Title
Essays on Health Insurance, Household Liquidity, and the Demand for Medical Care

Author
Niedzwiecki, Matthew John

Publication Date
2013

UNIVERSITY OF CALIFORNIA, SAN DIEGO

Essays on Health Insurance, Household Liquidity, and the Demand for Medical Care

A dissertation submitted in partial satisfaction of the
requirements for the degree
Doctor of Philosophy

in

Economics

by

Matthew John Niedzwiecki

Committee in charge:

Professor Julie Cullen, Chair
Professor Gordon Dahl
Professor Melissa Famulari
Professor Joshua Graff Zivin
Professor Krislert Samphantharak

2013

The dissertation of Matthew John Niedzwiecki is approved, and it is acceptable in quality and form for publication on microfilm and electronically:

Chair

University of California, San Diego

2013

DEDICATION

Thanks to all the family and friends who supported me through the many ups and downs of graduate school. Special thanks to my partner, Coral Waters, for always pushing me to reach my potential.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

VITA

2004	B.A. in Economics Economics and Mathematics *magna cum laude*, Brandeis University
2008	M.A. in Economics, University of California, San Diego
2013	Ph.D. in Economics, University of California, San Diego

ABSTRACT OF THE DISSERTATION

Essays on Health Insurance, Household Liquidity, and the Demand for Medical Care

by

Matthew John Niedzwiecki

Doctor of Philosophy in Economics

University of California, San Diego, 2013

Professor Julie Cullen, Chair

My dissertation uses empirical methods to investigate the determinants of medical care demand and to understand how government action affects population health. The first chapter looks at a recent health insurance expansion in the state of Massachusetts and finds that, in the short run, the newly insured seek more hospital care. In the long run, as the supply of primary care physicians expands, care may shift to more efficient points of services. The second and third chapters examine the effect of cash liquidity on the demand for medical care and health insurance. I find liquidity to play an important role

in determining the timing of health investments and that the uninsured are more sensitive to liquidity constraints, likely because they face higher prices. Cash-on hand is also shown to be an important fact in maintaining continuous private health insurance coverage.

Chapter 1

The Effect of Health Insurance Coverage on the Demand for Hospital Care: Evidence from a Natural Experiment in Massachusetts

1.1 Introduction

Despite the vast amount of research on the uninsured, we still lack good evidence of the causal effect of having insurance on medical care utilization and health among adults, a group which makes up an overwhelming majority of those currently uninsured in the United States. This paper seeks to estimate the causal effect of health insurance coverage on the utilization of hospital care, through both emergency department (ED) and inpatient services, using variation in insurance coverage induced by the 2006 Massachusetts health insurance reform. I also investigate certain categories of hospital services I predict will be most responsive to preventive care and management of chronic health conditions, in an attempt to shed light on the effects of insurance coverage on underlying patient health. More precisely, I will estimate local average treatment effect (LATE) for those induced to take-up insurance because of the subsidies and individual mandate, a group of great interest given the pending implementation of the Patient Protection and Affordable Care Act (PPACA) in the coming years. The PPACA is focused on expanding coverage using methods very similar to those used in Massachusetts, which makes understanding the effects of health insurance coverage on medical care utilization and health crucial for predicting the effects of recent national reform and guiding future health care policy.

The net effect of health insurance coverage on hospital use is ambiguous, because it depends on the balance of at least two competing effects. First, insurance coverage reduces the own-price of hospital care which increases demand. At the same time, insurance reduces the price of a wide variety of health care services including physicians office visits and pharmaceuticals, which may function as substitutes for hospital care. The cross-price effect would reduce the demand for hospital care and potentially result in reduced medical spending if hospital spending decreases more than spending on substitutes increases.

In addition to the instantaneous price changes, other effects of insurance coverage may develop over time. First, increased consumption of health care services may improve underlying patient health through better management of chronic conditions, thereby reducing demand for hospital services. It may take some time for health to improve, which makes it important to examine effects over many years. Second, for those individuals who gain insurance coverage, it may take some time for them to learn about the benefits and costs, monetary and non-monetary, for services provided by hospitals. Over time, the newly insured may also have more medical conditions diagnosed and subsequently treated, or they may discover new methods of treatment for existing conditions. If this is true, we may see utilization trend up or down for some time following the beginning of insurance coverage. Given the potential for many competing forces of different signs and unknown magnitudes, theory leaves us uncertain as to how utilization of hospital services will change following policy reforms to expand coverage, and so we must turn to the data.

Our paper looks explicitly at the extensive margin – insured or uninsured – using a recent policy change as quasi-random variation among a population representative of those currently uninsured in the United States. I split the state into 312 cells based on gender, age group (19-30, 31-50, and 51-64), and geography (52 Public Use Microdata Areas) and compare "treatment" cells, which see the largest increase in average insurance coverage, to "control" cells, which see little to no change in insurance coverage.[1] The youngest age group, 19-30 year olds, and those earning less than 300% of the federal poverty line (FPL) see large increases in coverage, while the oldest age group, 51-64 year olds, and those earning more than 300% FPL see little to no increase in coverage. I use the Current Population Survey March Supplement (CPS) to estimate the effect of the Massachusetts reform on individual coverage. Then,

[1]Those 65+ would appear to be a good "control" group since they are always Medicare eligible and do not see a change in insurance coverage during the time period studied. However, the introduction of Medicare Part D, which provides prescription drug coverage, occurs in 2006, right around the beginning of the MA reform, making it difficult to distinguish the effects of the near simultaneous reforms.

I predict individual-level health insurance coverage in the American Community Survey, which allows for geographic variation, to create a measure of average insurance coverage for each cell before and after the reform. I then match these cells to hospital discharge records available in the Massachusetts Case-Mix Database to examine how changes in health insurance induced by the reform, affect changes in the demand for hospital care. Given the panel structure of the data, I am able to include cell-level fixed effects to control for time invariant unobservables that affect pre-existing levels of utilization within the cells, effectively comparing the change utilization rates for the "treatment" and "control" cells pre-reform and post-reform. I also control for state-wide trends in utilization by including dummies for each quarter of each year. Given that the policy was fully implemented by July 2007, I have several years of post-reform data to look at how the results evolve over time and do so by allowing insurance coverage to have a different effect in the latter half of the post-reform period in the data.[2]

Ours is not the first paper to look at the Massachusetts reform as a source of variation. Our use of within-state variation contrasts the use of cross-state variation in Kolstad and Kowalski (2012). They show that inpatient admissions which originate in the ED decrease in Massachusetts relative to control states, but I am able to show that this pattern is likely due to a state-wide trend because both my "treatment" and "control" cells experience the same temporary decline, which begins well before the legislation for the health reform was passed. Using within-state variation has the advantage that it is better able to control for state-wide trends that affect all groups equally. It has the disadvantage that I must rely on the assumption that my "control" group and my "treatment" group, which have different demographics, would otherwise be trending in the same way. In cross-state identification, the assumption is that other states – which likely match better on demographics, but

[2]The first half of the post-reform period is taken to be July 2007 to December 2008 and the second half is January 2009 to June 2010.

have different hospitals and state-wide policies – provide a good counterfactual to the Massachusetts experience.

Miller (2012) uses both across-state and within-state variation for identification, but she does so at the county level, leaving fewer observations.[3] Using within-state variation at the county level, regressions she finds a large and statistically significant decrease in ED use. When I attempt to replicate the regressions using the same data sources, I find a much smaller effect (about 25% as large) with much larger standard errors. In the replication, I show that clustering the standard errors dramatically reduces their size, as compared to White standard errors, which indicates potential small sample bias due to the fact that there are only 14 clusters. The estimated coefficients are not statistically significant from zero at the 10% level even with the much smaller clustered standard errors.

Our paper contributes to the literature in two important ways. First, it uses the Massachusetts Case Mix Database to examine the causal effect of insurance coverage on several different types of hospital care utilization in a two-sample IV framework, exploiting within-state variation in insurance coverage changes across a large number of cells induced by the reform. I find significant increases of about 38.6% in emergency department use, even for the subset emergency room visits during regular doctors' office hours, Monday through Friday, 8:00am-5:00pm. ED visits during regular doctors' hours increase by 50%, indicating that, at least in the short run, if there is a shift away from emergency room use towards primary care, it is not enough to overcome the lower price insurance coverage affords for ED visits. The subset of inpatient admissions originating in the emergency department are not found to change, but inpatient admissions overall increase by 50%. Potentially avoidable hospitalizations are found to decrease initially, but the estimate is somewhat imprecise. Second, I am able to look at the evolution of utilization

[3] There are only 14 counties in Massachusetts, compared to 52 Public Use Microdata Areas. She also does not exploit variation in the change in insurance coverage across different demographic groups.

over a three year time horizon to investigate the effects of health insurance past the very short run. Allowing the effect of insurance coverage to have a different effect for the last 18 months of data, I find that ED use does not decline, and if anything appears to be increasing.[4] The increase in utilization of inpatient services appears to be consistent over the three-year post period.

Section 2 provides a literature review on the causal effect of insurance coverage on the demand for health care, Section 3 describes the policy change and data, Section 4 outlines the identification strategy and regression framework, Section 5 describes the results, Section 6 provides discussion, and Section 7 concludes.

1.2 Literature Review

The RAND Health Insurance Experiment (HIE) (Manning et al. (1987)) randomly assigned the coinsurance rate (the fraction of the hospital bill paid by the patient), the deductible, and the annual out-of-pocket maximum payment for patients to investigate the effects of price on the demand for medical care and on health. They find a price elasticity of demand for nearly all services equal to about -0.2, proving that individuals are price sensitive even for something as crucial as medical care. While the HIE can tell us about how changes in cost-sharing (the intensive margin) affect demand, all participants in the experiment were covered by health insurance. The RAND HIE therefore cannot speak to the effects of changes along the extensive margin – going from uninsured to insured. Furthermore, the experiment was conducted from 1974-1982, before the passage of Emergency Medical Treatment and Active Labor Act (EMTALA) in 1986, which required hospitals to stabilize patients

[4]The final 18 months corresponds to January 2009 – June 2010. While discharge data is available through September 2010, I drop the final 3 months of data in order to account for discharges that occur up to 3 months after admission. Admissions which occur near the end of the dataset may not appear in the data because the patient has been in the hospital for longer than three months.

with emergency conditions, regardless of their ability to pay. Over the last thirty years, medical practice and technology have changed significantly.

Recent empirical work that has estimated causal effects of insurance coverage on utilization has mainly used policy variation in Medicaid or SCHIP eligibility or age discontinuities in eligibility for Medicare to identify causal effects of health insurance coverage on utilization and health. Currie and Gruber (1996a) and Currie and Gruber (1996b) found that Medicaid expansions in the 1980's and 1990's increased utilization of basic health care services among newly eligible children. They also found a decrease in child mortality and improvements in birth outcomes, particularly among the lowest income of the newly insured. Dafny and Gruber (2005) find that hospital use increases among children who become eligible for Medicaid, but that hospitalizations for avoidable conditions do not increase as much as those for unavoidable conditions. Card et al. (2008) use the discontinuity of Medicare eligibility at age 65 to identify the causal effect of gaining insurance coverage on utilization of medical care. They find large increases in inpatient and outpatient care at age 65 among all groups, especially minorities and those of low education. Those with low levels of coverage before age 65 saw a marked increase in low-cost medical services like doctors visits, while those with higher levels of coverage before age 65 saw a large increase in high-cost services like elective surgeries. It unclear, however, to what extent the population affected by previous Medicaid and Medicare reforms, which included very low-income children and their families and those sixty-five and older, respectively, is representative of those who are currently uninsured.

More recently, Finkelstein et al. (2011) use a 2008 Medicaid expansion in Oregon which, due to over-subscription, resulted in a lottery drawing for Medicaid coverage. In a well-designed experiment, the first large-scale, truly random assignment of health insurance since RAND, the authors find health insurance coverage leads to an increase in primary care utilization and hospitalizations, but lack statistical power to examine emergency room use.

Those who gained insurance coverage also reported better health and lower stress. While the source of variation in health insurance coverage is sound, the population in the study may not be entirely representative of those currently uninsured. The Oregon study looks at those uninsured who signed up for the lottery or were enrolled by someone else, a group who potentially have more frequent contact with medical care providers. Another important different is that the persons eligible for Medicaid in Oregon is poorer than a large number of uninsured. If income and assets affect the price uninsured individuals ultimately pay at the emergency room through bankruptcy law, as Mahoney (2011) shows, the Medicaid eligible uninsured persons may use the medical care system differently than those with higher incomes and more assets who have more to lose.

Anderson et al. (2012) exploit a common insurance rule which allows for the coverage of dependents under nineteen who are not enrolled in school full-time by a parent's insurance plan, creating a discontinuity in coverage at their nineteenth birthday. The authors find that losing insurance coverage is associated with a 40% decline in the use of emergency room services and a 60% decline in the use of inpatient services. While their estimates are well-identified, precise, and internally valid, they represent the local average treatment effect for nineteen year olds who are not full-time students, which raises potential concerns about how broadly this result applies to the rest of the adult population, aged 19-64, who were uninsured[5].

Long and Stockley, eds (2010) present difference-in-differences estimates for the effect of health care reform in New York and Massachusetts using data from the National Health Interview Survey (NHIS). While the NHIS is not designed for state level estimates and has only a small sample of Massachusetts residents, they find an increase in health insurance coverage of 5.0% among low income adults and a 2.7% increase in insurance coverage

[5]According to Garcia et al. (2010) 21.5% of 19-64 year olds were uninsured at any given time in 2009.

for all adults. Regarding utilization, the authors find a statistically significant decrease in delay of necessary care and an increase in visits to nurse practitioners and physicians assistants. They also see an increase of 7.4% in the number of individuals who have used the ED, but this estimate is not statistically significant, which may be a result of a small sample size. The NHIS also does not allow for estimates of the total number of ED visits.

Kolstad and Kowalski (2012) use the policy change in Massachusetts and a cross-state difference-in-differences identification strategy to examine the effect on insurance coverage and the utilization of inpatient care. They find that the policy significantly increased insurance coverage among those in the general population and among those hospitalized, especially among younger adults and those living in low income zip codes. Insurance coverage is found to significantly reduce the number of hospitalizations which originating with an emergency department visit, but the data do not allow for examination of the vast majority of emergency department visits which are discharged through the ED and do not result in an inpatient admission. They also look at health effects by examining hospitalization categories that are deemed potentially avoidable (hereafter PAH) by the Agency for Healthcare Research and Quality's (ARHQ) Prevention Quality Indicators (PQI) and find many categories of PAH decline while one increase. The data they use looks at 18 months following the full implementation of the Massachusetts insurance reforms, July 2007 through December 2008.

Miller (2012) uses the Massachusetts Case-Mix Database to examine the effects of health insurance coverage on the volume and mix of ED visits in Massachusetts. She uses both within-state and across-state difference-in-differences estimation strategy to estimate the change in utilization before and after the Massachusetts policy reform and finds a large statistically significant decline in emergency room utilization of 8.3-13% per capita for the entire state population. When scaled by the fraction who actually gained coverage, the estimated reduction in use is 0.85 to 1.28 visits per newly insured per year, a level exceeding the pre-reform utilization level of 0.44 visits per person per year for the uninsured. The large reduction is attributed to decreased use by the previously insured who could have upgraded to more generous coverage or other spill-over effects. The within-state design compares the fourteen counties within the state, using pre-reform measures of health insurance coverage from the Small Area Health Insurance Estimates (SAHIE) as a measure of treatment intensity. Counties with the lowest levels of coverage pre-reform are those which gain the most coverage after the reform.

1.3 Massachusetts Health Care Reform

1.3.1 Massachusetts Health Care Reform

Chapter 58 of the Acts of 2006, An Act Providing Access to Affordable, Quality, Accountable Health Care, was signed into law in April of 2006 and rolled out over the next fifteen months. The bill, which included an individual mandate for health insurance coverage, an employer mandate, expansion of state Medicaid programs, subsidies for private health insurance, and insurance market reforms, was designed with the goal of increasing insurance coverage in the state to near universal levels.[6]

The individual mandate, the first of its kind in the United States, requires that all persons above 150% of the Federal Poverty Line must provide proof of adequate health insurance or face a tax penalty.[7] Insurance plans deemed adequate for compliance with the individual mandate are required to meet certain criteria including a maximum deductible and coverage for services like preventative care visits and prescription drugs. The Massachusetts

[6]Doonan and Tull (2010) provide a more detailed explanation of the policy than the brief summary provided here.

[7]In the first year, 2007, the penalty was $219, equal to the loss of the personal exemption on state income taxes. In 2008, penalties were increased to one-half of the price of the least expensive insurance plan available to the individual, which depends on the individual's income, but did not exceed $912 for any individual.

Connector Board was placed in charge of determining "minimum creditable coverage" (MCC). MCC means coverage of primary care, preventative care, emergency room care, hospital inpatient care, ambulatory care, mental health care, and prescription drug coverage. Deductibles were capped at $2,000 for an individual ($4,000 for families) and $250 for prescription drugs ($500 for families) (Doonan and Tull (2010)). Exceptions for religious reasons or extreme financial hardship were granted in some circumstances.[8]

The state also increased coverage through public insurance expansion and subsidies for those purchasing on the individual market. Public insurance, the State Children's Health Insurance Plan (SCHIP) and MassHealth, the state Medicaid program, was extended to all residents earning below 150% FPL. SCHIP extended coverage to all children under age nineteen whose parents earned less than 300% FPL. MassHealth also increased the rates at which doctors and hospitals were paid for services, which Hahn (2012) has shown increases take-up of public insurance. To increase private coverage, the state also extended a large number of new insurance subsidies to residents who earned less than 300% FPL. For those earning less than 150% FPL, premiums were $0 with low cost-sharing.

In addition to the individual mandate and public insurance expansion, firms operating in Massachusetts with more than 10 employees were required to offer health insurance for all or pay a fine of $295 per employee. To meet the mandate requirements, firms were required to set up Section 125 plans, which allowed employees to purchase health insurance plans on the individual market with pre-tax dollars, to contribute at least 33% towards each worker's premium cost, and to reach a minimum insurance take-up rate of 25% among all workers. To further mitigate crowd-out due to public insurance expansion and subsidies for non-group insurance, the state government provided subsidies employer sponsored health insurance through the Insurance Partnership Program to

[8]Financial hardships include high mortgage payments, but no specific guidelines are given.

levels equal to the amount the government would subsidize the worker had he or she purchased coverage on the individual market.

The Massachusetts state government also added new regulations for health insurance companies operating in Massachusetts. The individual and small group markets were merged, reducing the average cost for individuals significantly and raising the cost for small groups. The state also established the Commonwealth Connector, a single online market place where all individual and small group plans can be easily compared and purchased.

As a result, from the 2006 to 2008 CPS data shows that health insurance coverage for adults nineteen to sixty-four years old in Massachusetts rose from 87.1% to 92.8%, with low-income and younger persons seeing larger increases in insurance coverage. The low income adult population not previously eligible for public insurance saw insurance coverage increase by 13.43 percentage points.[9]

1.4 Data, Identification and Empirical Framework

To estimate the causal effect of health insurance coverage on hospital utilization by adults, nineteen to sixty-four years old, the ideal data set would allow us to run the following regression:

$$Y_{it} = X_{it}'\beta + HI_{it}\theta + \alpha_i + \lambda_t + \eta_{it} \tag{1.1}$$

where Y_{it} represents the medical utilization outcome of interest, e.g. the number of visits to the emergency department, by individual i in time period t. X_{it} is a vector of demographic covariates, HI_{it} is an indicator function equal to one

[9]Long (2008) find an increase of insurance coverage from 87% to over 97%. overall in the state and estimate that about 5% of Massachusetts residents gained coverage in the first year alone. They also find the largest gains for low income groups. Between fall 2006 and fall 2007, individuals earning less than 300% FPL saw insurance coverage increase from 76% to 87%, while those earning less than 100% FPL saw coverage increase from 70% to 90%.

if individual i was covered by any health insurance during time t. Throughout this paper, "any health insurance" coverage refers to privately purchased health insurance coverage or any form of state or federal government health insurance, including Medicaid, Medicare, and any health insurance provided by the military, no matter how generous.[10] This indicator includes individuals who were only covered for part of the time period.[11] α_i captures all individual characteristics which are time-invariant, λ_t captures all time-varying effects which are constant across individuals within the state, and η_{it} represents idiosyncratic error.

Because I lack a single data set that contains all the required information, I must combine the following sources: (1) Current Population Survey March Supplement (CPS), (2) the American Community Survey (ACS), (3) the Massachusetts Department of Health Care Financing and Policy's Acute Care Hospital Case-Mix Database, and the (4) U.S. Census. The data is combined in four steps with more details to follow:

1. Use the CPS to predict individual-level health insurance coverage pre- and post-reform

2. Use the fitted values from the first regression to predict individual-level health insurance coverage, pre- and post-reform in the ACS. The ACS, unlike the CPS, identifies geographic location at the PUMA level, providing spatial variation

3. Collapse both the ACS data and the Case-Mix data to the cell level (PUMA, age group, gender)

[10]It would be interesting to separately identify the effects of private vs. public insurance, but due to data limitations, I am unable to do so. The CPS codes subsidized private coverage purchased through the health insurance exchanges as public insurance according to email correspondence. Furthermore, I am unable to identify insurance plan characteristics such as cost sharing provisions for various types of services.

[11]The CPS March Supplement data only indicate whether an individual was covered at any time during the past year but does not indicate which months he or she was or was not covered

4. Merge ACS and hospital data

I will estimate the effect of average health insurance coverage on average utilization for cells c identified based on zip code of residence, gender, and age, all of which are contained in the discharge summaries.

$$\overline{Y_{ct}} = \overline{X'_{ct}}\alpha + \overline{HI_{ct}}\theta + \mu_c + \lambda_t + \overline{\eta_{ct}} \tag{1.2}$$

Simply estimating this regression using cross-sectional variation in actual health insurance coverage is problematic due to the fact that coverage is endogenously determined in part by expected health care utilization, captured by the μ_j term, which becomes part of the error term $\overline{\eta_{ct}}$ without time variation. If there is a positive correlation between insurance coverage and unobserved factors which lead to increased utilization, e.g. sicker individuals are more likely to be insured, we would be over-estimating the causal effect of insurance coverage on utilization. It may also be the case, that if we allow for other forms of unobserved heterogeneity, such as risk preferences, those with the highest degree of risk aversion are also the ones with the lowest expected medical care utilization, which could lead to advantageous selection, a phenomenon documented in Finkelstein and McGarry (2006). Furthermore, it is not enough to use all time variation in insurance coverage, because some changes in insurance coverage may be correlated with health status. By isolating the variation in insurance coverage due to an exogenous government policy, we can be more confident we are estimating the causal effect of insurance.

In order to get an unbiased estimate of θ, I instrument for health insurance coverage using the 2006 Massachusetts policy change as plausibly exogenous variation in coverage among the previously uninsured.[12] The 2006

[12]Note that the causal estimate for θ here is a local average treatment effect (LATE) for those who gained coverage as a result of the 2006 reform. To the extent that this group is different from those who will gain coverage as a result of the Patient Protection and Affordable Care Act (PPACA), or that institutional details and regulations differ in Massachusetts as compared to other parts of the nation, this estimate may be a good or bad predictor of the effect of pending national legislation.

policy change in the state of Massachusetts greatly expanded insurance coverage among residents, but affected some groups more than others. In particular the poor, young, and previously ineligible for public insurance saw the largest gains.[13]

I instrument for health insurance coverage in the first stage using the following regression:

$$HI_{it} = X_{it}^{'}\gamma + D_{it}^{'} \times post_t\delta + post_t\psi + \varepsilon_{it} \tag{1.3}$$

Where $post_t$ is an indicator for the period after the reform was fully implemented and D is a subset of the covariates X_{it} that predict the change in insurance coverage, including poverty status, eligibility for pre-reform public coverage, age and gender. The interaction terms are excluded in the second stage. Therefore, the identifying assumption is that the only way poverty status, eligibility for insurance pre-reform, age, and gender would have effect on demand for medical care services *differently* before and after reform is through their impact on insurance status. The aspects of the reform which affect all individuals within the state are controlled for using time dummies for each quarter of data. Through the interactions, insurance coverage is predicted to change considerably for some groups, and not at all for others, creating a treatment group (those who see large gains in fraction insured) and a control group (those who see little to no gains in fraction insured) who exist contemporaneously within the same state, and who visit the same hospitals and clinics. Exploiting the differences in treatment intensity within Massachusetts and across groups living within the state provides a different source of variation than across states difference-in-differences and allows for a more precise way to control for common trends across all persons living within the state. However, the within-state variation cannot estimate the aspects of the reform

[13] In the data, "previously eligible for public insurance" is defined as (1) having a dependent under age nineteen living at home with income below 150% of the Federal Poverty Line or (2) eligibility for SSI, Social Security, or TANF payments.

which affect all groups equally.

One potential disadvantage of using within-state variation is that there may be spillover effects due to the increase in the number of insured, but I will argue that these are likely small. Groups which see large increases in coverage and groups with small or zero change in coverage face the same general equilibrium supply side responses because they attend the same hospitals. If these general equilibrium effects act on the treatment and control groups in the same way, my estimates will only reflect the effect caused by changes in insurance coverage. For example, if two cells exclusively visit the same hospital, and one cell sees a large increase in insurance coverage while the other does not, the increased crowding at the ED would affect both cells' decision to seek care. In order for the estimates to be biased, I would need one group to see crowding as a larger disincentive. Furthermore, the estimated increases in ED and inpatient utilization are likely to have small effects on overcrowding of EDs relative to the secular trend that has been increasing utilization per capita over recent years. In a back-of-the-envelope calculation, 5% of adult population, which makes up about 60% of the state population, gains insurance coverage, and each newly insured increased utilization by 40%, which would result in an increase of $(0.056) \cdot (0.40) \cdot (0.60) = 0.0134$, or a 1.34% increase in emergency room use above trend.[14]

I estimate insurance coverage using a three-step estimation strategy. In the first stage, I use the Current Population Survey (CPS) March Supplement, which includes individual level demographic and health insurance coverage data at an annual level to estimate a linear probability model for the change in insurance coverage before and after the reform. Then, using coefficient estimates from the CPS, in the second step I predict individual level health insurance coverage, pre-reform and post-reform, in the American Community Survey (ACS), which contains the same demographic variables used to predict

[14]The adult population is less likely than the elder (65+ years old) or children (0-18 years old) to use the emergency room, so the calculation of 1.34% is likely an over estimate of the increased crowding due to the reform.

health insurance coverage in the CPS and geographic variation at the level of the Public Use Micro-data Area (PUMA), but lacks data on health insurance coverage.[15].

In the final step, due to data limitations, I must aggregate hospital visits to the {PUMA}×{gender}×{age group} cell level and merge with average cell characteristics, including predicted fraction insured, population, and other demographics. There are 52 PUMAs, 3 age groups, 2 genders, and 84 months, resulting in 26,208 data points.[16] [17] For age groups, I divide the sample into three groups: (1) 19-30, (2) 31-50, and (3) 51-64 year olds. Using cell-level variables, we are left with the following regression:

$$\overline{Y_{ct}} = \overline{X'_{ct}}\alpha + \overline{\hat{HI}_{ct}}\theta + \mu_c + \lambda_t + \overline{\eta_{ct}} \tag{1.4}$$

where $\overline{Y_{ct}}$ represents utilization of service M per capita per month among members of cell j during time period t (time unit is one month). Note that all variables contained the vector of controls X_{ct} are binary, so, for example, we would not use "average income" as a regressor, but rather the "fraction below 300% FPL." In the main regression, I replace X_{it} with an average based on five years of ACS data.[18] Using the 5-year average 2005-2010 to represent the PUMA characteristics from 2001-2010 sacrifices time variation in X_{ct} in order to achieve more precise estimates of average cell characteristics. X_j is unidentified when cell fixed effects μ_j are added, resulting in the following

[15]The ACS first began to include questions about health insurance coverage in 2008, which unfortunately does not include the pre-reform period. OLS estimates are not available for this reason.

[16]Splitting cells by race was not possible due to the fact that categorization of race changed in October 2006 (beginning of fiscal year 2007) leading to a large number of patients not being classified for a significant part of FY 2007.

[17]All demographic data are binary at the individual level, so the cell level variables will represent the fraction who fall under a certain category, not an average as with continuous variables such as income

[18]Regressions were also run which predicted HI vary by year for each cell and point estimates were not statistically significantly different, but standard errors were much larger due to the introduction of noise in the estimate of the average annual subpopulation characteristic due to the random nature of each yearly sample.

regression:

$$\overline{Y_{ct}} = \overline{\hat{HI}_{ct}}\theta + \mu_c + \lambda_t + \overline{\eta_{ct}} \tag{1.5}$$

If average population characteristics are not changing significantly at the PUMA level over the time period, averaging over the period gives a more accurate estimate of the average subpopulation characteristics because of the larger sample. Looking at the means table for the first stage regression predicting change in insurance coverage based on observables, in Table (1.1), I see that average demographics within the state remain fairly constant over this time period, suggesting that using average characteristics over the study period may not be an unrealistic simplification. Unemployment data come from estimates provided by the state government of Massachusetts and vary by month. With time effects and cell fixed effects, identification comes entirely from time variation in predicted insurance coverage within each cell due to the reform.

Measures of monthly unemployment at the PUMA level are added to control for different trends in unemployment, especially in light of the macro-level shock of the financial crisis which occurs shortly after the health reform takes effect, which may affect different geographic locations in different ways. These controls are present in all specifications. In addition, the local unemployment rate is interacted with the age-gender subgroup identifier for each cell, allowing the local unemployment rate to have different effects for older vs. younger cells and men vs. women. The second set of results is not reported as the coefficients are virtually unchanged.

The period from April 2006 to June 2007, which corresponds to the implementation period, is omitted from the regressions as insurance coverage is likely changing rapidly and the data on coverage is annual, not monthly. Furthermore, there is higher potential for hospital reporting errors during this period of transition.[19]

[19] Several large gaps in reporting are avoided if this period is ignored.

1.5 Results

1.5.1 Change in Insurance Coverage

The policy increased insurance coverage by 5.64 percentage points among adults, especially for younger individuals, aged 19 to 30, and the poor who were not previously eligible for public insurance. There is also considerable geographic variation in predicted coverage changes, due to the reform, across PUMAs for different subgroups in Figure (1.1), (1.2), and (1.3). Furthermore, there is even greater variation across age groups, as young adults saw much larger gains to insurance coverage than did older adults. Figure (1.4) shows the large changes in the fraction of insured among males 19 to 30 years old, while Figure (1.5) shows little to no change in insurance coverage among females 51 to 64 years old. In Figure (1.6), I see a jump in insurance coverage beginning in March 2008, the first period following full implementation of the reform, which began with the enforcement of the individual mandate in July 2007.

The results of estimating equation (3.8) are displayed in Table (1.3).[20] In Table (1.3), column 4, the estimated policy effect on insurance coverage by low-income adults who were not previously eligible for public insurance, Social Security benefits, Supplemental Security Income benefits, or Medicare, is to increase coverage by about 10.5 percentage points.[21] Those who were previously eligible for public insurance, do not see an increase insurance coverage at the same rate as evidenced by the negative coefficient on "Post Reform × Eligible for Medicaid Pre-Reform." The coefficient is 60% as large as the ef-

[20] As a check on whether I am actually estimating the levels and variation in coverage correctly, I also use the measured HI status in the ACS, which is available for years 2008-2010, as a measure of post reform HI coverage. The levels and changes are close to what is estimated with the CPS data in the TSIV framework and coefficient estimates are essentially unchanged.

[21] An individual is considered eligible prior to the reform if his income places that at or below 150%FPL and he has at least one child under 18 at home. In addition to these requirements, there is an asset test, but that information is unavailable in the data, leading to an overestimate of the number of individuals who were previously eligible for Medicaid.

fect on those under 300% FPL, which could be a result of increased in take-up of public insurance, misclassification of Medicaid eligibility, or both. Young adults, aged 19 to 30 also see significant gains, likely reflecting their previously low levels of insurance coverage coupled with the individual mandate to purchase coverage. The partial F-statistic in specification 4 is 20.58, well above the threshold for weak instruments.[22] The effects of the policy on insurance coverage among subgroups within the state remains relatively constant over the entire post reform period, years 2008-2011 in the CPS justifying my choice to only measure insurance coverage as pre-reform and post-reform, not changing year by year.

1.5.2 Emergency Department Visits

For many uninsured individuals, the emergency department is their only source of hospital care (Newton and et al. (2008)). By acquiring insurance coverage, individuals gain access to other sources of care, such as care at a doctor's office, potentially reducing their demand for emergency department visits. However, due to moral hazard, it is not clear how the demand for ED visits should change; the own-price effect should increase demand, but the cross-price effect, which lowers the price of other medical services, could have a substitution effect.[23] Furthermore, if health improves because individuals are better managing chronic conditions, we may see a decrease in demand.

Looking column 1, the main specification, at Table (1.5), there is an increase of 23.54 (38.6%) emergency department visits per 100 persons per year resulting from increased insurance coverage.[24] While newly insured individuals

[22]In addition, since I am predicting individual coverage in the first stage, but then aggregating up to the cell level in the second stage, the predicted relationship should be stronger than suggested by the partial F-statistic.

[23]It is unclear how the uninsured view the price of an ED visit since the hospital price of medical care is often hidden and individuals rarely pay the full amount even when uninsured, but it is unlikely that the price of an ED visit will increase when the individual becomes insured.

[24]In all results, we look at the specification in column 1 of the corresponding table, which includes time effects and cell fixed effects. Robustness is examined in columns (2) and

may be increasing their use of office-based care for some ED care, on net, emergency room use is increasing within the time period studied here. Given that primary care physician supply is likely to be inelastic in the short run, the increase ED utilization may be a result of excess demand or the long waits, and could be expected to diminish as physician supply adjusts to the new equilibrium. If we allow insurance coverage to have a different effect on utilization for the last eighteen months, we see the effect on ED use persists, and even appears to increase. In the last 18 months, ED use increases by 28.94 visits per 100 persons per year over the baseline, an increase of 48.4%.

In Figure (1.9), I regress emergency room use per capita on all control variables, excluding health insurance, and plot the difference in mean residuals[25] between the 25% of cells which see the largest increase in fraction insured and the 25% of cells which see the smallest increase. A linear fit line is plotted through each set of residuals, and the slope and intercept are allowed to differ before and after the reform. One can see that the cells which saw the largest change in coverage typically had an average residual below those who had continuously high levels of coverage in the pre-reform period. Then, around the time of the reform, whose implementation period is designated by the two vertical dashed lines, the groups predicted to see the largest increases in insurance coverage increased their utilization and maintained high utilization in the post reform period.

(3), which includes both subgroup specific time trends and an an interaction between a linear time trend and the fraction of individuals insured prior to the reform. Columns (4)-(6) repeat the exercise, allowing the effects of insurance coverage on the newly insured to vary over time (first 18 months post reform are compared with months 19-36) to separate transitory from more permanent effects. Coefficient estimates are bootstrapped as detailed in the Appendix.

[25]Residuals come from equation (1.5) excluding health insurance.

1.5.3 Emergency Department, Monday-Friday 8:00am-5:00pm

I next restrict I analysis to those ED visits which occur between the hours of 8:00am and 5:00pm, common business hours for primary care physicians. Visits to the ED during these hours are more likely to substitute for primary care and could be more responsive to insurance coverage. ED visits during regular doctors hours include any discharge from the emergency department for an individual admitted during the hours 8:00am-5:00pm, Monday through Friday.[26] In column 1, insurance coverage increases the number of ED visits during regular doctors's hours by 11.2 visits per 100 persons per year (50.15%). Looking at the evolution of ED visits during regular doctors' hours, I find that weekday visits remain higher and even increase. In Figure (1.9), the difference in mean residuals for the cells with high and low insurance increases are plotted. The cells which see the largest increase in health insurance coverage see their utilization increase right around the time of the reform and for all periods that follow.

1.5.4 Inpatient Hospitalizations from ED Visits

Unlike visits to the emergency department, which require only the decision of the patient, the decision to admit someone to inpatient care from the ED requires action on the part of the hospital. By law, hospitals are required to stabilize all individuals in serious condition who enter the ED, regardless of insurance coverage, but no more. When the fraction of insured individuals in the population increases, hospitals have an incentive to increase the intensity of treatment because they are now better reimbursed for the services they provide, which may lead to more inpatient admissions from the ED. While it

[26]Data on the day of the week and time of admission are only available for those ED visits which do not result in outpatient observation or inpatient hospitalization making it impossible to categorize visits by time of arrival for emergency department visits which are transferred to outpatient observation or inpatient hospitalizations due to data limitations.

is true that insurance coverage does increase the total number of emergency department visits, it is unlikely that individuals would have forgone treatment for conditions serious enough to warrant inpatient admission, so demand for these types of visits should be highly inelastic. At the same time, it is possible that insurance coverage leads to improvements in underlying health, which could lower the number of severe negative health shocks which would lead to an ED visit.

I examine all inpatient admissions originating with an ED visit and find, in column 1 of Table (1.7), that insurance coverage increases the number of ED visits that result in hospitalization by 0.82 per 100 persons per year, but this estimate is not statistically significant at traditional confidence levels. Estimates of the effect on charges yield results that are sensitive to the time trend controls included, and I cannot reject the null hypothesis of no change when using the specification in column 3, which includes controls for subgroup specific time trends and a linear trend interacted with the pre-reform level of insurance coverage. Figure (1.9) plots the difference in mean residuals, but finds no visual break around the time of the reform.

Unlike other services, inpatient admissions from the ED exhibit a puzzling decline statewide, seen in Figure (1.8), beginning between late 2005 and late 2006, which breaks from an otherwise steadily increasing or flat trend. Inpatient admissions from the ED then begin to trend upward beginning in 2007. It is the only such outcome measure which exhibits this behavior. The fact that it does not appear to affect cells within the state differentially based on change in insurance coverage, as seen in Figure (1.8), leads to the possibility that there was some other change in the way Massachusetts hospitals admitted patients from the emergency room, which may be of concern for estimates relying on cross-state variation in insurance coverage. There may also be some concern with the reliability of this measure as it requires discharging hospitals to have information on the source of the visit, which may have occurred at another location if the patient was transferred.

1.5.5 Inpatient Hospitalizations, Overall

Hospital admission overall, like the specific subset of inpatient admissions from the ED, is affected by a combination of patient health, and patient and doctor incentives. As a large fraction of medical care spending, the effect of insurance coverage on this type of utilization has important ramifications for overall medical spending in the U.S.

In the data, inpatient admissions include all patients discharged from hospitals, regardless of entry point. Looking at total inpatient hospitalizations in column 1 of Table (1.8) I find that increased insurance coverage leads to an increase of 3 inpatient hospitalizations per 100 persons per year (49.8%). When I allow the for the effect of insurance coverage to change over time, I find that the increase in the number of hospitalizations diminishes slightly in the last eighteen months of data, but the difference in the coefficients is not statistically significant. In Figure (1.9), the difference in mean residuals plot shows an relative increase in utilization just after the reform was passed for those cells who saw the largest increases in insurance coverage. This effect persists for the post period.

1.5.6 Potentially Avoidable Inpatient Hospitalizations

Finally, I look at the effect of insurance coverage on the number of inpatient hospitalizations which are deemed by the AHRQ's Prevention Quality Indicators to be the result of insufficient preventative care and proper management of chronic conditions. Avoidable hospitalizations include inpatient admission for avoidable complications from conditions such as hypertension, asthma, and complications due to diabetes. It is unreasonable to think that all hospitalizations for such conditions could be avoided, but overall, this subset of inpatient admissions should be most likely to decline with increased use of medication and routine doctor visits.

In the data, PAHs are defined by the diagnosis and procedures performed as described in Agency for Healthcare Research and Quality (2007). In column 1 of Table (1.9) insurance coverage decreases the number of PQI hospitalizations by 1.26 per 100 persons per day. The magnitude of the estimated coefficient seems rather large, but it could reflect further heterogeneity among previously uninsured patients. Figure (1.9) shows the difference in mean residuals, documenting a relative decline in PAHs for those cells which saw the largest increase in insurance coverage. Given that PAH hospitalizations are much more common for older populations, because of the increased incidence of chronic conditions among this group, it is possible that for this outcome, I control group is trending differently than the treatment group in ways not captured by a subgroup specific linear time trend. If this were the case, the results would driven by an increase in PAH admissions by older persons and not a decline in PAH admissions by younger persons.

1.6 Discussion

I find different results from other recent work on the Massachusetts reform in Kolstad and Kowalski (2012) and Miller (2012). Kolstad and Kowalski (2012) find the same increases in insurance coverage, particularly among the young and poor, but find a decrease in hospitalizations from the ED. I find no effect but the small decrease they find is well within the 95% confidence interval of I estimates. Examining the trend in hospital admissions from the ED in the Case-Mix data through June 2010 in Figure (1.7), a break in the upward trend starts in 2005 for Massachusetts, well before the reform began in April 2006, but not for other hospitals in the Kolstad and Kowalski (2012) sample, suggesting some Massachusetts specific trends may influence results in a cross-state difference-in-differences design. Figure (1.8) indicates that such trends are present for all subgroups, including those 51-64 year olds who see very little change in insurance coverage. Using within-state controls may better control for such confounding influences as long as those cells which saw relatively high or low gains in insurance coverage did not patronize hospitals most affected by this unknown influence.

Miller (2012) finds that reform caused a large decrease in emergency room use, sharply contrasting with the results in this paper, which find a significant increase, despite the fact that both papers use the same outcome data and within-state variation, albeit at different geographic levels. In I attempt to replicate, in Table (1.10), her main regression (Table 2 in Miller (2012)), I find a negative effect of increased insurance coverage on ED use, but the magnitude of this coefficient is smaller and the standard errors are much larger. Furthermore, clustering seems to result in standard errors that are less conservative (smaller) than White standard errors, indicating possible small-sample bias[27]. In addition, the magnitude of the results in Miller (2012) are quite large. Miller finds a *decrease* of 0.16 ED visits per quarter newly insured person relative to baseline of 0.11, which amounts to a $5-8\%$ reduction in statewide ED visits, a strikingly large result that does not appear visually in the data, which appear to increase every year from 2002 to 2010.

Our estimates on the effect of insurance coverage on hospital use align rather closely with those found in Anderson et al. (2012), but in this paper I am able to examine how the effect changes over a longer time period.[28] While I examine the effects over a three year period, the health effects of preventive care and regular care for chronic conditions may not show up for quite some time, leading short run studies to underestimate these cost-saving potential offsets. Further study of the long-run effects of health insurance, beyond the still modest three year window examined here, is crucial to current health policy.

[27] I use the standard Stata command “vce(cluster).”

[28] In Anderson et al. (2012), the authors found decreases of 40% and 60% for ED and inpatient use, respectively, when insurance coverage is dropped. The same calculation performed here results in a -23.08/(59.77+23.08)=-28% change in ED visits (compare to their -40%); and a -3.025/(6.07+3.025)=-33% (compare to their -60%)

1.7 Conclusion

In the short run, expanding health insurance coverage will lead to an increase in the demand for the emergency department and inpatient care. Given that Massachusetts previously had high rates of insurance, this implies even larger effects on aggregate demand for health insurance in other states, where the predicted change in fraction insured would be larger. Our results provide some suggestive evidence that potentially avoidable hospitalizations may decline.

Given that the uninsured are a diverse group, including some who are young and healthy as well as those who are very sick, it would be interesting to identify the effects of insurance coverage separately for each group. Unfortunately, the empirical design used in this paper cannot separately identify these effects. One might expect the offset effects of preventative care and disease management, and the ensuing health effects provided, to be larger for those individuals with many chronic conditions than for the young and healthy. Future work is clearly needed.

Finally, this paper, along with Kolstad and Kowalski (2012) tries to relate utilization patterns to underlying population health. In both settings, potentially avoidable hospitalizations, as defined by the AHRQ are examined as an outcome that may potentially decline insurance coverage improves underlying health. There exists an abundance of information in hospital discharge data, but it is still unclear how this exactly maps to health. Further work, with contributions from the medical care community, will be important to identify ways to use the wealth of administrative data to better understand population health in policy evaluation.

1.8 Appendix

1.8.1 Data Set Descriptions

Current Population Survey, March Supplement (CPS)

The CPS March Supplement provides a measure insurance coverage within the state of Massachusetts over time (at the annual level) and for subgroups of interest, which is crucial because the identification strategy relies on variation in changes in insurance coverage for different groups, identified by age, gender, and geography. The six subgroups which I examine are (1) males 19 to 30 years old, (2) females 19 to 30, (3) males 31 to 50, (4) females 31 to 50, (5) males 51 to 64, and (6) females 51 to 64. I use ten years of CPS data to establish the heterogenous effect of the policy on the six different subgroups. All variables used in the first stage are binary indicator variables. Health insurance questions in the March CPS are retrospective to the previous 12 months and ask participants to state if they were ever covered by a given type of insurance during that time period. The insurance coverage indicator is equal to one if an individual was ever covered by any type of health insurance in the previous year, public or private. The post reform period corresponds to the 2008-2011 CPS as March 2008 is the first March Supplement survey to occur after the mandate took effect in July 2007. The sample includes adults, nineteen to sixty-four years of age, living in Massachusetts in years 2002 to 2011, for a total of 19,978 observations.

American Community Survey (ACS)

The ACS when combined over three or more years, is similar to the decennial census, and allows us to estimate population characteristics at small geographic levels such as the Public Use Microdata Area (PUMA), a consistent group of contiguous census blocks with a combined population of 100,000 or more. In Massachusetts, there are 52 PUMAs, with 6 subgroups in each, making for 312 units of observation, or cells. The ACS is crucial for getting spatial variation in changes to insurance cover as a result of the reform. The ACS did not begin measuring rates of insurance coverage until 2008, however, but it does include many of the same demographic information as the CPS, which allows us to predict a pre-reform and post-reform insurance coverage variable for each person in the ACS, and then aggregate these predictions to get the predicted number insured in each cell. Cell characteristics are determined using the average characteristics from 2005-2010 to reduce sampling noise from year to year. Data for individuals living in group quarters was not collected in 2005, so it is imputed using 2006 data.

Massachusetts Department of Health Care Financing and Policy's Acute Care Hospital Case-Mix Database

The Case-Mix Database is collected by the state and contains discharge summaries for all patients who admitted to the 80 acute care hospital locations in Massachusetts.[29] Healthcare facilities not included in this dataset are community health centers, urgent care centers, day treatment facilities, and Veterans Administration (VA) Hospitals. Each entry includes a summary of the visit including patient's zip code of residence, basic demographic information (gender, age, race, etc.), date of admission and discharge, diagnosis (ICD-9-CM), procedures performed, charges, and insurance status, among other details. Using the geographic and demographic identifiers, I am able to match each visit for a Massachusetts resident to one of the 312 cells. The data does not include any VA hospital admissions and does not include visits to hospitals by Massachusetts residents to hospitals outside of the state. Data for one hospital was missing for 3 months due to a conversion to a new computer system; missing observations were imputed according to a procedure outlined later in appendix.

[29] There are 71 organizations, but some have multiple campuses.

2000 & 2010 U.S. Census

This data set provides estimates of cell-level population. Population for intercensal years is interpolated using a linear trend.

Massachusetts Labor Force and Unemployment Data

The state government of Massachusetts provides monthly estimates of the number of employed and unemployed persons by city of residence.[30] These estimates are aggregated to the PUMA level and applied to all subgroups within the PUMA.

1.8.2 Imputation for Missing Hospital Values

Some hospitals are missing data for months in the sample. For missing months, visit counts, by each zip code, age and gender, are replaced with the mean utilization rate in the corresponding period, pre-reform or post-reform, for that particular month. Given that the estimation strategy is based on a pre-reform/post-reform change in insurance coverage, this will not bias the results if the reports are missing at random.

1.8.3 Bootstrap Procedure

Coefficient estimates are bootstrapped to account for sampling variation in the first stage, where the fraction insured in each cell is imputed. The steps are outlined below:

1. Take a random sample from (1) CPS [19,978 obs.] and (2) ACS datasets [236,282 obs.]

2. Run individual-level regression in CPS to get predicted HI pre- and post-reform. Note: no geographic data, outside of state of residence, is avail-

[30] The smallest four towns do not provide monthly estimates; annual estimates for these towns are used instead.

able in the CPS dataset:

$$HI_{it} = X_{it}^{'}\gamma + \left\{ D_{it}^{'} \times post_t \right\} \delta + post_t + \varepsilon_{it} \tag{1.6}$$

3. Predict HI at individual level in the ACS data. ACS includes geographic identifiers:

$$\widehat{HI}_{it} = X_{it}^{'}\widehat{\gamma} + \left\{ D_{it}^{'} \times post_t \right\} \widehat{\delta} + \widehat{post_t} \tag{1.7}$$

4. Aggregate $\widehat{HI}_{it}$ over $i \in s$, a geography-gender-agegroup identifier (e.g. males 19-30 years old, living in one of fifty-two PUMAs), and other RHS variables from the first stage, excluding post interaction terms used as instruments, to create geography-gender-agegroup cell level characteristics. Divide by cell population to create average cell characteristics.

$$\frac{1}{N_{ct}} \sum_{i \in j} \widehat{HI}_{it} = \overline{\widehat{HI}_{ct}} \tag{1.8}$$

$$\frac{1}{N_{ct}} \sum_{i \in j} X_{it} = \overline{X_{ct}} \tag{1.9}$$

5. Match average predicted HI and other characteristics to outcome data

6. Take random sample of {PUMA}×{gender}×{age group} cells, clustered at the PUMA level.

7. Regress utilization per capita on average HI and other characteristics, including cell and time fixed effects.[31]

$$\overline{Y_{ct}} = \overline{\widehat{HI}_{ct}} * \theta + \mu_c + \lambda_t + \eta_{ct} \tag{1.10}$$

[31] There is a small attenuation bias due to sampling variation of the cell characteristics which predict the change in insurance coverage because (X,Y) pairs, i.e. average cell characteristics and average cell utilization. The bootstrapped coefficient estimates align closely with the two-sample instrumental variables estimates

1.9 Figures

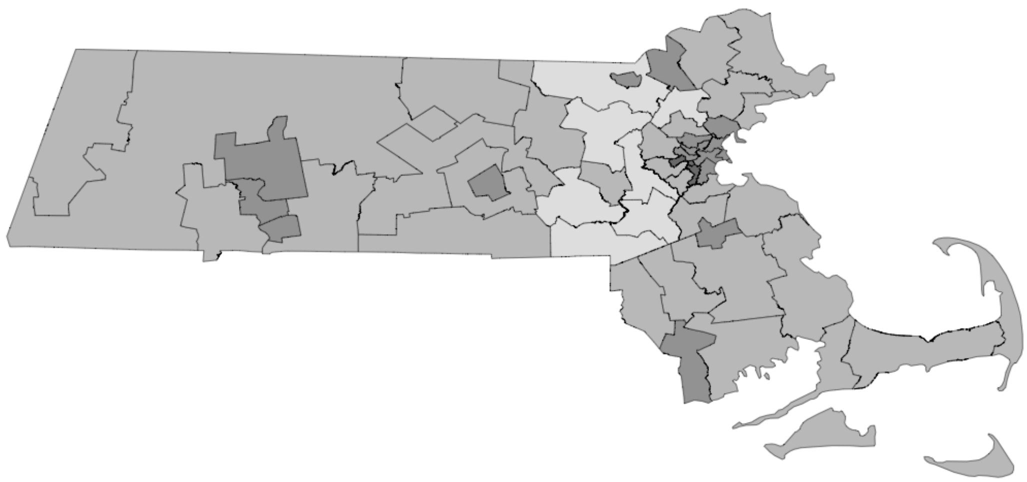

Figure 1.1: Non-Elderly Adult Health Insurance Coverage, Pre-Reform

Fraction of adults, 19 to 64, insured for each of 52 PUMAs in Massachusetts, prior to April 2006, the passage of the reform bill. Values range 0.77 to 0.92. Lighter colors represent higher average insurance coverage. Color scale represents values from 0.65 (black) to 1.0 (white).

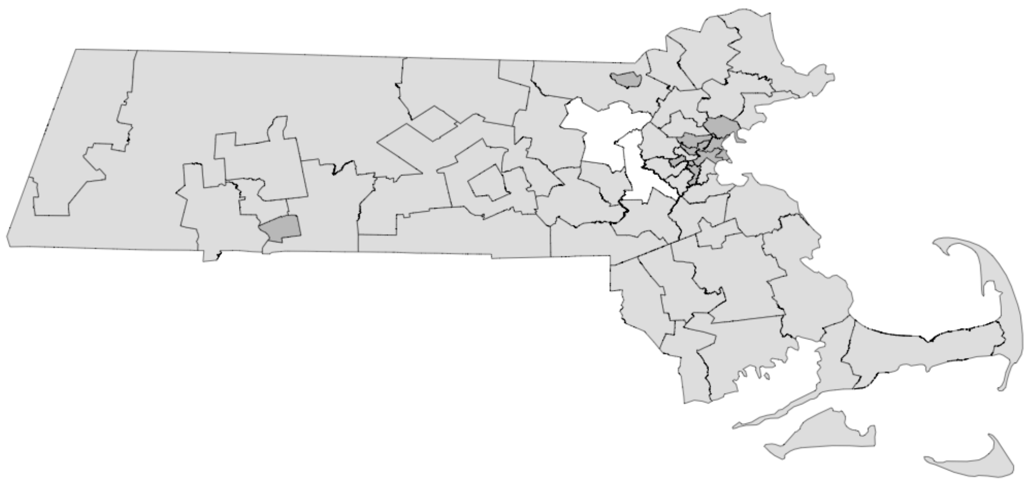

Figure 1.2: Non-Elderly Adult Health Insurance Coverage, Post-Reform

Fraction of adults, 19 to 64, insured for each of 52 PUMAs in Massachusetts, after the reform is fully implemented in July 2007. Values range 0.88 to 0.96. Lighter colors represent higher average insurance coverage. Color scale represents values from 0.65 (black) to 1.0 (white).

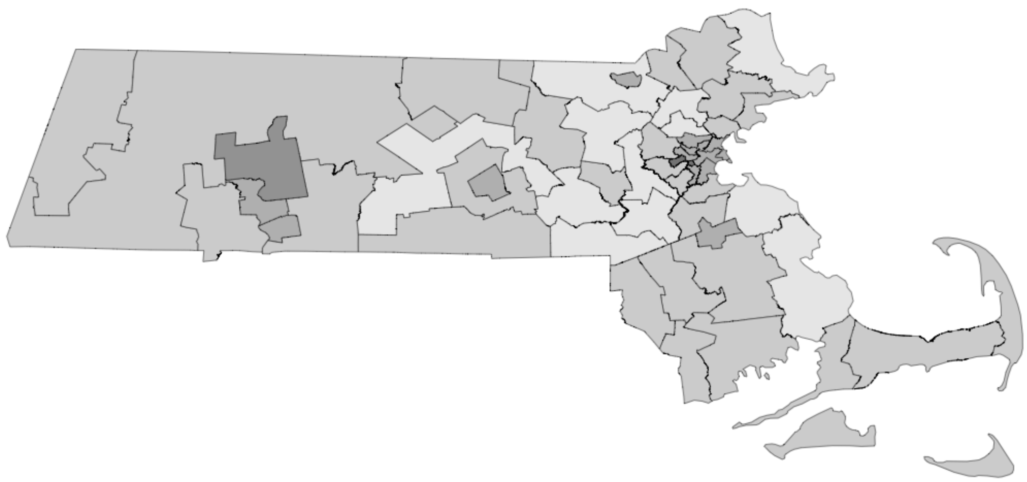

Figure 1.3: Change in Non-Elderly Adult Health Insurance Coverage

Change in the fraction of adults, 19 to 64, insured for each of 52 PUMAs in Massachusetts. Values range +0.03 to +0.12. Darker colors represent larger increases average insurance coverage. Color scale represents values from 0.0 (white) to 0.18 (black).

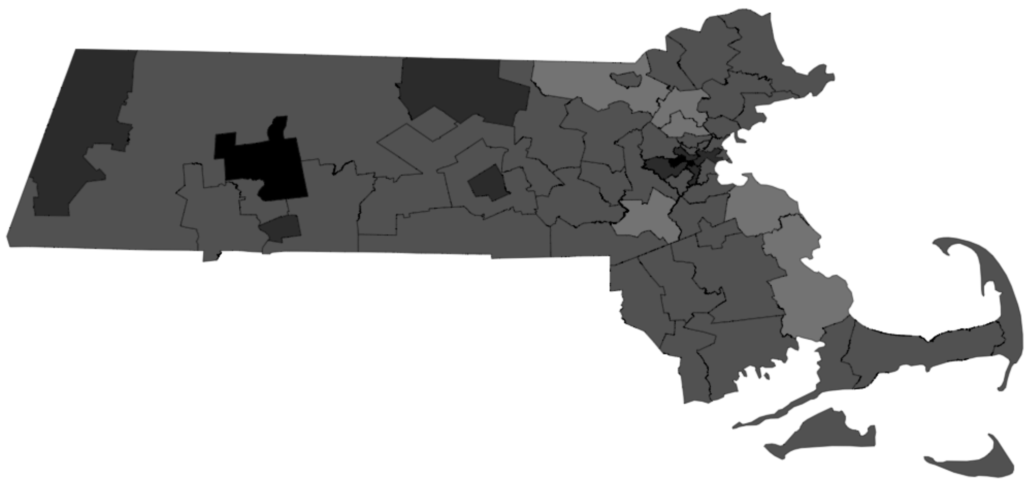

Figure 1.4: Change in Health Insurance Coverage, Males 19 to 30

Change in the fraction of males, 19 to 30, insured for each of 52 PUMAs in Massachusetts. Values range +0.12 to +0.18. Darker colors represent larger increases average insurance coverage. Color scale represents values from 0.0 (white) to 0.18 (black).

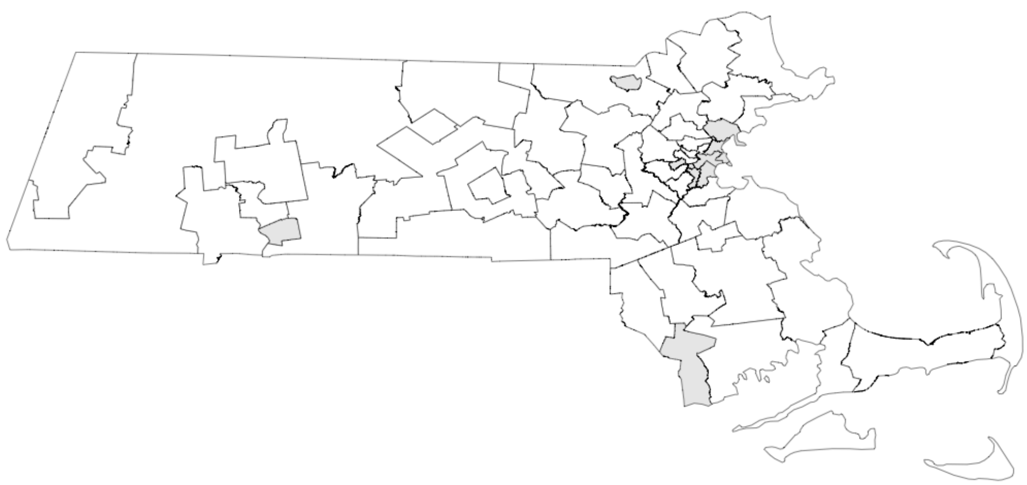

Figure 1.5: Change in Health Insurance Coverage, Females 51 to 64

Change in the fraction of females, 51 to 64, insured for each of 52 PUMAs in Massachusetts. Values range +0.01 to +0.03. Darker colors represent larger increases average insurance coverage. Color scale represents values from 0.0 (white) to 0.18 (black).

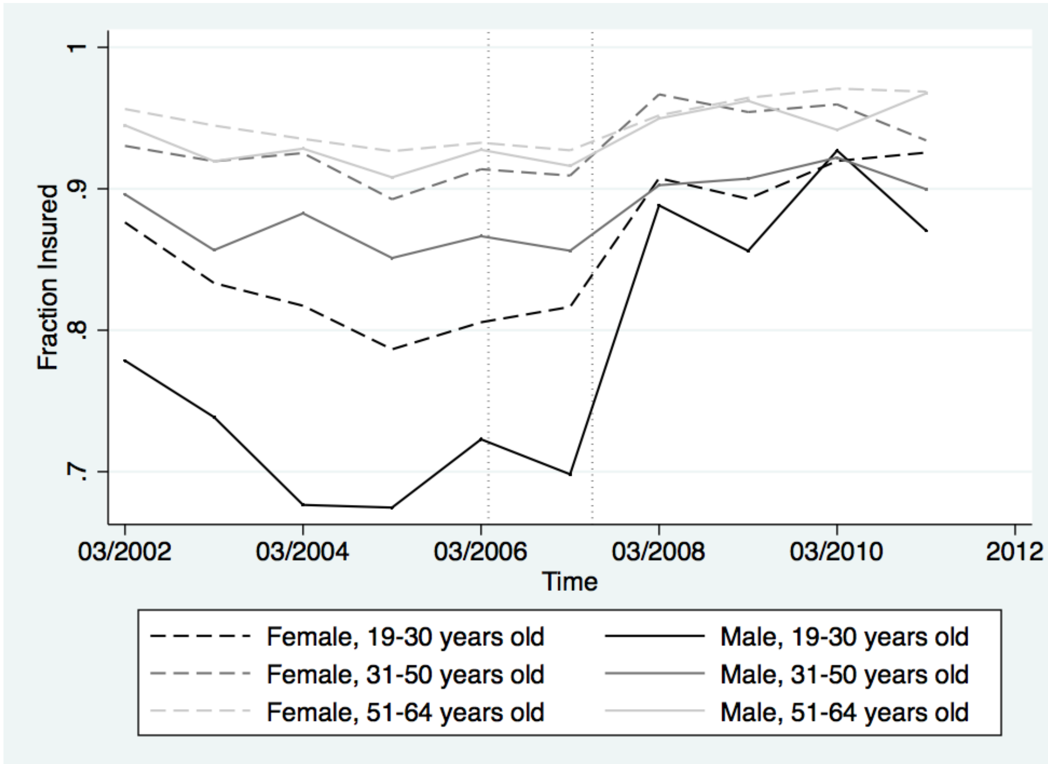

Figure 1.6: Adult Insurance Coverage in Massachusetts

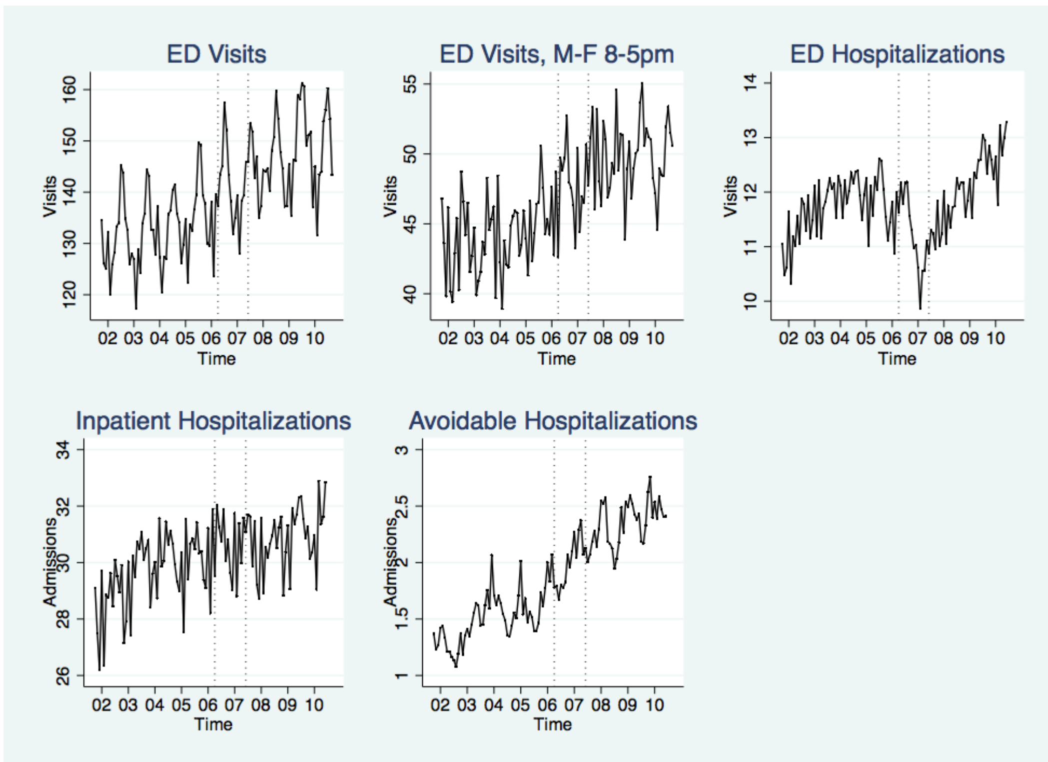

Figure 1.7: State-wide Adult Demand for Hospital Care, Visits per Month (1,000s)

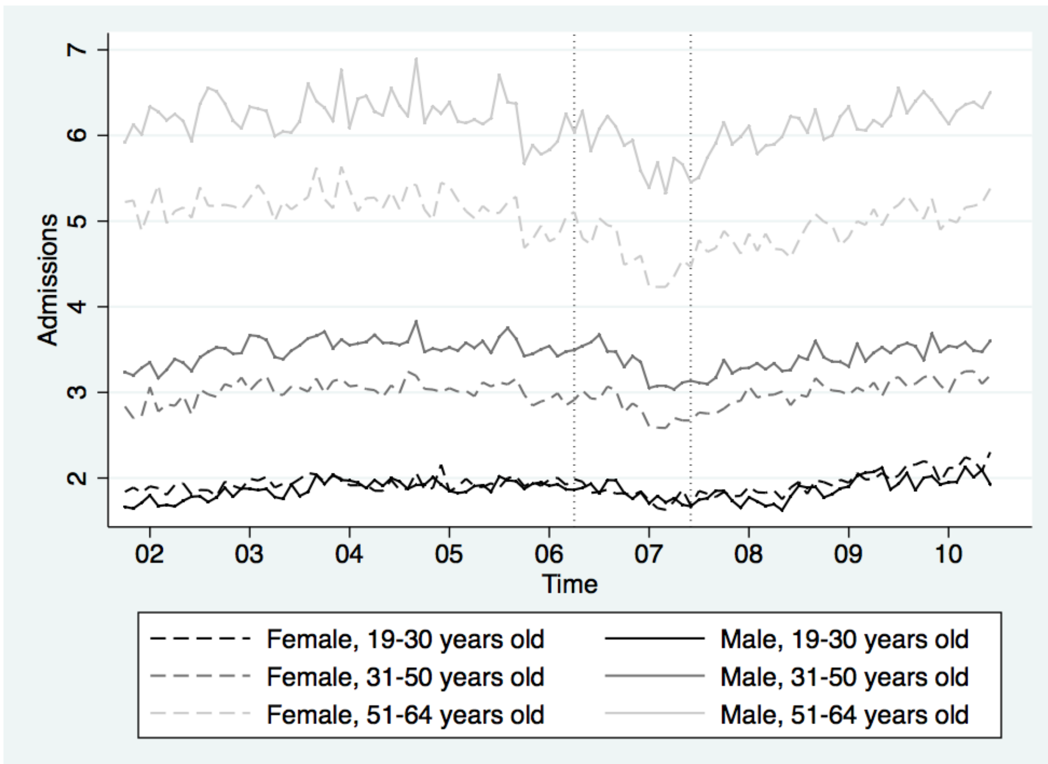

Figure 1.8: Inpatient Admissions through the ED, by Subgroup (Seasonal Trends Removed)

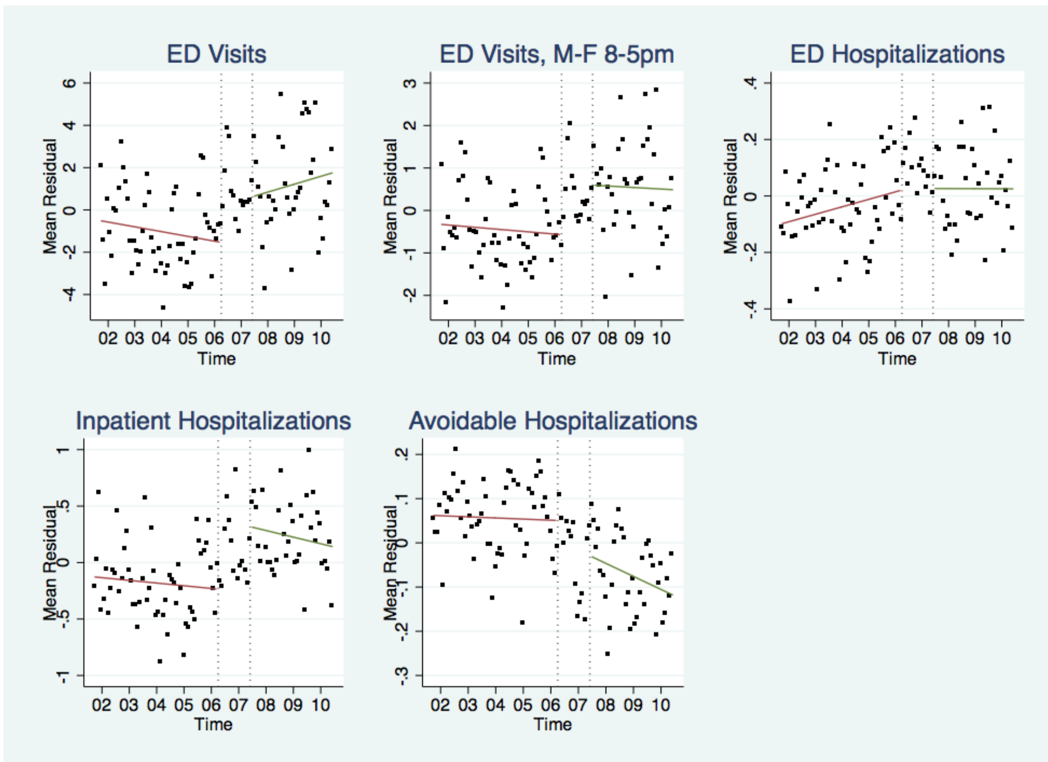

Figure 1.9: Difference In Mean Residuals (HI variable omitted), Top 25% HI Gains vs. Bottom 25% HI Gains

1.10 Tables

Table 1.1: Means Table, CPS March Supplement and ACS

	CPS Pre-Reform	CPS Post-Reform	ACS Pre-Reform	ACS Post-Reform
Covered by Health Insurance	0.870	0.930	-	-
	(0.335)	(0.253)		
< 300% FPL	0.336	0.342	0.327	0.348
	(0.472)	(0.474)	(0.469)	(0.476)
Eligible for Medicaid Pre-Reform	0.104	0.114	0.089	0.093
	(0.305)	(0.318)	(0.285)	(0.291)
Veteran	0.064	0.047	0.076	0.062
	(0.245)	(0.212)	(0.265)	(0.242)
Wage Worker	0.675	0.662	0.681	0.676
	(0.468)	(0.472)	(0.466)	(0.467)
Self-employed	0.074	0.065	0.073	0.067
	(0.262)	(0.248)	(0.260)	(0.250)
Unemployed	0.035	0.057	0.044	0.064
	(0.185)	(0.231)	(0.206)	(0.244)
Bachelor's Degree or Higher	0.388	0.417	0.374	0.386
	(0.487)	(0.493)	(0.484)	(0.486)
Some College	0.241	0.244	0.266	0.284
	(0.427)	(0.430)	(0.442)	(0.451)
HS Diploma or Equivalent	0.277	0.265	0.265	0.243
	(0.447)	(0.441)	(0.441)	(0.428)
Not a U.S. Citizen	0.102	0.087	0.103	0.098
	(0.302)	(0.283)	(0.304)	(0.298)
Has a Dependent Child, Under 19 Years Old	0.445	0.439	0.426	0.401
	(0.497)	(0.496)	(0.494)	(0.490)
N	12,698	7,280	77,395	119,763

Means table for binary variables used in first-stage regressions in two-sample instrumental variables regression. CPS Pre-Reform period refers to CPS March Supplement data from 2002-2007; CPS Post-Reform period refers to CPS March Supplement data from 2008-2011; ACS Pre-Reform period refers to ACS data from 2005-2006; and ACS Post-Reform Period refers to data from 2008-2010. No health insurance data appears in the ACS until 2008, so it is instrumented by demographics data and demographics interacted with an indicator variable for the post-reform period and not reported here.

Table 1.2: PUMA Subgroup Characteristics

	5th	25th	50th	75th	95th
Fraction Insured Pre-Reform	0.710	0.772	0.877	0.918	0.948
Fraction Below 300% FPL	0.150	0.220	0.311	0.426	0.601
Fraction Previously Eligible for Public Insurance	0.018	0.044	0.072	0.119	0.214
Fraction U.S. Military Veteran	0.004	0.013	0.024	0.098	0.323
Fraction Wage Worker	0.551	0.623	0.678	0.726	0.776
Fraction Self-Employed	0.009	0.027	0.056	0.097	0.154
Fraction Unemployed	0.024	0.034	0.045	0.061	0.105
Fraction with Bachelors Degree	0.152	0.236	0.351	0.461	0.682
Fraction with Some College	0.148	0.230	0.276	0.330	0.419
Fraction with HS Diploma	0.106	0.207	0.266	0.315	0.384
Fraction Not a U.S. Citizen	0.016	0.034	0.070	0.136	0.256
Fraction with a Child Under 19 Living at Home	0.059	0.208	0.345	0.510	0.697

5th/25th/50th/75th/95th percentile measures for PUMA subgroup (gender $\times$ age group $\times$ geography) characteristics. Characteristics are taken as fixed, using the ACS 5 year average from 2005-2010.

Table 1.3: First Stage: Predicted Health Insurance Coverage at Individual Level

	(1)	(2)	(3)	(4)
Post Reform	0.0564***	0.0293***	0.0293***	-0.00199
	(0.00451)	(0.00450)	(0.00450)	(0.00816)
Post Reform × < 300 FPL		0.0799***	0.105***	0.0926***
		(0.0108)	(0.0140)	(0.0141)
Post Reform × Eligible for Medicaid Pre-Reform			-0.0775***	-0.0596***
			(0.0183)	(0.0183)
Post Reform × 19 to 30 Years Old				0.0801***
				(0.0133)
Post Reform × 31 to 50 Years Old				0.00367
				(0.00902)
Post Reform × Male				0.0233***
				(0.00894)
< 300 FPL	-0.133***	-0.165***	-0.175***	-0.171***
	(0.00736)	(0.00901)	(0.00984)	(0.00984)
Eligible for Medicaid Pre-Reform	0.112***	0.110***	0.142***	0.136***
	(0.00967)	(0.00966)	(0.0131)	(0.0131)
19 to 30 Years Old	-0.0827***	-0.0829***	-0.0836***	-0.117***
	(0.00700)	(0.00699)	(0.00699)	(0.00970)
31 to 50 Years Old	-0.0360***	-0.0363***	-0.0365***	-0.0393***
	(0.00506)	(0.00505)	(0.00505)	(0.00665)
Male	-0.0294***	-0.0295***	-0.0293***	-0.0386***
	(0.00468)	(0.00467)	(0.00466)	(0.00639)
Observations	19,978	19,978	19,978	19,978
R-squared	0.114	0.117	0.118	0.122
Demographic Controls	Yes	Yes	Yes	Yes
Partial F	-	54.65	29.10	20.58

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Robust standard errors in parentheses. The sample is composed entirely of adults, 19 to 64 years of age, residing in Massachusetts in the CPS March Supplement 2002-2011. All right-hand-side variables are binary. Policy effects are argued to come through "post"-interactions only. Partial F-stat is a test that the coefficients on regressors interacted with the indicator variable for the post reform period are jointly statistically significantly different from zero. The fourth specification is used in the first stage.

Table 1.4: Utilization by Insured vs. Uninsured Persons

Hospital Service	Insured Baseline	Uninsured Baseline	Level Change	Percentage Change
Emergency Department Visits	37.81	59.77	23.07***	+38.60%***
			(7.869)	
ED Visits (Weekday, 8am-5pm)	12.37	22.33	11.199***	+50.15%***
			(4.171)	
Inpatient Admissions from ED	3.65	3.65	0.781	+20.40%
			(0.886)	
Inpatient Admissions, All	9.81	6.07	3.025***	+49.84%***
			(0.896)	

Data on utilization comes from the Massachusetts Case-Mix Database. All discharges for which there is no insurance payment are labeled as uninsured. The number of insured and uninsured are based on census population projections and CPS insurance rates. Baseline utilization rates for insured and uninsured persons is take from pre-reform levels projected forward to the post-reform time period, using growth in utilization based on pre-reform time trends.

Table 1.5: Emergency Department Visits, per 100 Persons per Year

	(1)	(2)	(3)	(4)	(5)	(6)
Fraction Insured	23.079***	25.033**	22.145***	17.772***	24.408**	22.659***
	[7.86]	[11.564]	[6.745]	[6.718]	[11.414]	[6.862]
Δ Fraction Insured × Late Post Period				10.615***	12.953***	12.339***
				[3.482]	[4.545]	[3.938]
R-squared	0.276	0.312	0.313	0.277	0.314	0.314
Subgroup Specific Time Trend	No	Yes	Yes	No	Yes	Yes
Pre-Reform HI Time Trend	No	No	Yes	No	No	Yes

Robust standard errors in parentheses, clustered at the PUMA level (52 clusters).
Bootstrapped coefficients and standard errors in italics.
*** p<0.01, ** p<0.05, * p<0.1

Estimates of the effect of health insurance coverage on emergency department services, per 100 persons, per year, and for charges associated with such utilization, per person per year. Bootstrapped coefficients (200 reps) and standard errors, clustered at the Public Use Microdata Area level, are estimated according to the procedure outlined in the appendix. All regressions include fixed effects at the unit of observation (PUMA×subgroup), quarterly dummies, and controls for monthly employment at the PUMA level. Reported R-squared is Stata's "within" R^2 and refers to non-bootstrapped regressions. Variation in insurance coverage comes through the specification shown in column 4 of Table (1.3). Subgroup specific time trends, columns (2) and (5), allow for different linear trends in each of the six subgroups. Pre-reform HI time trend, column 3 and 6, allows for an interaction between pre-reform fraction insured and the linear time trend. Data cover the time period October 1, 2001 to March 31, 2006 and July 1, 2007 to December 31, 2009; April 1, 2006 through June 30, 2007 are removed, corresponding to the period in which the policy was phased in, and data from January 1, 2010 through September 30, 2010 are removed to allow for a possible lag in discharges from time of admission. "Late Post Period" refers to January 1, 2009 through December 31, 2009, which starts 18 months after the reform is fully implemented.

Table 1.6: Emergency Department Visits, Weekdays 8:00am-5:00pm, per 100 Persons per Year

	(1)	(2)	(3)	(4)	(5)	(6)
Fraction Insured	11.199***	16.453**	10.756***	9.556***	16.180**	10.915***
	[4.171]	[7.209]	[3.873]	[3.630]	[7.092]	[3.92]
Δ Fraction Insured × Late Post Period				3.285*	5.621**	3.768*
				[1.787]	[2.635]	[2.001]
R-squared	0.195	0.243	0.245	0.195	0.243	0.246
Subgroup Specific Time Trend	No	Yes	Yes	No	Yes	Yes
Pre-Reform HI Time Trend	No	No	Yes	No	No	Yes

Bootstrapped coefficients and standard errors in italics.

*** p<0.01, ** p<0.05, * p<0.1

Estimates of the effect of health insurance coverage on emergency department services, weekdays 8:00am-5:00pm, per 100 persons, per year, and for charges associated with such utilization, per person per year. Bootstrapped coefficients (200 reps) and standard errors, clustered at the Public Use Microdata Area level, are estimated according to the procedure outlined in the appendix. For more details on the general regression framework, refer Section 4 or notes in Table (1.5).

Table 1.7: Emergency Department Visits Resulting in Hospitalization, per 100 Persons per Year

	(1)	(2)	(3)	(4)	(5)	(6)
Fraction Insured	0.781	-1.822	-0.506	0.830	-1.772	-0.529
	[0.886]	[2.160]	[1.504]	[0.846]	[2.125]	[1.519]
Δ Fraction Insured × Late Post Period				-0.099	-1.018	-0.581
				[0.485]	[0.800]	[0.670]
R-squared	0.030	0.032	0.033	0.030	0.032	0.033
Subgroup Specific Time Trend	No	Yes	Yes	No	Yes	Yes
Pre-Reform HI Time Trend	No	No	Yes	No	No	Yes

Bootstrapped coefficients and standard errors in italics.

*** p<0.01, ** p<0.05, * p<0.1

Estimates of the effect of health insurance coverage on emergency department services which result in an inpatient admission, per 100 persons, per year, and for charges associated with such utilization, per person per year. Bootstrapped coefficients (200 reps) and standard errors, clustered at the Public Use Microdata Area level, are estimated according to the procedure outlined in the appendix. For more details on the general regression framework, refer Section 4 or notes in Table (1.5).

Table 1.8: Inpatient Hospitalizations, per 100 Persons per Year

	(1)	(2)	(3)	(4)	(5)	(6)
Fraction Insured	3.025***	3.794**	4.736***	3.472***	3.834**	4.718***
	[0.896]	[1.592]	[1.329]	[0.849]	[1.577]	[1.332]
Δ Fraction Insured × Late Post Period				-0.895	-0.767	-0.455
				[0.681]	[0.719]	[0.665]
R-squared	0.031	0.032	0.033	0.031	0.032	0.033
Subgroup Specific Time Trend	No	Yes	Yes	No	Yes	Yes
Pre-Reform HI Time Trend	No	No	Yes	No	No	Yes

Bootstrapped coefficients and standard errors in italics.

*** p<0.01, ** p<0.05, * p<0.1

Estimates of the effect of health insurance coverage on inpatient admissions, per 100 persons, per year, and for charges associated with such utilization, per person per year. Bootstrapped coefficients (200 reps) and standard errors, clustered at the Public Use Microdata Area level, are estimated according to the procedure outlined in the appendix. For more details on the general regression framework, refer Section 4 or notes in Table (1.5).

Table 1.9: Potentially Avoidable Hospitalizations, per 100 Persons per Year

	(1)	(2)	(3)	(4)	(5)	(6)
Fraction Insured	-1.257***	0.829	-0.887*	-1.08***	0.810	-0.897*
	[0.317]	[0.803]	[0.485]	[0.314]	[0.782]	[0.487]
Δ Fraction Insured $\times$ Late Post Period				-0.349**	0.319	-0.282*
				[0.133]	[0.310]	[0.164]
R-squared	0.134	0.151	0.170	0.134	0.151	0.170
Subgroup Specific Time Trend	No	Yes	Yes	No	Yes	Yes
Pre-Reform HI Time Trend	No	No	Yes	No	No	Yes

Bootstrapped coefficients and standard errors in italics.

*** p<0.01, ** p<0.05, * p<0.1

Estimates of the effect of health insurance coverage on the subset of inpatient admissions which are deemed "potentially avoidable" with adequate preventive care by the AHRQ, per 100 persons, per year, and for charges associated with such utilization, per person per year. See Appendix for more details on classification. Bootstrapped coefficients (200 reps) and standard errors, clustered at the Public Use Microdata Area level, are estimated according to the procedure outlined in the appendix. For more details on the general regression framework, refer Section 4 or notes in Table (1.5).

Table 1.10: Miller Replication, Emergency Department Visits per Capita

	(1)	(2)
Interim	0.0114	0.0114
	(0.0118)	(0.00659)
Post-Reform	0.0154	0.0154**
	(0.0136)	(0.00678)
Fraction Uninsured in 2005	0.128***	0.128
	(0.0409)	(0.170)
Interim × Fraction Uninsured in 2005	-0.0349	-0.0349
	(0.0846)	(0.0489)
Post-Reform × Fraction Uninsured in 2005	-0.0571	-0.0571
	(0.0983)	(0.0533)
Observations	392	392
R-squared	0.052	0.052

Heteroscedasticity-robust SE in column 1, Cluster-robust SE in column 2

*** $p<0.01$, ** $p<0.05$, * $p<0.1$

Replicates “Total ER Visits” from the first column in Table 2 of Miller (2012). Dependent variable is quarterly, adult (18-64), ER visits per capita in each of fourteen Massachusetts counties 2002-2008, using MA Case-Mix Database. Standard errors in column 1 are heteroscedasticity robust, while standard errors in column 2 are clustered at the county level and corrected using STATA’s ”vce(cluster)” command. All regressions are weighted by the average adult county population from 2002-2008 using U.S. Census Bureau Intercensal Estimates.

Chapter 2

(Re)Funding Health Care Spending: The Timing of EITC Refunds, Liquidity, and Investment in Health

2.1 Introduction

There exists a broad economics and epidemiological literature on the strong positive relationship between income and health. Deaton (2002), among many others, has documented the "health-wealth" gradient, but there has been little work that has shown a causal relationship, likely because it is hard to find exogenous variation in income. Stowasser et al. (2012) rule out, using panel data, some causal pathways for income to affect future health, but many remain open. Acemoglu et al. (n.d.) find that oil price shocks that created large income shocks for local economies in oil-producing counties serve to increase spending.

One potentially fruitful, but under-utilized avenue for investigating the relationship between income and health exists at the intersection of public and health economics. Given that we care about the relationship between health and *after*-tax income, one can use plausibly exogenous variation in taxes and tax credits to identify the causal effect of income on a wide variety of outcomes, as in Dahl and Lochner (2012). One potential concern about using variation in taxes affects prices is that there could be a series of behavioral responses, such as labor supply responses, which may make it difficult to isolate the pure income effect.

I am not the first to study the effect of the EITC on health. Hoynes et al. (2012) find large decreases in low birth weight and infant mortality among the children of low-income women after the EITC expands in 1993 and 1997. Hoynes et al. (2012) show a strong, positive effect on infant health due to the EITC, but are unable to unpack the effects and ignore the intra-year liquidity effects induced by the lump-sum nature of the payment. Amarante et al. (2011) also find a large positive impact of income on infant health, but in the context of cash payments to low-income, expectant mothers in Uruguay. They are able to unpack the effect to highlight the roles of mother's nutrition, labor supply, and decreased smoking.

LaLumia (2013) looks at within-year liquidity in the context of unemployment. She finds that, among those most likely to be eligible for the EITC, unemployment spells that begin in February are longer than those that begin at other times of the year. Gross and Tobacman (2013) look at the timing of stimulus checks in 2007 and find that, shortly after receipt, there is an increase in emergency room visits due to increases in injuries related to drug and alcohol abuse.

In this paper, I use a large, predictable increase in liquidity during the month of February to identify the effects of cash liquidity on medical care utilization. I find that an additional 9% of EITC recipients purchase care in February, a sizable increase. Two main categories of spending emerge: (1) individuals schedule more appointments in office-based settings, and (2) they seek more care in the emergency room. The first result indicates an increase in investment, while the second likely indicates an increase in injuries that require medical care. Others have shown that liquidity can lead to more ER visits due to injury, but this paper is the first, to my knowledge, to show how liquidity affects health investment, shedding new light on a potential mechanism underpinning the relationship between income and health.

The paper proceeds as follows. Section 2 covers the necessary background material on the Earned Income Tax Credit and public insurance eligibility for those eligible for EITC, Section 3 describes the identification, Section 4 describes the Medical Expenditure Panel Survey data, Section 5 analyzes the results, and Section 6 concludes.

2.2 Background

2.2.1 Earned Income Tax Credit

The Earned Income Tax Credit (EITC) is a large, refundable tax credit provided to the working poor in the United States. The size of the credit depends on earnings and the number of dependent children present in the household. In 2007 for a single parent with two children, the EITC provided a 40% subsidy to wages up to $11,790, then was phased out with an additional tax of 21.06% from $15,390 to $37,783, for a maximum credit of $4,716.[1] See Figure (**??**) for a graphical representation and comparison to families with one or no children. In 2002, the earnings level at which the credit started to phase out was extended $1,000 for married couples filing jointly, relative to single filers, and another change in 2005 extended this region to $2,000.

Tax filers have two options for claiming the EITC. The first option is receive the credit, net of other taxes, as a lump sum payment after filing taxes for the previous year. The second option is to forecast earnings for the upcoming year and request EITC be paid out in advance with each pay check. As documented in LaLumia (2013), in practice, the vast majority of tax filers who qualify for the EITC choose the first option, filing in January and receiving a lump-sum payment in February.[2] Some EITC recipients file slightly later and receive payment in March or April. Furthermore, the proliferation of businesses providing cash advances on tax returns, for a sizable fee, means that more EITC recipients may be receiving their money in January. Both are potential sources of measurement error.

2.2.2 Insurance Status Among EITC Recipients

Among children who live in families receiving EITC payments, many are eligible for Medicaid, the jointly-funded federal and state public insurance program for those with low income, but eligibility varies across states and over time.[3] In 1996, the minimum national standards for eligibility for pregnant women and children under thirteen was 133% of the Federal Poverty Line (FPL), or $22,893.33 in 2007 dollars, covering most but not all of those eli-

[1] All values reported in 2007 U.S. dollars.

[2] Receipt of the Child Tax Credit follows a similar pattern, but it is not investigated in this paper

[3] Medicaid also covers those with exceptionally high medical costs relative to income.

gible for EITC. Some states, such as Hawaii, expanded eligibility to state-run programs for all citizens up to 300% FPL, well past the range of incomes eligible for the EITC. Children thirteen and older, as well as non-pregnant adults, were only eligible for public insurance coverage if family income fell below the poverty thresholds for cash payments under pre-1996 Aid to Families with Dependent Children (AFDC), commonly referred to as welfare. Some states, such as Alabama, set this eligibility threshold at 13.4% FPL, or about $2,300 in annual earnings.

In 2007, with the expansion of the Children's Health Insurance Program (CHIP), most children under the age of nineteen in families earning up to 200% FPL, or $34,340, were eligible for free or highly subsidized public insurance, although children under nineteen living in North Dakota were only eligible up to 140% FPL. Adults with children in many states were still only eligible for public insurance at the very low pre-1996 AFDC levels, but some states began to expand insurance for families with higher incomes and childless adults, most notably the Massachusetts insurance expansion in 2006 which provides free public insurance to all residents up to 150% FPL and generous subsidies up to 300% FPL.

Medicaid and CHIP insurance, while reimbursing doctors and hospitals at low rates, requires very low co-payments or coinsurance rates for patients. Services covered and the availability of doctors seeing Medicaid/CHIP patients varies from state to state.

2.3 Identification

Identification hinges on the timing of the EITC payments, which are highly clustered around February and early March as shown through Treasury Reports in LaLumia (2013). The identifying assumption is that, conditional on all other covariates, the only way EITC recipients and non-EITC recipients differ in the month of February is through the large cash payment to EITC recipients. The main regression specification, with data that varies at the individual, i, month m, and year t level, is as follows:

$$Y_{imt} = \beta_1 EITC_i \times Feb_m + \beta_2 EITC_i + \beta_3 Feb_t + \beta_4 X_{imt} + \beta_5 T_t + \varepsilon_{imt} \quad (2.1)$$

The treatment effect is β_1, the coefficient on the interaction term, as the "treatment" is the receipt of the EITC payment in the month of February. Other covariates include a cubic polynomial for age, an indicator variable for gender, and an indicator variable for the individual's insurance status in a given month, private insurance or Medicaid. EITC income is calculated from lagged income variables and family structure variables using the NBER Taxsim as described in Feenberg and Coutts (1993). MEPS data includes a variable indicating whether or not an individual received any EITC payments, although there is no data on the actual amount received. Standard errors are clustered at the level of variation of the EITC, the subfamily unit, as defined by the MEPS.

The main specification is a probit regression where the dependent variable is an indicator for positive expenditures. In one set of regressions, the EITC variable is an indicator for positive EITC receipt, and in the other, it is continuous. In some regressions, I include additional interactions with an "uninsured" indicator variable to show how EITC receipt in Feb affects the uninsured differently.

To interpret the β_1 coefficient in the context of a probit regression, I calculate the average marginal effect for those who are EITC eligible in February. Before calculating the marginal effects, one can simply look at the sign of the coefficients in the probit difference-in-differences regression to learn the sign of the marginal effect as shown in Puhani (2012). Unlike some cases, as shown in Ai and Norton (2003), in the difference-in-differences framework, the treatment is the coefficient on the interaction term only. As shown in Puhani (2012) the treatment effect can be written as the difference of two cross-differences:

$$\Phi(\beta_1 + \beta_2 + \beta_3 + \beta_4 X_{it} + \beta_5 T_t) - \Phi(\beta_2 + \beta_3 + \beta_4 X_{it} + \beta_5 T_t) \tag{2.2}$$

Where Φ represents the normal CDF. When using non-linear models like the probit, I report average marginal treatment effects for the set I of EITC recipients in February: [4]

$$\frac{1}{N_I} \sum_{i \in I} [\Phi(\beta_1 + \beta_2 + \beta_3 + \beta_4 X_{it} + \beta_5 T_t) - \Phi(\beta_2 + \beta_3 + \beta_4 X_{it} + \beta_5 T_t)] \tag{2.3}$$

In addition to looking at the probability of positive spending, I examine the number of visits per month corresponding to certain subcategories of spending: (1) Office-Based, (2) ER, (3) Dental, (4) Outpatient, and (5) Inpatient. I use the negative binomial model to account for the discreteness of the outcome data and the restriction that counts be non-negative.

Moving beyond the probit and negative binomial models, with important structural assumptions, I also use fully-saturated linear models where I fully interact a set of dummy variables with indicators for (1) positive EITC receipt and (2) February. Some specifications also include dummies for insurance status (also fully interacted with the other variables). The advantage of the saturated linear model is that it makes no functional form assumptions. The downside is that it sacrifices precision by including only binary regressors.

To investigate the effect of EITC income on levels of spending, I currently use fully-saturated linear models.[5]

[4] One way to think about the average marginal treatment effect is to compute the difference in probability of positive spending for each EITC-eligible individual in each February with the interaction term set to zero vs. the interaction term set to one, and then averaging this effect over all EITC-eligible individual in each February in the data.

[5] I have estimated the model using Tobit estimation, but the data violates the normality assumption even after transforming the spending data to log(spending) because the outcome variable's distribution has a thick right tail.

2.4 Data

All data comes from the Medical Expenditure Panel Survey (MEPS) for years 1996-2007. The MEPS is a nationally representative, two-year, overlapping panel with detailed and accurate data on expenditures, including the source of payment, as well as extensive data on income by source. Importantly, the MEPS allows for medical care use to be disaggregated down to the episode level, with information about the day, month, and year when the event occurred, which allows for the matching of utilization to the timing of EITC receipt. Payment information provides a spending breakdown into categories of payer, including out-of-pocket, payments by various insurers, and total charges.

Because taxes are filed the following year and EITC is received shortly thereafter, it is crucial to have lagged income and family structure data as well as contemporary data on medical care utilization and spending. The two-year panel structure of the MEPS allows for this. I use the first year to estimate the EITC payment and match this to monthly outcomes in year two.

I look at five subcategories of utilization as defined in the MEPS Event level data: (1) Office-Based, (2) ER, (3) Dental, (4) Outpatient, and (5) Inpatient.[6][7]

- (1) Office-Based: includes all visits and expenditures in an office-setting. Laboratory services and other diagnostic tests provided in conjunction with these visits are included. E.g. physical therapy, allergy shots, psychotherapy/counseling among others.
- (2) ER: includes all visits to outpatient emergency departments. Laboratory services and other diagnostic tests provided in conjunction with these visits are included.
- (3) Dental: all out-patient visits to a dentist's office.

[6]Prescription drug results coming soon. Need to account for the fact that the events are at the "wave" level (wave $\approx$ 4-5 months) instead of monthly.

[7]Home health visits are ignored as they are rare for the non-elderly.

- (4) Outpatient: outpatient services provided in a hospital setting.
- (5) Inpatient: inpatient services provided in a hospital setting.

For a summary of the data, refer to Table (3.1). All individuals 60 years of age or older are dropped from the dataset in order to keep the control group more similar to the EITC group, which is largely younger parents and their children. In order to focus on the uninsured and those with Medicaid or private insurance, those with Medicare or those with military insurance are dropped as well.[8] All individuals in families which do not have a consistent membership for the full two years are dropped. Results are not sensitive to these exclusions.

2.5 Results

2.5.1 Utilization

Tables (2.2) and (2.3) show probit regressions where the dependent variable is an indicator for positive out-of-pocket spending (columns 1 and 2) and positive total spending (columns 3 and 4). In Table (2.2), column 1, the key regressor is the EITC indicator interacted with an indicator for the month of February. The EITC indicator equals one if family receives at least $100 in EITC in that year. The positive coefficient indicates that EITC recipients are predicted to be more likely to consume medical care during the month when they receive EITC. In column 2, I add an additional interaction for uninsured status. The positive coefficient on the triple interaction – EITC, Feb., and uninsured – shows that the demand for medical care for the uninsured is even more responsive to the influx of cash. It is likely that the higher price of care for the uninsured interacts with credit-constraints to reduce medical care utilization among this group. Similar patterns emerge in columns (3) and (4)

[8]Medicare is free public insurance for those sixty-five and older and the disabled. It is also available to those with certain serious medical conditions, e.g. end-stage renal failure.

where the outcome variable is total charges.

Table (2.3) shows results where the EITC variable is continuous. Likelihood of spending in February appears to increase with the size of the EITC payment, providing further proof that the EITC is the cause. Each $1,000 of EITC received is predicted to increase the likelihood of spending by 0.4 percentage points, or roughly a 4% increase from baseline. Similar patterns in columns (2) through (4) emerge as in Table (2.2).

Table (2.4) breaks spending into five categories and looks at the effect of EITC on each individually. The only categories which appear to respond are office-based visits (column 1) and ER visits (column 2). The probability of any office visit increases by 0.7 percentage points, or roughly 10% from baseline. The probability of any ER visit increases by 0.26 percentage points, or about 47% from baseline.

Table (2.5) uses a negative binomial regression model to look at counts of visits in a given month. EITC receipt increases the total number of office-based visits by about 0.0144 per month, or 4%, and the total number of ER visits by 19%. These have the same sign as the estimates in Table (2.4), but appear smaller, although they are not statistically significantly different at the 5% level. While not conclusive, this suggests that the increase in utilization is coming from low-volume users.

Results of the saturated models are reported in Tables (2.6)-(2.10). While generally imprecise, some linear combinations of estimators, which include both the "main" EITC-February effect and the interaction of the EITC-February term with each insurance status variable, give estimates that are significant at the 10% level and correspond roughly to the estimates produced by the non-linear models.

2.6 Conclusion

This paper provides evidence which strongly suggests that the timing of the EITC payment affects the timing of medical spending. Liquidity constraints matter, especially for those without insurance. I find evidence for two types of increases in medical spending: (1) an increase in the demand for emergency services, likely due to the increase in injuries and substance abuse and (2) an increase in health investment, shown by the increase in office-based visits. Other studies have shown a positive effect of EITC payments and infant health, but to my knowledge, this is the first to examine the health investment channel, in particular in relation to household cash liquidity.

Past studies, including Gross and Tobacman (2013) and Ruhm (2000), have focused on the short-run negative health consequences of consumption and economic activity. A narrow focus on the short-run costs has overlooked the potential long-run benefits, which are more difficult to quantify. This paper has shown an increase in one type of investment – physician's office visits – as a starting point for investigating other health investments that may be hampered by a lack of liquidity. Future research should go further to quantify the effects of other types of spending including prescription drugs.

Another future research question is related to how transfer programs like the EITC affect not only the timing of spending, but also the level of spending. Using changes in the EITC and other tax programs over time will allow us to investigate how they may stimulate an increased investment in health.

Ultimately, precisely quantifying the liquidity effects on medical care spending produces an important parameter estimate for determining optimal health insurance contracts, as shown in the case of unemployment insurance in Chetty (2008). Assuming the liquidity effect is zero causes estimates of distortionary moral hazard in health insurance to be too high. Taking this to the data, I hope to exploit variation in cost-sharing across state-level Medicaid programs in the early twenty-first century. The combination of spatial and time variation for these policy changes will hopefully allow me to estimate the liquidity and moral hazard elasticities and speak to the issue of optimal health insurance design, at least in the context of Medicaid.

2.7 Figures

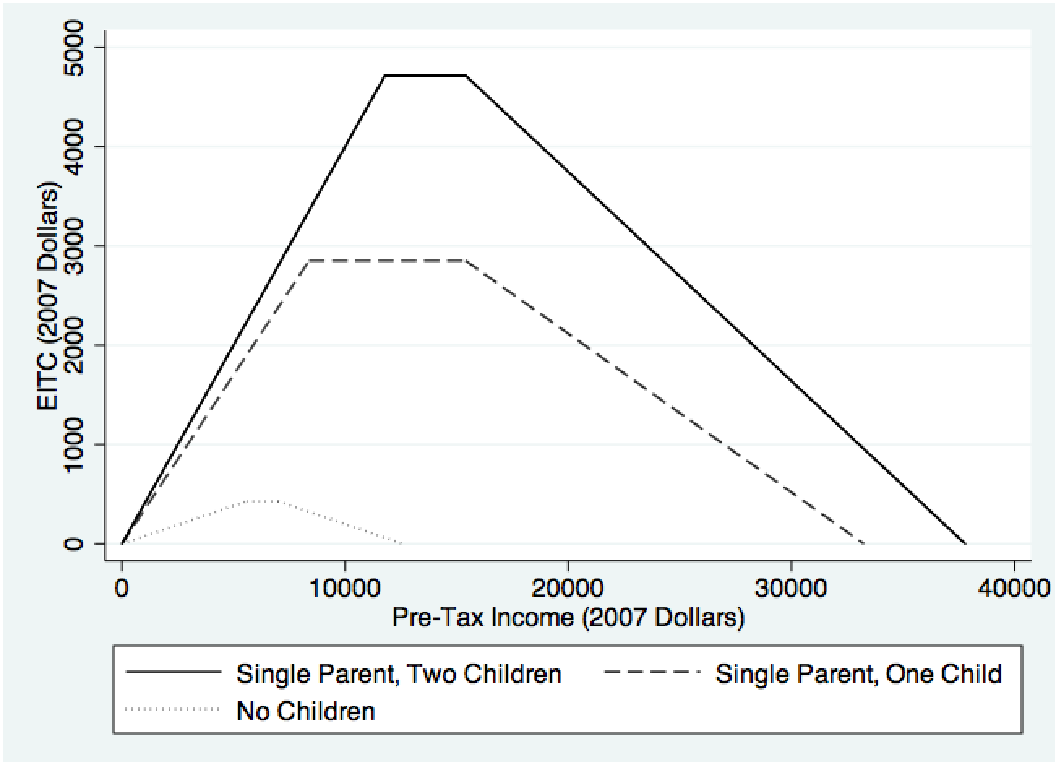

Figure 2.1: EITC Program Parameters, 2007

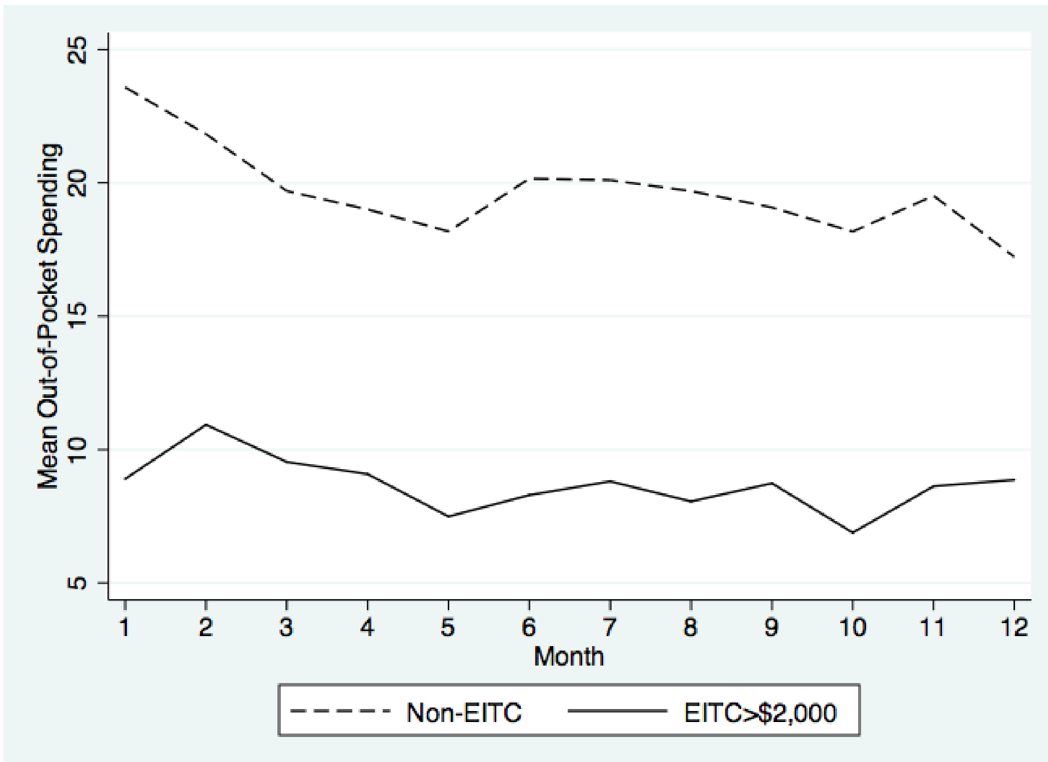

Figure 2.2: Out-of-Pocket Spending by Calendar Month, EITC vs. Non-EITC

2.8 Tables

Table 2.1: Summary Statistics

	Full Sample	EITC> 0	EITC= 0
Age	29.595	25.253	31.326
	(16.911)	(15.808)	(17.026)
Male	0.498	0.473	0.509
	(0.5)	(0.499)	(0.5)
Medicaid	0.116	0.25	0.063
	(0.321)	(0.433)	(0.244)
Private Health Insurance	0.703	0.47	0.795
	(0.457)	(0.499)	(0.404)
Uninsured	0.19	0.298	0.147
	(0.392)	(0.457)	(0.354)
Positive EITC Receipt	0.264	0.927	0
	(0.441)	(0.26)	(0)
EITC ($2007)	460.027	1614.459	0
	(1075.044)	(1480.609)	(0)
Out-of-Pocket Expenses ($2007)	23.031	15.529	26.02
	(234.515)	(255.658)	(225.468)
Dental Expenses (Out-of-Pocket)	9.133	5.811	10.457
	(122.999)	(97.631)	(131.731)
Office Visit Expenses (Out-of-Pocket)	9.159	5.502	10.616
	(99.741)	(83.696)	(105.421)
ER Expenses (Out-of-Pocket)	1.096	1.285	1.021
	(39.661)	(50.611)	(34.338)
Outpatient Expenses (Out-of-Pocket)	1.597	0.918	1.868
	(62.047)	(44.611)	(67.755)
Inpatient Expenses (Out-of-Pocket)	1.888	2.007	1.84
	(150.824)	(207.419)	(121.113)
Total Expenses ($2007)	136.331	101.017	150.404
	(1691.108)	(1436.457)	(1782.28)
Dental Expenses (Total)	20.958	13.614	23.885
	(193.178)	(153.369)	(206.846)
Office Visit Expenses (Total)	41.432	28.146	46.727
	(324.711)	(222.237)	(357.315)
ER Expenses (Total)	7.501	7.425	7.531
	(152.594)	(162.228)	(148.581)
Outpatient Expenses (Total)	18.144	12.53	20.381
	(331.081)	(273.094)	(351.508)
Inpatient Expenses (Total)	45.689	37.09	49.115
	(1558.19)	(1349.762)	(1633.842)
N	1458920	530585	928335

Table 2.2: Spending Indicator, EITC Indicator

	(1)	(2)	(3)	(4)
	Pos. OOP	Pos. OOP	Pos. Total	Pos. Total
Pos. EITC × Feb.	0.00866***	0.00401	0.00980***	0.00722*
	(0.00228)	(0.00256)	(0.00265)	(0.00301)
Positive EITC Receipt	-0.0354***	-0.0694***	-0.0446***	-0.0413***
	(0.00127)	(0.00134)	(0.00165)	(0.00175)
February	0.00862***	0.00933***	0.0120***	0.0128***
	(0.00115)	(0.00126)	(0.00145)	(0.00158)
Medicaid	-0.141***		0.0180***	
	(0.000960)		(0.00253)	
Uninsured	-0.0995***	-0.0775***	-0.125***	-0.130***
	(0.00110)	(0.00132)	(0.00128)	(0.00163)
Age	-0.00500***	-0.00293***	-0.00735***	-0.00760***
	(0.000296)	(0.000317)	(0.000349)	(0.000346)
Age Squared	0.000151***	0.000126***	0.000239***	0.000242***
	(0.0000113)	(0.0000120)	(0.0000136)	(0.0000136)
Age Cubed	-0.000000877***	-0.000000820***	-0.00000157***	-0.00000156***
	(0.000000123)	(0.000000130)	(0.000000151)	(0.000000150)
Male	-0.0508***	-0.0462***	-0.0689***	-0.0695***
	(0.000787)	(0.000824)	(0.00102)	(0.00102)
Pos. EITC × Unins.		0.0404***		0.000832
		(0.00358)		(0.00367)
Feb. × Unins.		-0.00268		-0.00522
		(0.00319)		(0.00390)
Observations	1518312	1518312	1518312	1518312
Year Effects	Yes	Yes	Yes	Yes

Standard errors, clustered at family level, in parentheses
(d) for discrete change of dummy variable from 0 to 1
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Marginal effects for probit regression. EITC variable is an indicator variable set equal to one for those predicted to receive greater than $100 in EITC income. Dependent variable is an indicator for positive spending in a given month. Columns (1) and (2) look at out-of-pocket spending while columns (3) and (4) look at total spending.

Table 2.3: Spending Indicator, EITC in Dollars (2007)

	(1)	(2)	(3)	(4)
	Pos. OOP	Pos. OOP	Pos. Total	Pos. Total
EITC (\$2007) × Feb.	0.00000395***	0.00000287**	0.00000278**	0.00000228
	(0.000000920)	(0.00000110)	(0.00000103)	(0.00000118)
EITC (\$2007)	0	0	0	0
	(0.000000619)	(0.000000847)	(0.000000698)	(0.000000750)
February	0.00908***	0.00923***	0.0135***	0.0136***
	(0.00108)	(0.00119)	(0.00134)	(0.00147)
Medicaid	-0.125***		0.0161***	
	(0.000832)		(0.00231)	
Uninsured	-0.0886***	-0.0775***	-0.128***	-0.131***
	(0.000939)	(0.00117)	(0.00127)	(0.00142)
EITC (\$2007) × Feb. × Unins.		0.00000256		0.00000221
		(0.00000208)		(0.00000238)
EITC (\$2007) × Unins.		0.0000166***		0.000000377
		(0.00000123)		(0.00000133)
Feb. × Unins.		0.00111		-0.00106
		(0.00286)		(0.00345)
Observations	1518312	1518312	1518312	1518312
Year Effects	Yes	Yes	Yes	Yes

Standard errors, clustered at family level, in parentheses
(d) for discrete change of dummy variable from 0 to 1
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Marginal effects for probit regression. EITC variable is the level of EITC income in 2007 USD. Dependent variable is an indicator for positive spending in a given month. Columns (1) and (2) look at out-of-pocket spending while columns (3) and (4) look at total spending. Columns (2) and (4) include additional interactions for uninsured status.

Table 2.4: Spending Indicator, by Subcategory, EITC Indicator

	(1)	(2)	(3)	(4)	(5)
	Office-Based	ER	Dental	Outpatient	Inpatient
Pos. EITC × Feb.	0.00706***	0.00257***	0.00161	0.000544	-0.000210
	(0.00203)	(0.000617)	(0.00137)	(0.000553)	(0.000273)
Positive EITC Receipt	-0.0285***	-0.000177	-0.00955***	-0.00115***	-0.000372***
	(0.00113)	(0.000149)	(0.000512)	(0.000199)	(0.0000825)
February	0.00969***	-0.000872***	-0.000764	-0.000157	0.0000972
	(0.00101)	(0.000224)	(0.000622)	(0.000248)	(0.000153)
Medicaid	-0.0983***	-0.00379***	-0.0302***	-0.00449***	-0.000939***
	(0.000718)	(0.000101)	(0.000288)	(0.000152)	(0.0000869)
Uninsured	-0.0703***	-0.0000892	-0.0221***	-0.00416***	-0.000592***
	(0.000827)	(0.000151)	(0.000354)	(0.000131)	(0.0000766)
Observations	1518312	1518312	1518312	1518312	1518312

Standard errors, clustered at family level, in parentheses

(d) for discrete change of dummy variable from 0 to 1

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Marginal effects for probit regression. EITC variable is an indicator variable set equal to one for those predicted to receive greater than $100 in EITC income. Dependent variable is an indicator for positive spending in a given month in each of the five categories listed. See text for more details on categories.

Table 2.5: Visits, by Subcategory

	(1) Office-Based	(2) ER	(3) Dental	(4) Outpatient	(5) Inpatient
Pos. EITC × Feb.	0.0144*	0.00347***	0.00248	0.00222	0.000625
	(0.00579)	(0.000960)	(0.00247)	(0.00193)	(0.000523)
Positive EITC Receipt	-0.0781***	0.000649*	-0.0226***	-0.00599***	-0.000696***
	(0.00417)	(0.000323)	(0.00101)	(0.000999)	(0.000155)
February	0.0203***	-0.000876	0.000269	0.000333	-0.0000513
	(0.00303)	(0.000460)	(0.00111)	(0.000948)	(0.000273)
Medicaid	0.0657***	0.0117***	-0.0307***	0.0185***	0.00801***
	(0.00748)	(0.000653)	(0.00104)	(0.00257)	(0.000433)
Uninsured	-0.161***	0.00162***	-0.0598***	-0.0130***	-0.00126***
	(0.00326)	(0.000365)	(0.000609)	(0.000800)	(0.000161)
Observations	1518312	1518312	1518312	1518312	1518312

Standard errors, clustered at family level, in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Negative binomial regression for counts of visits in a given month. EITC variable is an indicator variable set equal to one for those predicted to receive greater than $100 in EITC income. Dependent variable is a count of visits corresponding to one of five categories. See text for more details on categories.

Table 2.6: Positive Spending: Fully-Saturated Linear Model

	(1)	(2)	(3)	(4)
	Pos. OOP	Pos. OOP	Pos. Total	Pos. Total
Pos. EITC × Feb.	0.00160	0.00354	0.00538*	0.00128
	(0.00195)	(0.00361)	(0.00251)	(0.00397)
Positive EITC Receipt	-0.0831***	-0.0558***	-0.0712***	-0.0661***
	(0.00133)	(0.00217)	(0.00165)	(0.00248)
February	0.0102***	0.0132***	0.0129***	0.0143***
	(0.00132)	(0.00174)	(0.00155)	(0.00193)
Pos. EITC × Feb. × Unins.		0.00660		0.00970
		(0.00489)		(0.00545)
Pos. EITC × Feb. × Medicaid		0.00192		0.00790
		(0.00449)		(0.00747)
Pos. EITC × Unins.		0.0264***		0.0347***
		(0.00276)		(0.00322)
Pos. EITC × Medicaid		0.0413***		0.0157**
		(0.00270)		(0.00479)
Feb. × Unins.		-0.00828**		-0.00899**
		(0.00293)		(0.00328)
Feb. × Medicaid		-0.0146***		0.00192
		(0.00273)		(0.00528)
Constant	0.165***	0.204***	0.237***	0.264***
	(0.000922)	(0.00109)	(0.00105)	(0.00120)
Observations	1518312	1518312	1518312	1518312

Standard errors, clustered at family level, in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Fully saturated linear probability model. Dependent variable is an indicator for positive spending. EITC variable is an indicator variable set equal to one for those predicted to receive greater than $100 in EITC income. February and Uninsured indicators included but not reported.

Table 2.7: OOP Spending: Fully-Saturated Linear Model

	(1)	(2)	(3)	(4)
	OOP Exp.	OOP Exp.	Total Exp.	Total Exp.
Pos. EITC × Feb.	1.150	0.902	18.38	10.83
	(1.364)	(1.532)	(11.62)	(11.41)
Positive EITC Receipt	-11.29***	-8.437***	-54.68***	-36.30***
	(0.439)	(0.697)	(2.985)	(4.831)
February	0.394	-0.413	-6.747	-13.12*
	(1.114)	(0.757)	(5.242)	(5.492)
Pos. EITC × Feb. × Unins.		-1.429		24.12
		(5.782)		(28.89)
Pos. EITC × Feb. × Medicaid		0.572		-28.88
		(1.969)		(31.54)
Pos. EITC × Unins.		0.108		7.539
		(1.274)		(6.099)
Pos. EITC × Medicaid		6.746***		-41.87***
		(0.907)		(10.31)
Feb. × Unins.		4.752		14.10
		(5.381)		(9.930)
Feb. × Medicaid		-0.713		33.80
		(1.029)		(28.93)
Constant	22.91***	26.14***	142.3***	155.3***
	(0.307)	(0.343)	(2.180)	(2.649)
Observations	1518312	1518312	1518312	1518312

Standard errors, clustered at family level, in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Fully saturated linear probability model. Dependent variable is an indicator for positive spending. EITC variable is an indicator variable set equal to one for those predicted to receive greater than $100 in EITC income. February and Uninsured indicators included but not reported.

Table 2.8: Visits, by Subcategory, Fully-Saturated Linear Model

	(1) Office-Based	(2) ER	(3) Dental	(4) Outpatient	(5) Inpatient
Pos. EITC × Feb. × Unins.	0.0117	-0.00158	0.00594	0.00202	0.0000355
	(0.0115)	(0.00198)	(0.00409)	(0.00319)	(0.000997)
Pos. EITC × Feb. × Medicaid	0.00993	-0.000685	0.00514	0.00525	0.000608
	(0.0145)	(0.00282)	(0.00546)	(0.00552)	(0.00166)
Pos. EITC × Feb.	-0.00298	0.00348**	-0.00171	0.000124	0.0000495
	(0.00752)	(0.00122)	(0.00323)	(0.00223)	(0.000677)
Pos. EITC × Unins.	0.0485***	0.000612	0.0190***	0.00369	-0.000372
	(0.00824)	(0.000654)	(0.00212)	(0.00201)	(0.000313)
Pos. EITC × Medicaid	-0.0113	-0.00416***	0.0284***	-0.0169***	-0.00256***
	(0.0135)	(0.00101)	(0.00266)	(0.00413)	(0.000556)
Feb. × Unins.	-0.00196	0.00147	-0.00282	-0.00121	0.000267
	(0.00759)	(0.00108)	(0.00252)	(0.00205)	(0.000639)
Feb. × Medicaid	-0.00345	0.00261	-0.00103	0.000479	0.00136
	(0.0110)	(0.00196)	(0.00368)	(0.00457)	(0.00126)
Positive EITC Receipt	-0.108***	0.00115**	-0.0317***	-0.00790***	-0.000476*
	(0.00610)	(0.000386)	(0.00173)	(0.00133)	(0.000214)
Medicaid	0.0198	0.0126***	-0.0490***	0.0153***	0.00571***
	(0.0110)	(0.000768)	(0.00186)	(0.00376)	(0.000446)
Constant	0.365***	0.0107***	0.105***	0.0281***	0.00489***
	(0.00333)	(0.000159)	(0.000848)	(0.000731)	(0.000102)
Observations	1518312	1518312	1518312	1518312	1518312

Standard errors, clustered at family level, in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Fully saturated linear probability model. Dependent variable is an indicator for positive spending. EITC variable is an indicator variable set equal to one for those predicted to receive greater than $100 in EITC income. February and Uninsured indicators included but not reported.

Table 2.9: OOP Spending: Fully-Saturated Linear Model

	(1)	(2)	(3)	(4)	(5)
	Office-Based	ER	Dental	Outpatient	Inpatient
Pos. EITC × Feb. × Unins.	0.598	-0.622	1.299	0.140	-2.715
	(1.571)	(1.085)	(1.639)	(1.020)	(5.020)
Pos. EITC × Feb. × Medicaid	-0.111	-0.394	0.418	-0.304	0.670
	(0.876)	(0.272)	(1.484)	(0.430)	(0.732)
Pos. EITC × Feb.	0.503	0.590*	0.474	0.486	-1.169*
	(0.844)	(0.246)	(1.064)	(0.409)	(0.473)
Pos. EITC × Unins.	0.496	0.452	1.010	-0.00472	-1.735*
	(0.536)	(0.276)	(0.557)	(0.255)	(0.830)
Pos. EITC × Medicaid	3.122***	-0.0539	3.050***	0.365*	0.588
	(0.366)	(0.0909)	(0.469)	(0.178)	(0.512)
Feb. × Unins.	0.722	0.828	0.243	0.364	2.444
	(1.118)	(0.889)	(0.979)	(0.611)	(4.981)
Feb. × Medicaid	-0.690	-0.0939	0.918	0.0822	-0.651
	(0.409)	(0.124)	(0.724)	(0.222)	(0.500)
Positive EITC Receipt	-3.918***	-0.0966	-3.347***	-0.684***	-0.270
	(0.325)	(0.0541)	(0.430)	(0.149)	(0.300)
February	0.239	-0.152*	-0.794	-0.216	0.545
	(0.349)	(0.0775)	(0.537)	(0.190)	(0.326)
Constant	10.64***	0.800***	11.07***	1.876***	1.613***
	(0.172)	(0.0271)	(0.217)	(0.110)	(0.119)
Observations	1518312	1518312	1518312	1518312	1518312

Standard errors, clustered at family level, in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Fully saturated linear probability model. Dependent variable is an indicator for positive spending. EITC variable is an indicator variable set equal to one for those predicted to receive greater than $100 in EITC income. February and Uninsured indicators included but not reported.

Table 2.10: Total Spending: Fully-Saturated Linear Model

	(1)	(2)	(3)	(4)	(5)
	Office-Based	ER	Dental	Outpatient	Inpatient
Pos. EITC × Feb. × Unins.	-3.980	2.164	1.709	-9.766	34.89
	(3.063)	(2.522)	(2.196)	(5.969)	(28.30)
Pos. EITC × Feb. × Medicaid	-4.377	0.903	0.894	-5.464	-19.93
	(5.130)	(1.721)	(2.394)	(5.840)	(29.76)
Pos. EITC × Feb.	3.274	1.230	0.212	6.675	-0.753
	(2.133)	(1.171)	(1.707)	(4.896)	(9.759)
Pos. EITC × Unins.	7.699***	-0.377	4.420***	1.905	-5.484
	(1.608)	(0.644)	(0.899)	(1.259)	(5.021)
Pos. EITC × Medicaid	3.502	-1.222	6.542***	-3.278	-33.10***
	(2.201)	(0.835)	(0.955)	(2.039)	(8.275)
Feb. × Unins.	4.314	1.712	0.173	5.136	2.270
	(2.305)	(1.275)	(1.290)	(3.476)	(8.649)
Feb. × Medicaid	6.357	-1.223	1.603	1.444	24.94
	(3.962)	(0.975)	(1.468)	(3.084)	(27.83)
Positive EITC Receipt	-15.74***	0.384	-7.535***	-5.280***	-7.775*
	(1.338)	(0.482)	(0.775)	(1.082)	(3.927)
Medicaid	-8.971***	0.676	-17.78***	-5.911***	37.95***
	(1.663)	(0.491)	(0.578)	(1.721)	(6.861)
Constant	48.90***	7.748***	26.08***	22.28***	49.13***
	(0.824)	(0.201)	(0.371)	(0.567)	(2.072)
Observations	1518312	1518312	1518312	1518312	1518312

Standard errors, clustered at family level, in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Fully saturated linear probability model. Dependent variable is an indicator for positive spending. EITC variable is an indicator variable set equal to one for those predicted to receive greater than $100 in EITC income. February and Uninsured indicators included but not reported.

Chapter 3

Cashing In and Staying Covered? The Earned Income Tax Credit, Liquidity, and Health Insurance Coverage

3.1 Introduction

With the cost of private health insurance coverage rising, many low-income individuals, even those who work for employers who contribute towards coverage, are becoming priced out of coverage. According to a study in Agency for Health Care Research and Quality (2004), employee contributions increased on average by 56 to 65% from 1996 to 2002. While a rising number of firms are offering insurance plans, the fraction of workers who qualify is falling, and the number enrolled is falling further still due to the increasing cost burden. Cutler (2003) suggests that nearly all of the decline in take-up of private insurance – those who do not enroll when employers offer – is due to the rising share of cost born by employees. Gruber and McKnight (2003) show that from 1982 to 1996, the number of employers fully covering health insurance premiums falls from 44% to only 28%. Rising Medicaid eligibility, rising medical costs, and falling tax rates are found to explain a large fraction of this change.

Low-income workers are likely to be a group most affected by these changes. Chan and Gruber (2010) study premium subsidies in the context of the Massachusetts Insurance Experiment and find an implied elasticity of private plan take-up with respect to price among low-income workers of about -0.65, so that a 1% increase in price leads to a 0.65% decrease in take-up. As prices rise, low income workers are more likely to shed private coverage, and, if eligible, take-up public coverage.

The Earned Income Tax Credit (EITC), a program targeted towards low income workers, may help mitigate this private coverage decline, and massive future increase in crowd-out due to public insurance expansions, by increasing the disposable, after-tax income available to low-income families for the purchase and maintenance of private coverage. Furthermore, the timing of the EITC payments, which are large and frequently received as a lump sum around the month of February, presents an interesting opportunity to study how cash liquidity, the effect of having large amounts of assets in cash or easily convertible to cash, effects the choice of health insurance plan.

I am not the first to study the effect of the EITC on health insurance or health. Hoynes et al. (2012) found positive effects on low birth weight and infant mortality among low-income women when the EITC was expanded in 1993. Hoynes et al. (2012) show a broad, positive effect on health due to the EITC, but are unable to unpack the effects and ignore the intra-year liquidity effects induced by the lump-sum nature of the payment.

LaLumia (2013) looks at within-year liquidity in the context of unemployment and finds that unemployment spells which begin in February are longer than those that begin at other times of the year, among those most likely to be eligible for the EITC.

In this paper, I use variation in the timing of EITC receipt to estimate the effect of cash liquidity on maintaining, and not dropping, private insurance coverage among the EITC population. I contribute to the literature by showing that liquidity matters for insurance coverage as families that receive large EITC payments are less likely to drop private coverage in the months following EITC receipt. An additional $\$1,000$ of EITC income increases the probability of maintaining private coverage on average by $1.4 - 1.9$ percentage points. I also use an instrumental variables method developed in Dahl and Lochner (2012) to estimate the effect of increasing EITC generosity on coverage. I find, surprisingly, that increasing EITC generosity decreases private coverage and increases Medicaid coverage and discuss possible reasons for this result in the conclusion.

Section 2 describes the necessary background on the EITC, its evolution over time, as well as insurance coverage patterns of the EITC eligible, Section 3 describes the data from the SIPP, Section 4 explains the identification strategy, Section 5 review the results, and Section 6 concludes.

3.2 Background

3.2.1 Earned Income Tax Credit

The Earned Income Tax Credit (EITC) is a large, refundable tax credit provided to the working poor in the United States. The size of the credit depends on earnings and the number of dependent children present in the household. In 2007 for a single parent with two children, the EITC provided a 40% subsidy to wages up to $11,790, then was phased out with an additional tax of 21.06% from $15,390 to $37,783, for a maximum credit of $4,716.[1] See Figure (3.1) for a graphical representation and comparison to families with one or no children.

Tax filers have two options for claiming the EITC. The first option is receive the credit, net of other taxes owed, as a lump sum payment upon filing taxes for the previous year. The second option is to forecast earnings for the upcoming year and request EITC be paid out in advance with each pay check. As documented in LaLumia (2013), in practice, the vast majority of tax filers who qualify for the EITC choose the first option, file in January and receive their lump sum payment in February.[2] Some EITC recipients file slightly later and, while the peak of receipt is in February, many also receive payment in March. Furthermore, with the proliferation of businesses providing cash advances on tax returns for a sizable fee, some EITC recipients may receive their money in January. Both are potential sources of measurement error in the actual timing of EITC receipt.

3.2.2 Changes in the EITC, 1990-2006

The EITC has grown immensely throughout the late twentieth and early twenty-first centuries. From tax year 1990 to tax year 2005, the real

[1] All values reported in 2007 U.S. dollars.

[2] Receipt of the Child Tax Credit follows a similar pattern, but it is not investigated in this paper.

dollar amount (in 2012 dollars) of the maximum credit has grown more than three-fold, from about \$1600 to about \$5200. The largest changes at the federal level occured as part of the Omnibus Budget Reconciliation Act (OBRA) of 1990 and OBRA 1993. Smaller changes continued at the federal level, including a 2002 change in the earnings level at which the credit started to phase out, extended this region by $\$1,000$ for married couples filing jointly, and another change in 2005 that extended this region to $\$2,000$.

In addition to federal changes, states begin to offer their own State Earned Income Tax Credits, which most commonly added a fixed percentage to federal EITC checks. Many states offer additional EITC payments of $5-10\%$ of the federal credit, with some states such as New York (30%) and Minnesota (up to 45%) offering considerably more. In all, 25 states and the District of Columbia offer EITC on top of federal credits.

3.2.3 Insurance Status Among EITC Recipients

Among children who live in families receiving EITC payments, many are eligible for Medicaid, the jointly-funded federal and state public insurance program for those with low income, but eligibility varies across states and over time.[3] In 1996, the minimum national standard for eligibility for pregnant women and children under thirteen was 133% of the Federal Poverty Line (FPL), or $\$22,893.33$ in 2007 dollars, covering most but not all of those eligible for EITC. Some states, such as Hawaii, expanded eligibility to state-run programs for all citizens up to 300% FPL, well past the range of incomes eligible for the EITC. Children thirteen and older, as well as non-pregnant adults, were only eligible for public insurance coverage if family income fell below the poverty thresholds for cash payments under pre-1996 Aid to Families with Dependent Children (AFDC), commonly referred to as welfare. Some states, such as Alabama, set this eligibility threshold at 13.4% FPL, or about $\$2,300$

[3] Medicaid also covers those with exceptionally high medical costs relative to income.

in annual earnings.

In 2007, with the expansion of the Children's Health Insurance Program (CHIP), most children under the age of nineteen in families earning up to 200% FPL, or $34,340, were eligible for free or highly subsidized public insurance, although children under nineteen living in North Dakota were only eligible up to 140% FPL. Adults with children in many states were still only eligible for public insurance at the very low pre-1996 AFDC levels, but some states began to expand insurance for families with higher incomes and childless adults, most notably the Massachusetts insurance expansion in 2006 which provides free public insurance to all residents up to 150% FPL and generous subsidies up to 300% FPL.

Medicaid and CHIP insurance, while reimbursing doctors and hospitals at low rates, requires very low co-payments or coinsurance rates for patients.[4]

3.3 Data

The data come from the 1990, 1991, 1992, 1993, 1996, 2001, and 2004 panels of the Survey of Income and Program Participation (SIPP) and in the final sample includes 603,705 individuals. All panels are monthly and cover a period of 28 to 48 months. The SIPP contains a wealth of very detailed income and tax variables, along with data on the source of health insurance coverage and family structure at the monthly level. For those who take-up private coverage from an employer, I have data on whether the employer contributed towards some or all of the premium, but information on the dollar-value of the contribution of employer and employee is unfortunately unavailable.

Families in the SIPP are constructed using the family and subfamily identifiers. If one of the primary earners enters or leaves during the time the family is in the sample, all observations from all individuals within the family

[4]It is important to note that services covered and the availability of doctors seeing Medicaid/CHIP patients varies from state to state.

are dropped. If children enter the family, they take on the family identifier of the primary earner(s). If children leave the sample, these month observations are dropped from the sample.

The SIPP provides data at a monthly level, but data is only collected every four months in a "wave." As such, data for months one, two, or three months prior to the interview month are reported with a high probability of error. Variable outcomes across months within a wave tend to be smoothed which exaggerates changes across waves. To avoid biasing transitions across states of insurance towards zero, I drop all non-seam months.[5]

Individuals are dropped from the sample if they do not appear for at least 24 months. Individuals who are 60 years or older, or who ever have Medicare or military insurance are dropped, leaving only those with private insurance, Medicaid, and the uninsured. Families are dropped if the composition of primary earners changes within the sample, e.g. families form or dissolve. Furthermore, I drop those who earn more than $\$100,000$ in 2012 dollars in their first year of the sample, in order to minimize the effects of outliers. Unlike others, I do not drop individuals who earn more than $\$100,000$ in subsequent years as this would lead us to select only high earners who for some reason see their income decline in each subsequent year, which may be correlated with health and thus the demand for health insurance. Individuals in states which cannot be individually identified are also dropped.

Additional data on Medicaid eligibility is constructed from press releases from the National Governor's Association, beginning in 1990, which highlight changes in federal and state Medicaid policy. A dataset containing the evolution of Medicaid eligibility for each state, from 1990-2007, is available from the author by request. Data on the Federal Poverty Limit for each of the 48 continental states, as well as Alaska and Hawaii, for each year from 1990-2007 comes from the Office of the Assistant Secretary for Planning and

[5]Indeed, I have run regressions where all months are included and the coefficient estimates for EITC effects in Estimation Strategy #1 are two orders of magnitude smaller.

Evaluation.

3.4 Estimation Strategy

I use two main identification strategies, as explored in LaLumia (2013), Chetty et al. (2011), and Dahl and Lochner (2012) to look at (1) the effects of intra-year liquidity on insurance coverage and (2) the effects of increasing EITC generosity on insurance coverage take-up.

3.4.1 Strategy 1: Exploiting the Timing of EITC Receipt within a Year

The first strategy exploits the timing of the EITC payment within a given year, tracing out the effect of EITC payment received in year t over each month within the year. I look at transitions from each of three insurance states – (1) private, (2) Medicaid, and (3) uninsured – from the previous September to the current wave in the Survey of Income and Program Participation (SIPP).[6] I focus only on "seam" months in the SIPP, which are those months which correspond to the interview month when individuals are more likely to accurately recall insurance status. For a table of baseline transition probabilities for all sample years, see Table (3.2). The hypothesis is that in months where family financial liquidity high, near the timing of EITC receipt, families will be less likely to transition out of private insurance if currently enrolled, and more likely to transition into private insurance if enrolled in Medicaid or uninsured. Testing the second part of the hypothesis is made more difficult by the unknown timing of individual firm's "open enrollment periods" which are a limited number of weeks or months when those without private coverage can choose to enroll, which will likely bias downward my estimates of the transition into private insurance coverage.

[6] September of the previous year is chosen as it is a fixed point and is guaranteed to be in a different interview wave.

$$HI_{ifmt} = \sum_{m=1}^{11} \gamma_m EITC_{ft}(P_{f(t-1)}) \times \{month = m\} + EITC_{ft}(P_{f(t-1)}) \\ + \\ sum_{m=1}^{11} nu_m \{month = m\} + \Phi(P_{f(t-1)}) + X'_{ifmt}\beta + \varepsilon_{ifmt} \tag{3.1}$$

where the notation is as follows: individual i, family f, month m, year t. Key coefficient is on interaction of EITC receipt and a month dummy for February, the "February-EITC effect". In all regressions, I control flexibly for lagged income using a fifth-order polynomial and an indicator for positive lagged income. The inclusion of the flexible control function means identification comes solely through the nonlinearities in the EITC schedule.

3.4.2 Strategy 2: Changes In EITC Payments across Years

The second strategy uses a control function approach to isolate the change in after-tax income across years due to changes in the tax code, separating the changes in income due to the natural evolution of income. Income is mean-reverting and the EITC benefits those with low-income, which makes it extremely important to use a flexible control function to adequately account for this. The control function is again a fifth-order polynomial and an indicator for positive lagged income. I begin with the following equation, where Inc_{it} represents after-tax income:

$$HI_{it} = \gamma Inc_{it} + X'_{it}\beta + \mu_i + \varepsilon_{it} \tag{3.2}$$

Which, in differences, becomes:

$$\Delta HI_{it} = \gamma \Delta Inc_{it} + \Delta X'_{it}\beta + \Delta\varepsilon_{it} \tag{3.3}$$

I can then decompose after-tax income into pre-tax income and taxes:

$$\Delta Inc_{it} = [Inc_{it} - Inc_{i,t-1}] = [P_{it} - T^t(P_{it})] - [P_{i,t-1} - T^{t-1}(P_{i,t-1})] \tag{3.4}$$

Using the method from Dahl and Lochner (2012), I instrument for change in income using only change through the tax code:

$$\Delta Inc_{it}^{IV} = -[T^t(E[P_{it}|P_{i,t-1}]) - T^{t-1}(P_{i,t-1})] \tag{3.5}$$

Including the contra function, I have the first stage regression:

$$\Delta Inc_{it} = -[T^t(E[P_{it}|P_{i,t-1}]) - T^{t-1}(P_{i,t-1})] + \Phi(P_{f(t-1)}) + \chi_{it} \tag{3.6}$$

and a second stage regression:

$$\Delta HI_{it} = \Delta Inc_{it}^{IV} + \Phi(P_{f(t-1)}) + \Delta X_{it}^{'}\beta + \Delta\varepsilon_{it} \tag{3.7}$$

3.5 Results

3.5.1 OLS

First, I examine a simple OLS regression with insurance status as the dependent variable and monthly family income, employment status and Medicaid eligibility statues, along with many controls, as independent variables.[7] I also control for year, state of residence, and the full interaction of year and state. In Table (3.3), the correlation between family monthly income and insurance status moves in the predicted direction and is highly significant. A $\$1,000$ increase in monthly income is associated with a 1.01 percentage point increase in the likelihood of having private health insurance. The probability of having Medicaid, conditional on eligibility, as well as the probability of being uninsured are decreasing in income.

3.5.2 Within-Year Liquidity

Next, turning our attention to the results from the estimation for intra-year liquidity, I estimate Equation (3.1), including covariate controls, and fixed

[7] I control for individual Medicaid eligibility status and employment status as well as the statuses of the primary earners in the household. Additional covariate controls include family size, marital status, race, gender, a cubic for age, dummies for firm premium contributions.

effects for year, state, and state-year. I estimate this regression separately for each lagged insurance status of the individual in order to report the probability of transitioning from one state to another, with liquidity, as proxied by the EITC-month interactions, as the main explanatory variable for whether or not a transition occurs.

In Table (3.4), I report the results for those individuals who reported private health insurance coverage during the interview month in the previous wave. Column 1 is a linear regression where the outcome variable is an indicator equal to one if the individual maintains private coverage and zero if the individual drops private coverage in that month. The first row shows a negative coefficient on the level of EITC income, indicating that those who receive large amounts of EITC income are more likely to transition out of private insurance coverage, conditional on all other covariates, including a flexible control function for lagged income. Note that those receiving a lower amount of EITC income may earn more or less than those with higher EITC income as the credit is phased in and then phased out. Looking at the EITC-month interactions, a pattern emerges where EITC recipients are more likely to maintain private coverage during the months of March and May, the months following the likely timing of the EITC-induced liquidity (statistically significant at the 1% or 0.1% level) . The EITC-January, EITC-February, and EITC-June interaction also have positive coefficients which are statistically significantly different from the EITC-December effect at the 5% level (December is the omitted month). In the same months, the probability that an EITC recipient transitions to Medicaid or uninsured decreases, but the effects are less precisely estimated and only a handful of months are statistically significantly different from December at the 1% or 5% level.

In Tables (3.5)-(3.6), I look at the liquidity effects of the transition probabilities for those who had Medicaid or were uninsured last period. Coefficients are small, but somewhat precisely estimated, and little information is gained by examining these tables. Given that individuals without private insurance are only able to enroll during "open enrollment periods," which typically occur only once per year at different months which are not guaranteed to overlap with peak liquidity, it is perhaps somewhat unsurprising that there is little to no transition into private insurance or between Medicaid and uninsured. Finally, in Table (3.7), I estimate the unconditional probability that an individual has one of the three types of insurance, conditional on the level of EITC receipt which is allowed to vary in its effect throughout the year. The general pattern is similar to that observed in Table (3.4), but many of the coefficients are imprecisely measured.

3.5.3 Across-Year, After-Tax Income

In Estimation Strategy #2, I instrument for the change in income, using bureaucratic changes in the EITC program, as well as other tax changes, at the federal and state level. This provides plausibly exogenous variation in income across years to identify the effect of changing the size of the tax credit over time, as opposed to looking at the effect of the timing of receipt within a given year.

In Table (3.8) I see that there is a strong, nearly one-to-one relationship between the predicted change in tax liability, in row (1), and the change in after tax income, controlling flexibly for lagged income. F-statistics in each specification are all well above 10, the cutoff for weak instruments. Each row corresponds to a different initial state, taken to be the insurance status in March of the year t when the first EITC payment is received. Each regression includes covariate controls as well as state, year, and state-year fixed effects, as is the case in all second stage regressions. The coefficient on Δ Net Tax Credits is not sensitive to the exclusion of any controls.

In Tables (3.9)-(3.11), I estimate Equation (3.7) for each lagged insurance status. The results for such regressions are puzzling in light of the results presented in Tables (3.4)-(3.7) which indicate cash liquidity increases the probability of having and maintaining private insurance coverage. In (3.9), it appears that increases in net tax credits actually increase the likelihood of enrolling in Medicaid and do nothing to preserve private coverage. In Table (3.11), for the previously uninsured, they are more likely to transition to Medicaid when after-tax income increases due to tax changes. Possible explanations are explored in the conclusion.

3.6 Conclusion

Cash liquidity, induced by a lump-sum EITC payment which most generally occurs around February, has a positive effect on maintaining private health insurance coverage for a two month window following receipt. Individuals in families that receive larger EITC payments are more likely to keep private coverage and less likely to transition to no insurance or Medicaid. When I examine the effect of increasing annual after-tax income through changes in tax policy, however, I find that increased EITC receipt causes an increase in the transition from private insurance to Medicaid.

Why is the year-over-year effect negative? Several explanations exist. First, while I control explicitly for state policy changes at the monthly level to control for Medicaid eligibility at the individual and family level, it is possible that there are errors in this dataset, or the timing of key reforms and the take-up of public coverage is different in practice than it is in legislation. It could also be that those who have private insurance and receive large amounts of EITC income are precisely those with the most tenuous link to private coverage. They are also likely to be those on the margin of private coverage, so that when prices rise and employer premium contributions do not rise enough to match, it is exactly these individuals who drop coverage.

In future work, I hope to better understand the temporal nature of liquidity to better forecast EITC policy counterfactuals which change the timing of the distribution of funds. Currently, very few individuals receive the

EITC in each paycheck for the year before taxes are filed. It is unclear from this work and other research as to what the effect would be of encouraging this type of distribution of the EITC or any other alternative division of the lump-sum payment. It is possible that the timing and size of the payments matters greatly for how they are spent. It is conceivable that many smaller payments would lead to an increase in the use of EITC funds for day-to-day consumption and a decrease in the use of funds for larger purchases like health insurance premiums or durables goods.

3.7 Figures

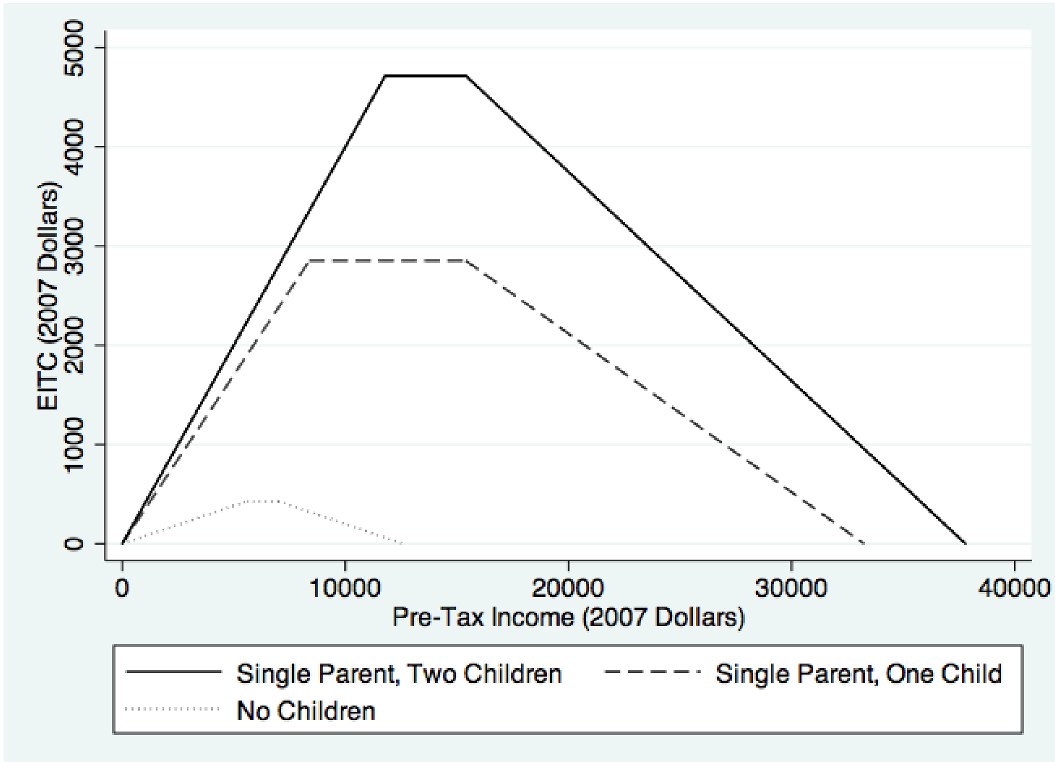

Figure 3.1: EITC Program Parameters, 2007

3.8 Tables

Table 3.1: Summary Statistics

	Full Sample	EITC> 0	EITC= 0
EITC ($2012)	458.549	1,544.017	0
	(951.432)	(1,171.354)	(0)
Monthly Family Income	2,856.363	1,515.452	3,422.822
	(1834.227)	(942.3926)	(1824.838)
Annual Family Income	34,719.105	18,821.468	41,434.961
	(21,319.521)	(11,445.812)	(20,959.969)
Private HI	0.735	0.494	0.837
	(0.441)	(0.5)	(0.369)
Medicaid	0.107	0.213	0.063
	(0.309)	(0.409)	(0.242)
Uninsured	0.157	0.293	0.1
	(0.364)	(0.455)	(0.3)
Medicaid Eligible (FPL-based)	0.152	0.332	0.076
	(0.359)	(0.471)	(0.264)
Firm Pays Part of HI Premium	0.452	0.294	0.518
	(0.498)	(0.456)	(0.5)
Firm Pays All of HI Premium	0.209	0.11	0.25
	(0.406)	(0.313)	(0.433)
Age	26.278	22.415	27.91
	(16.263)	(15.753)	(16.2)
Married	0.371	0.286	0.408
	(0.483)	(0.452)	(0.491)
Family Size	3.654	3.748	3.614
	(1.815)	(1.605)	(1.895)
Male	0.481	0.442	0.498
	(0.5)	(0.497)	(0.5)
Black	0.149	0.221	0.118
	(0.356)	(0.415)	(0.322)
Asian	0.032	0.038	0.029
	(0.175)	(0.191)	(0.169)
Native American	0.014	0.02	0.011
	(0.116)	(0.139)	(0.104)
Year	1997.14	1997.95	1996.79
	(4.874)	(4.711)	(4.901)
N	603,705	179,291	424,414

Table 3.2: Transition Matrices

All			
	Private_{t+1}	Medicaid_{t+1}	Uninsured_{t+1}
Private_t	0.96	0.01	0.03
Medicaid_t	0.07	0.83	0.10
Uninsured_t	0.14	0.06	0.80

EITC>0			
	Private_{t+1}	Medicaid_{t+1}	Uninsured_{t+1}
Private_t	0.90	0.03	0.07
Medicaid_t	0.08	0.80	0.12
Uninsured_t	0.12	0.08	0.80

EITC=0			
	Private_{t+1}	Medicaid_{t+1}	Uninsured_{t+1}
Private_t	0.98	0.00	0.02
Medicaid_t	0.05	0.88	0.07
Uninsured_t	0.17	0.04	0.79

Wave-to-wave transition probability matrix for each of the three insurance states. Each entry represents the baseline probability of transitioning to row j in period $t+1$ given that the individual was in state i in period t where i represents rows and j columns. The first table is for the entire sample, the second table is for those with positive EITC receipt, and the third table is for those with no EITC receipt.

Table 3.3: OLS Regressions

	(1) Private HI	(2) Medicaid	(3) Uninsured
Monthly Family Income ($1,000s)	0.0101***	-0.00109**	-0.00904***
	(0.000516)	(0.000386)	(0.000514)
Age	-0.000521	-0.00822***	0.00874***
	(0.000414)	(0.000520)	(0.000519)
Age Squared	0.00000599	0.000169***	-0.000175***
	(0.0000168)	(0.0000190)	(0.0000204)
Age Cubed	0.000000128	-0.00000105***	0.000000920***
	(0.000000203)	(0.000000213)	(0.000000239)
Male	-0.00371***	-0.0216***	0.0253***
	(0.00103)	(0.00117)	(0.00129)
Medicaid Eligible (FPL-based)	0.0137***	0.0732***	-0.0869***
	(0.00281)	(0.00399)	(0.00394)
Firm Pays Part of HI Premium	0.716***	-0.188***	-0.528***
	(0.00371)	(0.00257)	(0.00371)
Firm Pays All of HI Premium	0.713***	-0.184***	-0.529***
	(0.00376)	(0.00262)	(0.00376)
Family Size	-0.00168**	0.00683***	-0.00515***
	(0.000634)	(0.000604)	(0.000705)
State Fixed Effects	Yes	Yes	Yes
Year Fixed Effects	Yes	Yes	Yes
State×Year Fixed Effects	Yes	Yes	Yes
Observations	525076	525076	525076

Robust standard errors, clustered at the family level, in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Linear probability of an individual having (1) private health insurance, (2) Medicaid, or (3) being uninsured. Family monthly income measured in thousands of 2012 U.S. dollars. Data comes from 1990-2004 SIPP panels. Only the last month of the interview wave ("seam" month) is included. Race and marital status included but not reported. For more information, see text.

Table 3.4: EITC Liquidity Regressions, Privately Insured

	(1) Private HI	(2) Medicaid	(3) Uninsured
EITC ($2012)	-0.0157**	0.00443	0.0112**
	(0.00478)	(0.00313)	(0.00390)
EITC × Jan.	0.0159**	-0.00497	-0.0109*
	(0.00571)	(0.00356)	(0.00464)
EITC × Feb.	0.0142*	-0.00199	-0.0122**
	(0.00585)	(0.00384)	(0.00469)
EITC × Mar.	0.0189***	-0.00412	-0.0148***
	(0.00554)	(0.00364)	(0.00440)
EITC × Apr.	0.0145**	-0.00609	-0.00844*
	(0.00498)	(0.00316)	(0.00422)
EITC × May	0.00831	0.00144	-0.00976*
	(0.00598)	(0.00398)	(0.00470)
EITC × Jun.	0.00855	0.00135	-0.00990*
	(0.00608)	(0.00401)	(0.00477)
EITC × Jul.	0.00935	0.00244	-0.0118**
	(0.00595)	(0.00405)	(0.00456)
EITC × Aug.	0.00534	0.000275	-0.00562
	(0.00443)	(0.00316)	(0.00370)
EITC × Sep.	0.00477	0.00138	-0.00615
	(0.00636)	(0.00399)	(0.00510)
EITC × Oct.	0.00631	0.00156	-0.00787
	(0.00634)	(0.00420)	(0.00489)
EITC × Nov.	0.00186	0.000679	-0.00254
	(0.00661)	(0.00397)	(0.00541)
Observations	201592	201592	201592

Robust standard errors, clustered at the family level, in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Linear probability of an individual having (1) private health insurance, (2) Medicaid, or (3) being uninsured, conditional on being privately insured during the last reference month of the previous SIPP interview wave. Covariate controls and state, year, and state-by-year fixed effects included. Income measured in thousands of 2012 U.S. dollars. Data comes from 1990-2004 SIPP panels. Only the last month of the interview wave ("seam" month) is included. For more information, see text.

Table 3.5: EITC Liquidity Regressions, Medicaid Insured

	(1) Private HI	(2) Medicaid	(3) Uninsured
EITC ($2012)	0.00398	-0.00478	0.000803
	(0.00568)	(0.00808)	(0.00679)
EITC × Jan.	-0.00591	0.0160	-0.0101
	(0.00666)	(0.00957)	(0.00777)
EITC × Feb.	-0.00672	0.0140	-0.00728
	(0.00656)	(0.00958)	(0.00789)
EITC × Mar.	-0.00602	0.00263	0.00339
	(0.00662)	(0.0102)	(0.00872)
EITC × Apr.	-0.0143*	0.0199*	-0.00564
	(0.00590)	(0.00855)	(0.00694)
EITC × May	0.00161	-0.000607	-0.000999
	(0.00707)	(0.00983)	(0.00798)
EITC × Jun.	-0.00630	0.00563	0.000671
	(0.00675)	(0.00989)	(0.00836)
EITC × Jul.	-0.00275	-0.00161	0.00436
	(0.00695)	(0.0103)	(0.00866)
EITC × Aug.	-0.00862	0.0139	-0.00527
	(0.00547)	(0.00815)	(0.00679)
EITC × Sep.	0.00173	0.00160	-0.00332
	(0.00797)	(0.0108)	(0.00879)
EITC × Oct.	-0.00863	0.0171	-0.00842
	(0.00734)	(0.0108)	(0.00897)
EITC × Nov.	-0.0113	0.00943	0.00183
	(0.00749)	(0.0110)	(0.00928)
Observations	27330	27330	27330

Robust standard errors, clustered at the family level, in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Linear probability of an individual having (1) private health insurance, (2) Medicaid, or (3) being uninsured, conditional on having Medicaid during the last reference month of the previous SIPP interview wave. Covariate controls and state, year, and state-by-year fixed effects included. Income measured in thousands of 2012 U.S. dollars. Data comes from 1990-2004 SIPP panels. Only the last month of the interview wave ("seam" month) is included. For more information, see text.

Table 3.6: EITC Liquidity Regressions, Uninsured

	(1) Private HI	(2) Medicaid	(3) Uninsured
EITC ($2012)	-0.00836	0.00366	0.00470
	(0.00576)	(0.00465)	(0.00684)
EITC × Jan.	0.0144	-0.00932	-0.00509
	(0.00749)	(0.00555)	(0.00863)
EITC × Feb.	0.0109	-0.00625	-0.00462
	(0.00687)	(0.00558)	(0.00834)
EITC × Mar.	0.00456	-0.00398	-0.000577
	(0.00696)	(0.00566)	(0.00839)
EITC × Apr.	0.00917	-0.00324	-0.00593
	(0.00580)	(0.00517)	(0.00710)
EITC × May	0.0131	-0.00722	-0.00584
	(0.00775)	(0.00598)	(0.00888)
EITC × Jun.	0.00280	0.0000726	-0.00287
	(0.00700)	(0.00608)	(0.00851)
EITC × Jul.	0.00964	-0.00116	-0.00848
	(0.00745)	(0.00600)	(0.00899)
EITC × Aug.	-0.000419	-0.00139	0.00181
	(0.00537)	(0.00392)	(0.00626)
EITC × Sep.	0.00462	-0.00762	0.00300
	(0.00821)	(0.00620)	(0.00942)
EITC × Oct.	-0.000796	-0.00352	0.00432
	(0.00768)	(0.00615)	(0.00903)
EITC × Nov.	0.00255	0.00256	-0.00511
	(0.00803)	(0.00668)	(0.00963)
Observations	44838	44838	44838

Robust standard errors, clustered at the family level, in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Linear probability of an individual having (1) private health insurance, (2) Medicaid, or (3) being uninsured, conditional on being uninsured during the last reference month of the previous SIPP interview wave.Covariate controls and state, year, and state-by-year fixed effects included. Income measured in thousands of 2012 U.S. dollars. Data comes from 1990-2004 SIPP panels. Only the last month of the interview wave ("seam" month) is included. For more information, see text.

Table 3.7: EITC Liquidity Regressions

	(1) Private HI	(2) Medicaid	(3) Uninsured
EITC ($2012)	-0.0349***	0.0263***	0.00854*
	(0.00339)	(0.00369)	(0.00383)
EITC × Jan.	0.00788	-0.000761	-0.00712
	(0.00505)	(0.00544)	(0.00549)
EITC × Feb.	0.00387	0.00502	-0.00889
	(0.00479)	(0.00546)	(0.00549)
EITC × Mar.	0.00505	-0.00355	-0.00149
	(0.00485)	(0.00538)	(0.00558)
EITC × Apr.	0.00227	0.00503	-0.00730
	(0.00368)	(0.00339)	(0.00377)
EITC × May	0.00903*	0.000351	-0.00938
	(0.00453)	(0.00506)	(0.00496)
EITC × Jun.	0.00276	0.00407	-0.00683
	(0.00445)	(0.00512)	(0.00503)
EITC × Jul.	0.000541	0.00373	-0.00427
	(0.00433)	(0.00496)	(0.00497)
EITC × Aug.	0.00127	0.00403	-0.00529
	(0.00296)	(0.00268)	(0.00303)
EITC × Sep.	0.00730	-0.000478	-0.00682
	(0.00475)	(0.00518)	(0.00522)
EITC × Oct.	0.00318	-0.0000650	-0.00312
	(0.00460)	(0.00518)	(0.00521)
EITC × Nov.	-0.000877	0.00323	-0.00235
	(0.00441)	(0.00509)	(0.00510)
Observations	417397	417397	417397

Robust standard errors, clustered at the family level, in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Linear probability of an individual having (1) private health insurance, (2) Medicaid, or (3) being uninsured during the last reference month of the previous SIPP interview wave. Covariate controls and state, year, and state-by-year fixed effects included. Sample inclusion does not depend on previous insurance status. Income measured in thousands of 2012 U.S. dollars. Data comes from 1990-2004 SIPP panels. Only the last month of the interview wave ("seam" month) is included. For more information, see text.

Table 3.8: First Stage Regression

	Δ Family After-Tax Annual Income			
	(1)	(2)	(3)	(4)
	if Lag Priv = 1	if Lag Medicaid = 1	if Lag Unins = 1	All
Δ Net Tax Credits	1.083***	1.280***	1.460***	1.357***
	(0.155)	(0.251)	(0.217)	(0.114)
Positive Lag Wages	3063.0***	1029.2**	1606.1**	1353.1***
	(609.5)	(342.1)	(532.5)	(259.2)
Lag Wages	-0.294***	-0.306**	-0.157	-0.0125
	(0.0810)	(0.0998)	(0.108)	(0.0491)
Lag Wages^2	0.00000506	0.0000206*	0.00000252	-0.00000514
	(0.00000424)	(0.00000949)	(0.00000780)	(0.00000283)
Lag Wages^3	-5.75e-11	-7.94e-10*	-1.30e-10	1.22e-10
	(9.55e-11)	(3.36e-10)	(2.33e-10)	(6.77e-11)
Lag Wages^4	3.36e-16	1.16e-14*	2.00e-15	-1.18e-15
	(9.61e-16)	(4.81e-15)	(2.97e-15)	(7.06e-16)
Lag Wages^5	-9.59e-22	-5.67e-20*	-9.47e-21	3.91e-21
	(3.54e-21)	(2.34e-20)	(1.34e-20)	(2.66e-21)
Covariate Controls	Yes	Yes	Yes	Yes
State Fixed Effects	Yes	Yes	Yes	Yes
Year Fixed Effects	Yes	Yes	Yes	Yes
State×Year Fixed Effects	Yes	Yes	Yes	Yes
Observations	93314	15787	21442	135241

Robust standard errors, clustered at the family level, in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

First stage regression to estimate predicted change in after-tax annual family income using the change in predicted net tax credits as an instrument. All regressions include a control function for lagged income among other controls. For more information, see text.

Table 3.9: Second Stage Regression, Lag Private = 1

	(1) Keep Private	(2) Switch to Medicaid	(3) Switch to Uninsured
Instrumented Δ Family After-Tax Income ($1,000s)	-0.00650	0.00962***	-0.00236
	(0.00435)	(0.00285)	(0.00335)
Covariate Controls	Yes	Yes	Yes
State Fixed Effects	Yes	Yes	Yes
Year Fixed Effects	Yes	Yes	Yes
State×Year Fixed Effects	Yes	Yes	Yes
Observations	93314	93314	93314

Robust standard errors, clustered at the family level, in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$
Second stage regression for the probability of transitioning from private health insurance last year, to (1) retaining private health insurance, (2) switching to Medicaid, or (3) switching to uninsured.

Table 3.10: Second Stage Regression, Lag Medicaid = 1

	(1) Switch to Private	(2) Keep Medicaid	(3) Switch to Uninsured
Instrumented Δ Family After-Tax Income ($1,000s)	0.00698	-0.0169	0.0112
	(0.00801)	(0.0104)	(0.00845)
Covariate Controls	Yes	Yes	Yes
State Fixed Effects	Yes	Yes	Yes
Year Fixed Effects	Yes	Yes	Yes
State×Year Fixed Effects	Yes	Yes	Yes
Observations	93314	93314	93314

Robust standard errors, clustered at the family level, in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$
Second stage regression for the probability of transitioning from private health insurance last year, to (1) switching to private health insurance, (2) retaining Medicaid, or (3) switching to uninsured.

Table 3.11: Second Stage Regression, Lag Uninsured = 1

	(1) Switch to Private	(2) Switch to Medicaid	(3) Remain Uninsured
Instrumented Δ Family After-Tax Income ($1,000s)	0.00439	0.0179**	-0.0219**
	(0.00627)	(0.00573)	(0.00769)
Covariate Controls	Yes	Yes	Yes
State Fixed Effects	Yes	Yes	Yes
Year Fixed Effects	Yes	Yes	Yes
State×Year Fixed Effects	Yes	Yes	Yes
Observations	93314	93314	93314

Robust standard errors, clustered at the family level, in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Second stage regression for the probability of transitioning from private health insurance last year, to (1) switching to private health insurance, (2) switching to Medicaid, or (3) remaining uninsured.

Bibliography

Acemoglu, Daron, Amy Finkelstein, and Matthew Notowidigdo, "Income and Health Spending: Evidence from Oil Price Shocks," *Review of Economics and Statistics, Forthcoming.*

Adams, William, Liran Einav, and Jonathan Levin, "Liquidity Constraints and Imperfect Information in Subprime Lending," *American Economic Review*, 2009, *99* (1), 49–84.

Agency for Health Care Research and Quality, "Employer-Sponsored Health Insurance: Trends in Cost and Access," *AHRQ Research In Action*, 2004, *17* (1), 1–12.

Agency for Healthcare Research and Quality, "Guide to Prevention Quality Indicators: Hospital Admission for Ambulatory Care Sensitive Conditions," www.ahrq.gov/downloads/pub/ahrqqi/pqiguide.pdf March 2007.

Ai, Chunrong and Edward C. Norton, "Interaction Terms in Logit and Probit Models," *Economics Letters*, 2003, *80*, 123–129.

Amarante, Veronica, Marco Manacorda, Edward Miguel, and Andrea Vigorito, "Do Cash Transfers Improve Birth Outcomes? Evidence from Matched Vital Statistics, Program and Social Security Data," *NBER Working Paper Series*, 2011, *WP 17690.*

Anderson, Michael, Carlos Dobkin, and Tal Gross, "The Effect of Health Insurance Coverage on the Use of Medical Services," *American Economic Journal: Economic Policy*, 2012, *4* (1), 1–27.

Card, David, Carlos Dobkin, and Nicole Maestas, "The Impact of Nearly Universal Insurance Coverage on Health Care Utilization: Evidence from Medicare," *American Economic Review*, 2008, *98* (5), 2242–52.

Chan, David and Jonathan Gruber, "How Sensitive are Low Income Families to Health Plan Prices," *American Economic Review: Papers & Proceedings*, 2010, *100*, 292–6.

Chetty, Raj, "Moral Hazard versus Liquidity and Optimal Unemployment Insurance," *Journal of Political Economy*, 2008, *116* (2), 173–234.

_ **, John N. Friedman, and Jonah Rockoff**, "New Evidence on the Long-Term Impacts of Tax Credits," *Statistics of Income Paper Series*, 2011.

Currie, Janet and Jonathan Gruber, "Health Insurance Eligibility, Utilization of Medical Care, and Child Health," *The Quarterly Journal of Economics*, 1996, *111* (2), 431–466.

_ **and** _ , "Saving Babies: The Efficacy and Cost of Recent Changes in the Medicaid Eligibility of Pregnant Women," *The Journal of Politcal Economy*, 1996, *104* (6), 1263–1296.

Cutler, David, "Employee Costs and the Decline in Health Insurance Coverage," *Forum for Health Economics & Policy*, 2003, *6*, 1–27.

_ **and Jonathan Gruber**, "Does Public Insurance Crowd Out Private Insurance," *Quarterly Journal of Economics*, 1996, *111* (2), 391–430.

Dafny, Leemore and Jonathan Gruber, "Public Insurance and Child Hospitalizations: Access and Efficiency Effects," *Journal of Public Economics*, 2005, *89* (1), 109 – 129.

Dahl, Gordon and Lance Lochner, "The Impact of Family Income on Child Achievement: Evidence from the Earned Income Tax Credit," *American Economic Review*, 2012, *102* (5), 1927–56.

Deaton, Angus, "Policy Implications of the Gradient of Health and Wealth," *Health Affairs*, 2002, *21* (2), 13–30.

Doonan, Michael T. and Katharine R. Tull, "Health Care Reform in Massachusetts: Implementation of Coverage Expansions and a Health Insurance Mandate," *Milbank Quarterly*, 2010, *88* (1), 54–80.

Feenberg, Daniel R. and Elizabeth Coutts, "An Introduction to the TAXSIM Model," *Journal of Policy Analysis and Management*, 1993, *12* (1), 189–194.

Finkelstein, Amy and Kathleen McGarry, "Multiple Dimensions of Private Information: Evidence from the Long-Term Care Insurance Market," *The American Economic Review*, 2006, *96* (4), 938–958.

_ , **Sarah Taubman, Bill Wright, Mira Bernstein, Jonathan Gruber, Joseph P. Newhouse, Heidi Allen, Katherine Baicker, and The Oregon Health Study Group**, "The Oregon Health Insurance Experiment: Evidence from the First Year," *NBER Working Paper Series*, 2011, *WP 17190.*

Freeman, Joseph D., Srikanth Kadiyala, Janice F. Bell, and Diane P Martin, "The Causal Effect of Health Insurance on Utilization and Outcomes in Adults: A Systematic Review of US Studies," *Medical Care*, 2008, *146* (10), 1023–32.

Garcia, Tamyra Carroll, Amy B. Bernstein, and Mary Ann Bush, "Emergency Department Visitors and Visits: Who Used the Emergency Room in 2007," National Center for Health Statistics Data Brief 2010.

Gross, Tal and Jeremy Tobacman, "Dangerous Liquidity and the Demand for Health Care: Evidence from the 2008 Stimulus Payments," *Journal of Human Resources, Forthcoming*, 2013.

Gruber, Jonathan and Robin McKnight, "Why Did Employee Health Insurance Contributions Rise?," *Journal of Health Economics*, 2003, *22*, 1085–1104.

Hahn, Youjin, "The Effect of Medicaid Physician Fees on Take-up of Public Health Insurance among Children in Poverty," *Monash University Department of Economics, Discussion Paper Series 29/12*, 2012.

Hoynes, Hilary W., Douglas L. Miller, and David Simon, "Income, the Earned Income Tax Credit, and Infant Health," *NBER Working Paper*, 2012, (18206).

Kolstad, Jonathan T. and Amanda E. Kowalski, "The Impact of an Individual Health Insurance Mandate On Hospital and Preventive Care: Evidence from Massachusetts," *Journal of Public Economics*, 2012, *Forthcoming.*

LaLumia, Sarah, "The EITC, Tax Refunds, and Unemployment Spells," *American Economic Journal: Economic Policy*, 2013, *5* (2), 188–221.

Levy, Helen and David Meltzer, "The Impact of Health Insurance on Health," *Annual Review of Public Health*, 2008, *29* (1), 399–409.

Long, Sharon K., "On The Road To Universal Coverage: Impacts Of Reform In Massachusetts At One Year," *Health Affairs*, 2008, *27* (4), 270–284.

— **and Karen Stockley, eds**, *Using the National Health Interview Survey to Evaluate State Health Reform: Findings from New York and Massachusetts* Washington, DC 2010.

Mahoney, Neale, "Bankruptcy as Implicit Health Insurance," http://www.people.fas.harvard.edu/~nealemahoney/Research/Mahoney_Bankruptcy.pdf 2011.

Manning, Willard G., Joseph P. Newhouse, Naihua Duan, Emmett B. Keeler, and Arleen Leibowitz, "Health Insurance and the Demand for Medical Care: Evidence from a Randomized Experiment," *The American Economic Review*, 1987, *77* (3), 251–277.
Massachusetts Labor and Workforce Development Website

Massachusetts Labor and Workforce Development Website, *http://lmi2.detma.org/lmi/town202data.asp.*

Miller, Sarah, *"The Effect of Insurance on Emergency Room Visits: An Analysis of the 2006 Massachusetts Health Reform," Journal of Public Economics, 2012, 96 (11), 893–908.*

Newton, Manya F. and et al., *"Uninsured Adults Presenting to US Emergency Departments: Assumptions vs Data," JAMA, 2008, 300 (16), 1914–1924.*

Puhani, Patrick A., *"The Treatment Effect, the Cross Difference, and the Interaction Term in Nonlinear Difference-in-Differences Models," Economics Letters, 2012, 115, 85–87.*

Ruhm, Christopher J., *"Are Recessions Good for Your Health?," The Quarterly Journal of Economics, 2000, 115 (2), 617–650.*

Stowasser, Till, Florian Heiss, Daniel McFadden, and Joachim Winter, *NBER Investigations in the Economics of Aging, Chicago, IL: University of Chicago Press,*

Weissman, Joel S., Constantine Gatsonis, and Arnold M. Epstein, *"Rates of Avoidable Hospitalization by Insurance Status in Massachusetts and Maryland," JAMA, 1992, 268 (17), 2388–2394.*

Yelowitz, Aaron and Michael F. Cannon, *"The Massachusetts Health Plan: Much Pain, Little Gain," The Cato Institute March 2010.*