor Relations Review* 52(2): 252-270.

[42] Fang, Hanming, Michael P. Keane, and Dan Silverman, 2008. "Sources of Advantageous Selection: Evidence from the Medigap Insurance Market," *Journal of Political Economy*, **116**(2):303-350.

[43] Feldman, Roger D. and Bryan E. Dowd, 1982. "Simulation of a Health Insurance Market with Adverse Selection" *Operations Research*, **30**(6):1027-1042.

[44] Feldman, Roger and Bryan Dowd, 1991. "Must adverse seleciton cause premium spirals?" *Journal of Health Economics*, **10**:349-357.

[45] Feldman, Roger and Bryan Dowd, 2000. "Risk segmentation: goal or problem?" *Journal of Health Economics*, **19**:499-512.

[46] Feldman, Roger, Michael Finch, Bryan Dowd and Steven Cassou, 1989. "The demand for employment-based health insurance plans," *Journal of Human Resources*, **24**(1):117-142.

[47] Finkelstein, Amy and Kathleen McGarry, 2006. "Multiple dimensions of private information: evidence from the long-term care insurance market," *American Economic Review*, **96**(4):938-958.

[48] Finkelstein, Amy and James Poterba, 2004. "Adverse Selection in Insurance Markets: Policyholder Evidence from the U.K. Annuity Market," *Journal of Political Economy*, **112**(1):183-208.

[49] Finkelstein, Amy and James Poterba, 2006. "Testing for Adverse Selection with 'Unused Observables,'" NBER working paper 12112, March 2006.

[50] Fronstin, Paul, Lawrence G. Goldberg, and Philip K. Robins, 1997. "Differences in Private Health Insurance Coverage for Working Male Hispanics," *Inquiry* 34: 171-180.

[51] Gibbs, Deborah A., Judith A. Sangl, and Barri Burrus, 1996. "Consumer perspectives on information needs for health plan choice," *Health Care Financing Review* **18**(1):55-73.

[52] Gresenz, Carole Roan, Jeanette Rogowski, and Jose J. Escarce, 2007. "Social Networks and Access to Health Care Among Mexican-Americans," NBER Working Paper 13460.

[53] Hamilton, Erin R., Robert A. Hummer, Xiuhong H. You, and Yolanda C. Padilla, 2006. "Health Insurance and Health-Care Utilization of U.S.-Born Mexican-American Children," *Social Science Quarterly* 87(5): 1280-1294.

[54] Hanoch, Giora and Haim Levy, 1970. "Efficient Portfolio Selection with Quadratic and Cubic Utility," *Journal of Business* **43**(2):181-189.

[55] Hendee, William R., 1991. "Hispanic Health in the United States," *Journal of the American Medical Association* 265(2): 248-252.

[56] Hensher, David A., John M. Rose, and William H. Greene, 2005. *Applied Choice Analysis: A Primer.* New York: Cambridge University Press.

[57] Herrera, Guillermo A., Zhen Zhao, and R. Monina Klevens, 2001. "Variation in Vaccination Coverage Among Children of Hispanic Ancestry," *American Journal of Preventive Medicine* 20(4): 69-74.

[58] Hirth, Richard A., Reagan A. Baughman, Michael E. Chernew, and Emily C. Shelton, 2006. "Worker preferences, sorting and aggregate patterns of health insurance coverage," *International Journal of Health Care Finance and Economics*, **6**:259-277.

[59] Jann, Ben, 2005. "Standard Errors for the Blinder-Oaxaca Decomposition," working paper.

[60] Kahneman, Daniel and Amos Tversky, 1979. "Prospect Theory: An Analysis of Decision Under Risk," *Econometrica* **47**:263-291.

[61] Kandilov, Amy M. Gass, 2007. "Health Insurance and Labor Supply Among Recent Immigrants following the 1996 Welfare Reform: Examining the Effect of the Five-Year Residence Requirement," working paper.

[62] Ku, Leighton, and Brian Bruen, 1999. "The Continuing Decline in Medicaid Coverage," *New Federalism: Issues and Options for States*, Urban Institute.

[63] Liebman, Jeffrey and Richard Zeckhauser, 2008. "Simple Humans, Complex Insurance, Subtle Subsidies," NBER working paper 14330, September 2008.

[64] Manning, Willard G., Joseph P. Newhouse, Naihua Duan, Emmett B. Keeler, and Arleen Leibowitz, 1987. "Health Insurance and the Demand for Medical Care: Evidence from a Randomized Experiment," *American Economic Review*, **77**(3):251-277.

[65] Marquis, M. Susan and Martin R. Holmer, 1986. "Choice Under Uncertainty and the Demand for Health Insurance," The Rand Corporation, N-2516-HHS.

[66] Marquis, M. Susan and Martin R. Holmer, 1996. "Alternative Models of Choice Under Uncertainty and Demand for Health Insurance," *Review of Economics and Statistics* **78**(3):421-427.

[67] Medina, Jose F., and Chak-Tong Chau, 1998. "Credit Card Usage Behavior Between Anglos and Hispanics," *Hispanic Journal of Behavioral Sciences* 20(4): 429-447.

[68] Monheit, Alan C. and Jessica Vistnes, 1999. "Health Insurance Availability at the Workplace: How Important are Worker Preferences?" *Journal of Human Resources*, **34**(4):770-785.

[69] Monheit, Alan C., and Jessica Primoff Vistnes, 2000. "Race/Ethnicity and Health Insurance Status: 1987 and 1996," *Medical Care Research and Review* 57(Supp. 1): 11-35.

[70] Moran, John R., Michael E. Chernew, and Richard A. Hirth, 2001. "Preference Diversity and the Breadth of Employee Health Insurance Options," *Health Services Research*, **36**(5):911-934.

[71] Neumark, David, 1988. "Employers' Discriminatory Behavior and the Estimation of the Wage Discrimination," *Journal of Human Resources* 23: 279-295.

[72] Oaxaca, Ronald, 1973. "Male-Female Wage Differentials in Urban Labor Markets," *International Economic Review* 14(3): 693-709.

[73] Oaxaca, Ronald L., and Michael R. Ransom, 1994. "On Discrimination and the Decomposition of Wage Differentials," *Journal of Econometrics* 61: 5-21.

[74] Pauly, Mark V., 1968. "The economics of moral hazard: comment," *American Economic Review*, **58**(3):531-536.

[75] Pauly, Mark V., 1986. "Taxation, Health Insurance, and Market Failure in the Medical Economy," *Journal of Economic Literature*, **24**(2):629-675.

[76] Reimers, Cordelia, 1983. "Labor Market Discrimination Against Hispanic and Black Men," *Review of Economics and Statistics* 65: 570-579.

[77] Robins, James A., Andrea Rotnitzky, and Lue Ping Zhao, 1995. "Analysis of Semiparameteric Regression Models for Repeated Outcomes in the Presence of Missing Data," *Journal of the American Statistical Association* 90(429), 106-121.

[78] Rothschild, Michael and Joseph E. Stiglitz, 1976. "Equilibrium in competitive insurance markets: an essay on the economics of imperfect information," *Quarterly Journal of Economics*, **90**(4):347-369.

[79] Royalty, Anne Beeson and Neil Solomon, 1999. "Health Plan Choice: Price Elasticities in a Managed Competition Setting," *Journal of Human Resources* **34**(1):1-41.

[80] Rutledge, Matthew S., and Catherine G. McLaughlin, 2008. "Hispanics and Health Insurance Coverage: The Rising Disparity," *Medical Care* 46(10), 1086-1092.

[81] Scanlon, Dennis P., Michael Chernew, and Judith R. Lave, 1997. "Consumer Health Plan Choice: Current Knowledge and Future Directions," *Annual Review of Public Health* **18**: 507-528.

[82] Schoenbaum, Michael, Mark Spranca, Marc Elliott, Jay Bhattacharya, and Pamela Farley Short, 2001. "Health Plan Choice and Information about Out-of-Pocket Costs: An Experimental Analysis," *Inquiry* **38**:35-48.

[83] Schur, Claudia, and Jacob Feldman, 2001. "Running in Place: How Job Characteristics, Immigrant Status, and Family Structure Keep Hispanics Uninsured," Commonwealth Fund Working Paper 453.

[84] Shah, N. Sarita, and Olveen Carrasquillo, 2006. "Twelve-Year Trends in Health Insurance Coverage Among Latinos, By Subgroup and Immigration Status," *Health Affairs* 25(6): 1612-1619.

[85] Short, Pamela Farley and Amy K. Taylor, 1989. "Premiums, Benefits, and Employee Choice of Health Insurance Options," *Journal of Health Economics* **8**(3):293-311.

[86] Strombom, Bruce A., Thomas C. Buchmueller, and Paul J. Feldstein, 2002. "Switching Costs, Price Sensitivity, and Health Plan Choice," *Journal of Health Economics* **21**:89-116.

[87] Taylor, Amy K., Sharon Larsen, and Rosaly Correa-de-Araujo, 2006. "Women's Health Care Utilization and Expenditures," *Women's Health Issues* 16(2): 66-79.

[88] Van Hook, Jennifer, Susan L. Brown, and Maxwell Ndigume Kwenda, 2004. "A Decomposition of Trends in Poverty Among Children of Immigrants," *Demography* 41(4): 649-670.

[89] Waidmann, Timothy A., Bowen Garrett, and Jack Hadley, 2004. "Explaining Differences in Employer Sponsored Insurance Coverage by Race, Ethnicity, and Immigrant Status," Economic Research Initiative on the Uninsured Working Paper 42.

[90] Weinick, Robin M., Elizabeth A. Jacobs, Lisa Cacari Stone, Alexander N. Ortega, and Helen Burstin, 2004. "Hispanic Healthcare Disparities: Challenging the Myth of a Monolithic Hispanic Population," *Medical Care* 42(4): 313-320.

[91] Weinick, Robin M., Samuel H. Zuvekas, and Joel W. Cohen, 2000. "Racial and Ethnic Differences in Access to and Use of Health Care Services, 1977 to 1996," *Medical Care Research and Review* 57 (Supp. 1): 36-54.

[92] Wooldridge, Jeffrey M., 2000. "Inverse Probability Weighted M-Estimators for Sample Selection, Attrition, and Stratification," mimeo, Michigan State University Department of Economics.