THE OREGON HEALTH INSURANCE EXPERIMENT: EVIDENCE FROM THE FIRST YEAR

Amy Finkelstein
Sarah Taubman
Bill Wright
Mira Bernstein
Jonathan Gruber
Joseph P. Newhouse
Heidi Allen
Katherine Baicker
The Oregon Health Study Group

The Oregon Health Insurance Experiment: Evidence from the First Year
Amy Finkelstein, Sarah Taubman, Bill Wright, Mira Bernstein, Jonathan Gruber, Joseph P.
Newhouse, Heidi Allen, Katherine Baicker, and The Oregon Health Study Group
NBER Working Paper No. 17190
July 2011
JEL No. H51,H75,I1

ABSTRACT

In 2008, a group of uninsured low-income adults in Oregon was selected by lottery to be given the chance to apply for Medicaid. This lottery provides a unique opportunity to gauge the effects of expanding access to public health insurance on the health care use, financial strain, and health of low-income adults using a randomized controlled design. In the year after random assignment, the treatment group selected by the lottery was about 25 percentage points more likely to have insurance than the control group that was not selected. We find that in this first year, the treatment group had substantively and statistically significantly higher health care utilization (including primary and preventive care as well as hospitalizations), lower out-of-pocket medical expenditures and medical debt (including fewer bills sent to collection), and better self-reported physical and mental health than the control group.

Amy Finkelstein
Department of Economics
MIT E52-274C
50 Memorial Drive
Cambridge, MA 02142
and NBER
afink@mit.edu

Sarah Taubman
National Bureau of Economic Research
taubmans@nber.org

Bill Wright
Center for Outcomes Research and Education
Providence Health & Services
bill.wright@providence.org

Mira Bernstein
National Bureau of Economic Research
mbernst@nber.org

Jonathan Gruber
MIT Department of Economics
E52-355
50 Memorial Drive
Cambridge, MA 02142-1347
and NBER
gruberj@mit.edu

Joseph P. Newhouse
Division of Health Policy Research and Education
Harvard University
180 Longwood Avenue
Boston, MA 02115-5899
and NBER
newhouse@hcp.med.harvard.edu

Heidi Allen
Center for Outcomes Research and Education
Providence Health & Services
heidi.allen@providence.org .

Katherine Baicker
Professor of Health Economics
Department of Health Policy and Management
Harvard School of Public Health
677 Huntington Avenue
Boston, MA 02115
and NBER
kbaicker@hsph.harvard.edu

The Oregon Health Study Group
includes Matt Carlson (Portland State University),
Tina Edlund (Deputy Director, Oregon Health
Authority),Charles Gallia (Oregon DHS), Eric
Schneider (RAND), and Jeanene Smith (Office for
Oregon Health Policy and Research) in addition
to the other authors of this paper

An online appendix is available at:
http://www.nber.org/data-appendix/w17190

1. INTRODUCTION

In early 2008, Oregon opened a waiting list for a limited number of spots in its Medicaid program for low-income adults, which had previously been closed to new enrollment. The state drew names by lottery from the 90,000 people who signed up. This lottery presents the opportunity to study the effects of access to public insurance using the framework of a randomized controlled design

Although there are literally hundreds of studies comparing the health and health care utilization of insured and uninsured populations (see Institute of Medicine (2003)), inferring the impact of health insurance from such comparisons is difficult because individuals with and without insurance coverage differ in many ways – such as income, employment, or health – that are likely to be correlated with the outcomes of interest.[1] Random assignment of health insurance to some but not others would avoid such confounding, but has never been done before in the United States.[2]

In this paper we examine the effects of the Oregon Medicaid lottery after approximately one year of insurance coverage. We present comparisons of outcomes between the treatment group (those selected by the lottery who had an opportunity to apply for Medicaid) and the control group (those not selected and thus not able to apply for Medicaid). We also present estimates of the impact of insurance coverage, using the lottery as an instrument for insurance coverage.

We organize our analysis around the potential costs and benefits of health insurance. On the cost side, we examine the impact of health insurance on increased health care utilization. On the benefit side, we examine the impact of health insurance not only on health, which is a standard topic of analysis, but also on consumption smoothing, which we proxy for with measures of financial strain. There is remarkably

---

[1] A much smaller number of studies have used quasi-experimental variation in the availability of public insurance to assess the effects of insurance. While such studies can improve on observational studies, this work has focused primarily on children or the elderly because of the available variation and, as with all quasi-experimental studies, the validity of the results depends on untestable identifying assumptions.
[2] The famous RAND Health Insurance Experiment from the 1970s – the only other randomized controlled health insurance experiment that we know of in a developed country – was designed to investigate the marginal impact of varying insurance cost-sharing features among approximately 6,000 insured individuals, not the effect of insurance coverage itself (Newhouse et al., 1993, Manning et al., 1987).

little empirical work on the impact of health insurance on consumption smoothing or financial strain, even though risk spreading is arguably the primary purpose of health insurance (see e.g. Zeckhauser 1970).

A priori, the sign – let alone magnitude – of the impact of health insurance on these various outcomes is not obvious. Although the price and income effects of subsidized health insurance should increase health care utilization, it is possible that increasing primary care utilization or improving health could result in offsetting reductions in hospital or emergency department use. Similarly, although health insurance is expected to improve health through increases in the quantity and quality of health care, it is also possible that by reducing the adverse financial consequences of poor health, health insurance may discourage investments in health and thereby worsen health outcomes.

Moreover, for our low-income study population, the value of Medicaid relative to the existing safety net options is uncertain. The impact of Medicaid could be attenuated (or potentially non-existent) if public health clinics and uncompensated care allow individuals to consume de facto free medical care similar to that of the insured. Medicaid's impact would also be attenuated if – as is often conjectured – Medicaid itself is not particularly "good" insurance – at least in terms of being able to access health care providers.

Our analysis draws on administrative data from hospital discharge, credit report, and mortality records, as well as on responses to a large mail survey that we conducted. The administrative data are objectively measured and should not be biased by the treatment and control groups differentially reporting outcomes, but they only cover a relatively narrow set of outcomes. The survey data allow examination of a much richer set of outcomes than is feasible with administrative data alone, but, with a 50 percent response rate, are subject to potential non-response bias. Our available evidence on this issue is reassuring, but limited; response rates and available pre-randomization measures appear reasonably balanced between treatment and control responders.[3]

---

[3] Non-response or attrition bias is always an important concern in interpreting results from a randomized experiment using non administrative data. For example, in the RAND Health Insurance Experiment, response rates varied by 24 percentage points across treatment arms with different insurance coverage, from 87 percent for those randomized into the most generous plan to 63 percent for those randomized into the least generous plan (Newhouse et al., 1993).

Prior to looking at the data on outcomes for the treatment group, virtually all of the analysis presented here was pre-specified and publicly archived in a detailed analysis plan.[4] This was designed to minimize issues of data and specification mining and to provide a record of the full set of planned analyses. Although pre-specification of hypotheses is the norm for randomized controlled medical trials, to our knowledge is it extremely rare in evaluation of social policy experiments.[5]

About one year after enrollment, we find that those selected by the lottery have substantial and statistically significantly higher health care utilization, lower out-of-pocket medical expenditures and medical debt, and better self-reported health than the control group that was not given the opportunity to apply for Medicaid. Being selected through the lottery is associated with a 25 percentage point increase in the probability of having insurance during our study period. This net increase in insurance appears to come entirely through a gross increase in Medicaid coverage, with little evidence of crowd-out of private insurance. Using lottery selection as an instrument for insurance coverage, we find that insurance coverage is associated with a 2.1 percentage point (30 percent) increase in the probability of having a hospital admission, an 8.8 percentage point (15 percent) increase in the probability of taking any prescription drugs, and a 21 percentage point (35 percent) increase in the probability of having an outpatient visit; we are unable to reject the null of no change in emergency room utilization, although the point estimates suggest that such use may have increased. In addition, insurance is associated with three-tenths of a standard deviation increase in reported compliance with recommended preventive care such as mammograms and cholesterol monitoring. Insurance also results in decreased exposure to medical liabilities and out-of-pocket medical expenses, including a 6.4 percentage point (25 percent) decline in the probability of having an unpaid medical bill sent to a collection agency and a 20 percentage point (35 percent) decline in having any out-of-pocket medical expenditures. Since much medical debt is never

---

[4] Our analysis plan was archived on December 3, 2010 at http://www.nber.org/sap/20101203/ and at hypotheses@povertyactionlab.org. Some of those analyses yielded little of interest and therefore we describe them briefly, presenting the full results only in appendices. In the few instances in which the results suggested the performance of additional analyses that had not originally been planned, we have indicated this in the text and tables with a ^.

[5] Within economics, we know of only a few examples in developing countries (Alatas et al. (2010), Olken, Onishi and Wong (2010) and Schaner (2010)) and none in the United States.

paid, the financial incidence of expanded coverage thus appears to be both on the newly insured and on their medical providers. Finally, we find that insurance is associated with improvements across the board in our measures of self-reported physical and mental health, averaging two-tenths of a standard deviation improvement. These results appear to reflect improvements in mental health and also at least partly a general sense of improved well being; they may also reflect improvements in objective, physical health, but this is more difficult to determine with the data we now have available.

Our estimates of the impact of public health insurance apply to able-bodied uninsured adults below 100 percent of poverty who express interest in insurance coverage. This is a population of considerable policy interest. In 2011, fewer than half of the states offered Medicaid coverage to able-bodied adults with income up to 100 percent of poverty absent specific categorical requirements (Kaiser 2011). As part of the 2010 Patient Protection and Affordable Care Act, starting in 2014 all states will be required to extend Medicaid eligibility to all adults up to 133 percent of the federal poverty level, with no financial penalties for many individuals in this income range who do not take up coverage (Kaiser 2010a, Kaiser 2010b, US GPO 2010).

The rest of the paper is structured as follows. Section two provides background on the Oregon Medicaid program and the lottery design. Section three describes the primary data sources, and section four presents our empirical framework. Section five presents our main results. Section six discusses interpretation and extrapolation of our estimates. A number of appendices – referenced in the text – provide additional details.

2. OREGON'S MEDICAID LOTTERY

The Oregon Health Plan (OHP) – created by one of the first federal waivers of traditional Medicaid rules – currently consists of two distinct programs: OHP Standard and OHP Plus.[6] OHP Plus serves the categorically eligible Medicaid population, which includes (up to specific income thresholds) children

---

[6] See Oregon Office for Health Policy Research (2009) for more detail on the Oregon Health Plan. The Oregon Health Plan is well known for its efforts in the 1990s to create a list of covered services for Medicaid recipients, based on available funding and evidence of their effectiveness. In practice, however, this "rationing" has not been binding on Medicaid patients (Bodenheimer, 1997).

and pregnant women, the disabled, and families enrolled in Temporary Assistance to Needy Families (TANF). OHP Standard, which is the program that was lotteried, is a Medicaid expansion program to cover low-income adults who are not categorically eligible for OHP Plus. Specifically, it covers adults ages 19 – 64, not otherwise eligible for Medicaid, who are Oregon residents, are U.S. citizens or legal immigrants, have been without health insurance for six months, have income below the federal poverty level (FPL), and have assets below $2,000.

OHP Standard provides relatively comprehensive benefits with no consumer cost sharing. It covers physician services, prescription drugs, all major hospital benefits, mental health and chemical dependency services (including outpatient services), hospice care, and some durable medical equipment. Vision is not covered nor are non-emergency dental services. Wallace et al. (2008) estimate that, in 2001-2004, average annual Medicaid expenditures for an individual on OHP Standard were about $3,000. Most care is provided through managed care organizations. Monthly enrollee premiums range from $0 to $20 depending on income (with those below 10 percent of the FPL paying $0).

At its peak in early 2002, about 110,000 people were enrolled in OHP Standard, about one-third the size of OHP Plus enrollment at that time. Due to budgetary shortfalls, OHP Standard was closed to new enrollment in 2004. By early 2008, attrition had reduced enrollment to about 19,000 and the state determined it had the budget to enroll an additional 10,000 adults. Therefore, in January 2008 the state re-opened OHP Standard to new enrollment.

Because the state (correctly) anticipated that the demand for the program among eligible individuals would far exceed the 10,000 available new enrollment slots, it applied for and received permission from the Centers for Medicare and Medicaid Services to add the new members through random lottery draws from a new reservation list.[7] From January 28 to February 29, 2008, anyone could be added to the lottery

---

[7] In 2008, there were approximately 200,000 uninsured adults aged 19-64 below 100% FPL (see www.statehealthfacts.org last accessed February 2011). Adding some of these individuals to OHP Standard based on health status is not allowed by Federal law, which prevents states from determining eligibility for federal programs based on health care conditions. The state also considered enrolling individuals on a 'first-come, first-served' basis, but rejected that option because of concerns that this would put people without ready access to the information or the means to get on the list quickly at a disadvantage (Oregon Department of Human Services, 2008).

list by telephone, by fax, in person sign-up, by mail, or online. The state conducted an extensive public awareness campaign about the lottery opportunity. To keep barriers to sign-up low, the sign-up form (shown in Appendix Figure A2) requested limited demographic information on the individual and any interested household member, and no attempt was made to verify the information or to screen for program eligibility at sign-up for the lottery. A total of 89,824 unique individuals were placed on the list during the five week window it was open.

The state conducted eight lottery drawings from the list, roughly equally sized and evenly staggered from March through September 2008. Selected individuals won the opportunity – for themselves and any household member (whether listed or not) – to apply for OHP Standard coverage; treatment thus occurred at the household level. In total, 35,169 individuals – representing 29,664 unique households –- were selected by lottery. If individuals in a selected household submitted the appropriate paperwork within 45 days after the state mailed them an application and demonstrated that they met the eligibility requirements, they were enrolled in OHP Standard.[8] About 30 percent of selected individuals successfully enrolled. There were two main sources of slippage: only about 60 percent of those selected sent back applications, and about half of those who sent back applications were deemed ineligible, primarily due to failure to meet the requirement of income in the last quarter corresponding to annual income below the poverty level, which in 2008 was $10,400 for a single person and $21,200 for a family of four (Allen et al. 2010). If they did successfully enroll in OHP Standard, individuals could remain enrolled indefinitely, provided that they re-certified their eligibility status every six months.

3. DATA

We briefly describe each data source here. Additional details can be found in Appendix 1.

---

[8] The application inquired about Oregon residence, U.S. citizenship, insurance over the last six months, household income over the last two months, and assets (Oregon DHS, Form 7210). Documentation of identity and citizenship (in the form of passports, birth certificates, etc) and income (in the form of pay stubs, letters from employers, etc) was required (Oregon DHS, Form 7222). The state reviewed applications, first examining eligibility for OHP Plus and then, if not eligible for Plus, examining eligibility for OHP Standard. Plus does not have premiums and covers some additional services that Standard does not such as home health, elective hospital stays, physical, occupational and speech therapy and private duty nursing (Oregon Office for Health Policy Research, 2009).

3.1. Administrative data on outcomes: hospital discharge data, credit report data, and mortality data

We obtained standard individual-level hospital discharge data for the entire state of Oregon from January 2008 through September 2009 and probabilistically matched them to the lottery list based on information provided at the time of lottery sign-up on full name, zip code and date of birth. The data include a hospital identifier, dates of admission and discharge, source of admission, detail on diagnoses and procedures, and discharge destination. Similar discharge data have been used to study the impact of health insurance in other contexts (see e.g. Doyle (2005), Card et al (2008, 2009) and Anderson et al. (2010)). Although inpatient admissions are relatively rare (the annual admission rate for our controls is only about 5 percent), they are quite expensive, accounting for about one-quarter of total medical expenditures for 18 – 64 year olds.[9] We observe, on average, about 5 months of pre-randomization data.

We obtained detailed credit records from TransUnion's Consumer Credit Database. Credit bureaus like TransUnion collect rich and detailed information on virtually all formal consumer borrowing, yet the analysis of such data is still relatively rare in the economics literature and, to our knowledge, has never been done before in a health insurance context.[10] TransUnion used the full name, full address and date of birth reported at sign up to match 68.5 percent of lottery participants to their pre-randomization credit report in February 2008. The credit bureau was able to track over 97 percent of those found in the February 2008 file to their September 2009 file. Our primary outcomes of financial strain are measured in this 2009 file, which thus has an effective post-randomization "attrition rate" of 3 percent. We also observe pre-randomization outcomes measured in February 2008.

We obtained mortality data from Oregon's Center of Health Statistics for all deaths occurring in Oregon from January 1, 2008 through September 30, 2009 and probabilistically matched our sample using full name, zip code, and date of birth.

---

[9] Author calculations based on publicly available tables from the 2008 Medical Expenditure Panel Survey.
[10] Avery, Calem and Canner (2003) provide an excellent, detailed discussion of credit bureau data; much of our discussion of the data and our choice of analysis variables is guided by their work.

3.2 Survey data on outcomes

We supplement the outcome measures available in the administrative data with a mail survey that was sent out in seven "survey waves" over July and August 2009 to virtually all individuals selected by the lottery and a roughly equal number of unselected individuals. The complete survey instrument is shown in Appendix Figure A4. The basic protocol involved three mail attempts. In addition, we designed a more intensive protocol (conducted on approximately 30 percent of non-respondents) which included additional tracking efforts, mailings, and phone contacts. The response rate to the basic protocol was 36 percent; following the intensive protocol, the overall weighted response rate was 50 percent, with individuals who responded to the intensive follow-up weighted by the inverse probability of their being included in the intensive follow-up subsample.

In Section 5.3, we also briefly compare some of our estimates from this main survey to those from two earlier, virtually identical surveys of the same population: an "initial survey" conducted approximately one year earlier (i.e. shortly after random assignment), and a "six month" survey conducted about mid way between the initial and main survey. The six month survey was conducted on a 20 percent subsample of the sample used in the other two surveys. The earlier surveys used similar protocols but did not have an intensive follow up arm; the initial and six month surveys achieved response rates of 45 and 42 percent respectively.

3.3 Other data

We obtained pre-randomization demographic information that the participants provided at the time of lottery sign-up. Appendix Figure A2 shows the sign-up form. We use these data primarily to construct nine "lottery list variables" which we use below to examine treatment and control balance on pre-randomization demographics.[11] We also obtained state administrative records on the complete Medicaid enrollment history of lottery list participants from prior to the lottery through September 2009. We use

---

[11] These nine "lottery list variables" are year of birth; sex; whether English is the preferred language for receiving materials; whether the individuals signed themselves up for the lottery or were signed up by a household member; whether they provided a phone number on sign up; whether the individuals gave their address as a PO box; whether they signed up the first day the lottery list was open; the median household income in the 2000 census from their zip code; and whether the zip code they gave is within a census-defined MSA.

these data as our primary measure of the first stage outcome (i.e. insurance coverage). We obtained broader measures of insurance coverage (including private insurance) from our mail survey. Finally we obtained state administrative records on the Food Stamp benefit history of lottery list participants from prior to the lottery through September 2009.

3.4 Time frame of the study

In the administrative data we measure outcomes from the date individuals were notified that they were selected (i.e. their "notification date") through the end of September 2009. [12] This observation period represents, on average, 16 months (standard deviation of 2 months) after individuals are notified of their selection and, on average 14 months (standard deviation of 3 months) after insurance coverage is approved for those selected by the lottery who successfully enrolled in OHP Standard. [13] If an individual successfully obtained insurance through the lottery, coverage was applied retroactively to only a few days after the state mailed the application to the individual, which was on average about one month after the notification date and one month prior to the approval date. [14]

In our survey most outcomes were asked with a six month look-back period (e.g. number of doctor visits in the last six months) or based on "current" conditions (e.g. self-reported health). There is variation across individuals in when surveys were mailed and how long they took to respond, as well as their lottery draw (and hence notification date). Our average survey response occurs 15.3 months after notification date (standard deviation = 2.7) months or 13.1 months after insurance approval (standard deviation = 2.9 months).

---

[12] We randomly assigned lottery draws (and hence notification dates) to the control individuals as discussed in more detail in Section 4. A primary reason for measuring outcomes from notification date – which varies by up to seven months across lottery draw – rather than just measuring all outcomes from the earliest notification date of any lottery draw, is to increase the availability of "pre-randomization" hospital data for analysis. If we analyzed outcomes from the earliest notification date we would have less than 3 months of pre-randomization data; this approach gives us almost 9 months for some participants and on average 5 months.

[13] The timing difference between notification and approval reflects the approximately one month lag between when the state notified individuals they had been selected and when the state mailed them an application, and then the time it took for individuals to fill out and mail in an application, and for the state to process it.

[14] We suspect, and focus group interviews with selected individuals suggest, that selected individuals would have been unlikely to change their behavior while their applications were being processed; however the retroactive insurance coverage may have affected the financial burden associated with health care utilization during that time period.

3.5. Sample definition and descriptive statistics

Of the 89,824 individuals who were on the lottery list, we used pre-randomization data to exclude individuals who were not eligible for OHP Standard (because they gave an address outside of Oregon, were not in the right age range, or had died prior to the lottery), had institutional addresses, were signed up by third parties, would have been eligible for Medicare by the end of our study period, or were inadvertently included on the original list multiple times by the state. These exclusions brought our study population down to 74,922 individuals (representing 66,385 unique households). Of these, 29,834 individuals were selected by the lottery and the remaining 45,088 individuals are controls. Of these, we surveyed 29,589 treatment and 28,816 control individuals. Appendix Figure A1 shows the overlap between the full sample (which is used in the hospital discharge data and mortality analysis), the credit report subsample and the mail survey subsample.

Table 1 provides some demographic summary statistics for our control sample. Panel A, based on information provided prior to randomization on the sign up list, shows that our study population is 56 percent female; about one-quarter are 50 to 64 at the end of our study period (the average age is 41); 92 percent have a preferred language of English; and about three-quarters live in an MSA. Panel B reports additional demographic characteristics of the control group from the survey. These outcomes are only available for individuals who responded to the mail survey, and may therefore not be representative of the full sample. The population is 4 percent black, and 12 percent Hispanic. Almost one-fifth has less than a high school education, and another half has only a high school diploma or GED. Over half report that they are not currently working. Most strikingly, they appear to be in quite poor health: 18 percent report having ever been diagnosed with diabetes, 28 percent with asthma, 40 percent with high blood pressure; 56 percent screen positive for depression.[15]

---

[15] By contrast for a general adult population, 7 percent report ever being diagnosed with diabetes, 14 percent with asthma, 24 percent with high blood pressure, and 28 percent with depression. (These numbers are based on our calculation from the 2004-2009 Behavioral Risk Factor Social Surveillance Survey which uses virtually identical questions to our survey questions.) Moreover, diagnoses likely understate the level of poor health, particularly in our low-income population with limited access to providers who could diagnose disease.

Panel B also shows the distribution of household income (relative to the federal poverty level) and

insurance coverage. Both are important for the first-stage impact of lottery selection on insurance

coverage. About 70 percent report incomes below the eligibility cut-off of 100 percent of the federal

poverty level; this is consistent with our finding from analyzing application data that t income eligibility

requirements disqualified a non-trivial share selected individuals (Allen et al., 2010). Finally, about 30

percent of the controls report having insurance (which would also make them ineligible for OHP

Standard); 13 percent report having private insurance.

4. EMPIRICAL FRAMEWORK

4.1 Reduced form

4.1.1 Reduced form equation

We estimate the reduced form effect of winning the lottery with the following equation:

$$y_{ihj} = \beta_0 + \beta_1 LOTTERY_h + X_{ih}\beta_2 + V_{ih}\beta_3 + \varepsilon_{ihj} \tag{1}$$

where *i* denotes an individual, *h* denotes a household and $j \in J$ denotes a "domain" of related outcomes

(such as health or financial strain). For example y_{ij} might be the self-reported health of individual *i*, which

is one of the health measures in the health "domain" J. We define (sign) each outcome within a domain

so that higher values all have the same interpretation within a domain (e.g. more health care use, more

financial strain). As we discuss below, we will summarize the estimates within a domain by the

standardized treatment effect across outcomes in that domain; we will also report estimates for individual

outcomes and show p-values that are adjusted to account for the multiple outcomes examined within the

domain.

LOTTERY is an indicator variable for whether or not household *h* was selected by the lottery. The

coefficient on *LOTTERY* (β_1) is the main coefficient of interest, and gives the average difference in

(adjusted) means between the treatment group (the lottery winners) and the control group (those not

selected by the lottery).

We denote by X_{ih} the set of covariates that are correlated with treatment probability (and potentially with the outcome) and therefore must be controlled for so that estimates of β_1 give an unbiased estimate of the relationship between winning the lottery and the outcome. In all our analyses, X_{ih} includes indicator variables for the number of individuals in the household listed on the lottery sign up form (hereafter "household size"); although the state randomly sampled from individuals on the list, as mentioned above the entire household of any selected individual was then considered selected and eligible to apply for insurance. As a result, selected (treatment) individuals are disproportionately drawn from larger household sizes.[16] For outcomes in the survey data, X_{ih} also includes indicator variables for survey wave (and the interaction of these indicator variables with household size indicators) since the fraction of treatment individuals varies across the seven survey waves.[17]

We denote by V_{ih} a second set of covariates that can be included to potentially improve power by accounting for chance differences in variables between treatment and control group but that are not needed for β_1 to give an unbiased estimate of the relationship between winning the lottery and the outcome. There are three potential sources of such variables: the "lottery list" demographic variables; pre-randomization measures of outcomes in the credit report data and hospital discharge data; and the lottery draw to which the individual is assigned.[18] Our analysis of survey data will not control for any V_{ih} covariates; our analysis of administrative data will include lottery draw indicators as well as the pre-randomization measure for the outcome analyzed in the hospital and credit report data. Appendix 2.3

---

[16]The proportion of treated (respectively, control) individuals in household size 1 is 66.5 (respectively, 83.6), in household size 2 is 33.1 (respectively, 16.4) and in household size 3 is 0.5 (respectively, 0.04). We do not study winning household members who were not on the original list; such individuals do not affect probability of selection and are not included in our measure of household size. Virtually all listed household members are adults since only individuals aged 19-64 are eligible for OHP Standard.

[17] Our initial survey (conducted very shortly after the lottery began) included the treated individuals and an oversample of then-untreated individuals. Lower than expected take-up resulted in more draws than anticipated to reach the state's enrollment target, and thus there was a higher treatment probability in earlier waves. We kept the wave assignments constant between this initial survey and the main survey conducted the following year.

[18] In the administrative data, we measure outcomes from the notification date. For treatment individuals, notification date varies by lottery draw (which span a seven month period). For control individuals, we randomly assigned a lottery draw at the household level, stratified on household size, to match the distribution of lottery draws among the treatments so that, by construction, treatment probability is uncorrelated with lottery draw within household size.

shows that our results are not sensitive to other choices regarding the V_{ih} covariates in either the survey or administrative data.

In all of our analyses we cluster the standard errors on the household identifier since the treatment is at the household level. Analyses of survey data are weighted to account for the sampling design of the survey as described above (and in Appendix 1.7).

4.1.2 Handling many outcomes: standardized treatment effects and multiple inference[19]

We summarize multiple findings across related outcomes within a domain *J* by the average standardized treatment effect:

$$\sum_{j \in J} \frac{1}{J} \frac{\xi_{1j}}{\sigma_j} \tag{2}$$

where σ_j is the standard deviation of y_j in the control group and ξ_{1j} is the coefficient of interest for outcome j. (Specifically, for the reduced form in equation (1), the ξ_{1j}'s correspond to the β_{1j}'s in equation (1)).[20] In order to account for covariance in the estimates of ξ_{1j} / σ_j we estimate pooled OLS for all outcomes $j \in J$. An important limitation of standardized treatment effects is that they implicitly "weight" each outcome within a domain equally, which may not be desirable.

In addition to standardized treatment effects, in each domain we also report the underlying estimates on the individual outcomes (i.e. the β_{1j}'s) due to their ease of interpretation, as well as their individual interest. For the individual outcomes we report both per comparison p-values and "family-wise" p-values adjusted to account for the multiple outcomes examined within the domain.[21] The per-comparison p-value may be appropriate if one is interested in the specific hypothesis of the impact of insurance on a

---

[19] This approach draws heavily on Kling and Liebman (2004) and Kling Liebman and Katz (2007).

[20] To calculate the standard error on the standardized treatment effect we estimated pooled OLS (or pooled IV in the case of standardized treatment effects for the IV estimation discussed in Section 4.3) for all outcomes within the domain.

[21] The family-wise p value corresponds to the probability of rejecting the null hypothesis of no effect on a given outcome under the null family of hypotheses of no effect on any outcome in this domain. We calculate these family-wise error rate adjusted p-values based on 10,000 iterations of the free step-down resampling method of Westfall and Young (1993). This is more powerful than a standard Bonferroni correction because it does not assume independence across the outcomes within a domain and sequentially removes hypotheses from the family after they are rejected; see Kling and Liebman (2004) or Anderson (2008) for more detailed discussions as well as applications.

specific outcome (such as depression). This adjusted p-value is more appropriate for considering the test of the specific outcome as part of a set of tests on all the outcomes in the domain of that standardized treatment effect (such as overall health).[22] In practice, it is rare for one of our results to be statistically significant (e.g. p value of less than 0.05) in the per-comparison test and not in the multiple inference adjustment.^

4.2 Validity of the experimental design

Our causal inference rests on the twin assumptions that assignment of the ability to apply for OHP Standard was in fact randomized in the way described, and that the treatment and control individuals in the sub-samples we use to analyze outcomes are not differentially selected from the full sample.

The lottery's random selection process was performed by Oregon's Department of Human Services (DHS). We verified through independent computer simulations that we could replicate their procedure to within sampling error(Appendix 2.1 and Table A12 provide more detail); we also demonstrate that the procedure we used to draw our survey sample produced balance of treatment and control characteristics (see Appendix 2.2).

Differences in attrition (match rates or response rates) or in the pre-randomization characteristics of the treatment and control analysis samples would raise concerns about the key assumptions for causal inference. Table 2 therefore investigates treatment-control balance for three different samples in columns 2 through 4, respectively: the sample universe (which is the sample analyzed in the hospital discharge

---

[22] Needless to say, there is a fair degree of arbitrariness regarding what is grouped into the same standardized treatment effect, and readers may legitimately prefer some more or less conservative adjustments. One extreme would be adjustments that take account of all the outcomes analyzed, both in this paper and in any other (present or future) paper about the study. In addition to being impractical, we do not feel that the null hypothesis of "no impact of health insurance on anything" is a particularly useful one, and it would discourage investigation of effects that a priori are considered unlikely. Another extreme would be to treat each outcome analyzed as an independent hypothesis, and report only per-comparison (unadjusted) p-values. We also found this unappealing, since one might feel very differently if, for example, only one of twenty closely related self-reported health measures had a per-comparison p-value of less than 0.05. A virtue of our current grouping is that it was pre-specified in advance of analysis.

^ In the archived analysis plan we proposed presenting standardized treatment effects of related outcomes within a domain separately for both survey and administrative data, as well as a third standardized treatment effect using the outcomes from both survey and administrative data in a given domain. Given the major substantive and methodological differences between the two types of data, in this paper we have opted for reporting only the standardized treatment effects across outcomes within domains for the survey and administrative data separately. In practice this makes a negligible difference to the adjusted p-values; results available upon request.

data and the mortality data), the credit report subsample, and the survey respondents. A priori we were most concerned about the potential for imbalance between treatment and controls in the subsample of survey respondents, given the 50 percent non-response rate (compared to an effective match rate of over 97 percent in the credit report data).[23]

Panel A shows the balance of match rates. In the credit report data, the difference in (unconditional) match rates between treatment and control is a statistically insignificant 0.4 percentage points. In the survey respondent sample, there is a statistically significant 1.6 percentage point (standard error = 0.7) lower response rate for treated individuals, off of a 51 percent base.[24] By way of comparison, our estimated difference in response rates across treatment and control is much smaller than in the RAND Health Insurance Experiment; there, the overall response rate was higher (77 percent), but those randomized into more coverage had systematically higher response rates, with a 24 percentage point difference in response rate between the most and least comprehensive plan (87 percent versus 63 percent; Newhouse et al., 1993).

Among the matched or responding subsample, Panel B reports the treatment-control balance for various pre-randomization characteristics. The first row ("lottery list characteristics") shows the F-statistics and p-values on the treatment-control balance of all the nine lottery list demographics. The second row ("pre-randomization outcomes") shows the balance of pre-randomization outcomes that match ones that we subsequently analyze post-randomization. The number and definition of these outcomes varies according to the sample studied.[25] Finally, the third row ("both of the above") examines balance on the lottery list and pre-randomization outcomes combined. In each of the three samples we are

---

[23] Our response rate is quite good for a mixed mode mail and phone survey of a low-income population in the United States (for some comparisons see e.g. Beebe et al 2005, Brown et al. 1999, Carlson et al 2006, Fowler et al. 1999, Gallagher et al. 2005, Hartz et al. 2000, and AHQR 2001) although it of course leaves substantial scope for non-response bias arising from difference between treatment and control responders.

[24] Conditional on response, response time between treatment and controls is indistinguishable.

[25] For the hospital sample we examine 12 pre-randomization hospital outcomes measured on average for 5 months prior to the lottery. For the credit report sample we examine 10 pre-randomization credit report outcomes measured from February 2008 (the month before the lottery began) and with the same approximately 16 month look-back period as the post-randomization measures. For the survey sample we have no pre-randomization survey measures but we selected four pre-randomization measures from the credit report and hospital discharge data that are similar to subsequent survey measures. The selection of all these variables was pre-specified and is described in more detail in Appendix 2.2, where we also present treatment-control differences for each individual variable analyzed.

unable to reject the null of treatment-control balance on the lottery list variables, the pre-randomization measures, or the combined set of variables. All p-values are above 0.23. For the survey respondent subsample – where a priori we were most concerned about potential imbalance – the p-values are all at least 0.28.[26]

A separate question from the balance of treatment and controls within a subsample is how the various subsamples compare to each other. Differences in characteristics across sub-samples do not threaten causal inference, but may be important for comparing estimates across data sets as well as for extrapolatingto other contexts. Appendix Table A1 shows that survey respondents are on average almost two years older and about 3 percentage points more likely to be female than the full sample. There are no obvious observable differences between the credit report subsample and the full sample.

4.3 IV

The reduced form (or intent-to-treat) estimates from equation (1) provide an estimate of the causal effect of winning the lottery (i.e. winning the ability to apply for OHP Standard). This provides an estimate of the net impact of expanding access to public health insurance. We are also interested in the impact of insurance coverage. This is modeled as follows:

$$y_{ihj} = \pi_0 + \pi_1 INSURANCE_{ih} + X_{ih}\pi_2 + V_{ih}\pi_3 + v_{ihj} \tag{3}$$

where INSURANCE is a measure of insurance coverage and all other variables are as defined in equation (1). We estimate equation (3) by two stage least squares (2SLS), using the following first stage equation:

$$INSURANCE_{ihj} = \delta_0 + \delta_1 LOTTERY_{ih} + X_{ih}\delta_2 + V_{ih}\delta_3 + \mu_{ihj} \tag{4}$$

---

[26] Another potential way to examine non-response bias in the survey data could be to investigate whether the estimated impact of insurance on outcomes in the administrative data differs for survey responders compared to survey non-responders. Our concern with this approach is that it could confound potential non-response bias with potential heterogeneity in treatment effects across responders and non-responders, so that observed differences (or lack thereof) do not have clear implications for the presence of non-response bias. For this reason our analysis plan specified only the differences in pre-randomization measures presented here for investigating potential non-response bias. In practice, our subsequent attempt to also compare the estimates of the impact of insurance on later outcomes in the administrative data for survey responders and non responders yielded estimates that were too imprecise to be informative.

in which the excluded instrument is the variable *LOTTERY* with the first stage coefficient of δ_1.[27]

Because the model is just identified, the 2SLS estimate of π_1 is given by the ratio of the reduced form (equation 1) and first stage (equation 4) coefficients (β_1 / δ_1). We interpret the 2SLS estimates as a local average treatment effect or LATE. (Imbens and Angrist, 1994). In other words, the 2SLS estimate of π_1 identifies the causal impact of insurance among the subset of individuals who would obtain insurance on winning the lottery and would not obtain insurance without winning the lottery (i.e. the compliers). Table 3 reports our first stage estimates based on estimation of equation (4). The first row reports the estimates for the measure of insurance that is used in all of our IV estimates: whether the individual was ever on Medicaid (which includes both OHP Standard and OHP Plus) during our study period, as measured by the state's Medicaid enrollment files.[28] The results indicate a first stage of 0.26 for both the full sample (column 2) and the credit-report subsample (column 4) and a first stage of 0.29 for the survey respondents (column 6). All of these first stages have F-statistics above 500. The first stage coefficient is considerably less than 1, primarily reflecting the 30 percent take-up discussed above; in addition, a small percentage of the controls became eligible for OHP Plus over our study period.[29] Relative to our study population, compliers are somewhat older, more likely white, in worse health, and in lower socio-economic status (see Appendix table A23).

The subsequent rows of Table 3 report first stage estimates for alternative definitions of insurance. We focus the discussion on the results for the full sample (column 2); the other samples show very similar

---

[27] When we report standardized treatment effects for IV estimates, they are calculated based on the formula in equation (2) and using pooled IV estimates of equation (3) across outcomes.

[28] For purposes of measuring the first stage, the study period is defined as ending on September 30, 2009. For the administrative data the study period begins with the notification date (which varies by lottery draw). For the survey data the study period is defined from the first notification date (i.e. March 10, 2008), since within survey wave the notification date varies. Although the average response date to the survey was September 23, 2009, about 35 percent of our weighted responses occur after this, in large part because of the intensive follow-up timing which continued in some cases till as late as March 2010 (see Appendix 1.7). Defining the study period differently for individuals based on their response date makes little difference for our estimated first stage estimates. This can be seen by comparing the first stage estimates for survey respondents in row 8 of Table 3 (where insurance is defined as covered by Medicaid on the date closest to the survey response date) and row 4 of Table 3 (where insurance is defined as covered by Medicaid as of September 30, 2009); the estimated first stages among survey respondents are 0.18 and 0.19 respectively.

[29] As we discuss in more detail in the data Appendix 1.3, the way the state identifies lottery participants in the Medicaid enrollment files may cause it to slightly under-estimate enrollment among non-selected (control) individuals, and thus cause us to over-estimate our first stage by, we estimate, 2 percentage points or less.

patterns. Not surprisingly, the results in row 2 show that estimated first stage is virtually the same if we examine only coverage by OHP Standard (the program directly affected by the lottery) versus coverage by either public program, indicating that selection by the lottery is not associated with an increase in coverage by OHP Plus. Row 3 shows that on average the lottery is associated with an increase of 3.4 months with Medicaid.

Over time the difference in insurance coverage between treatment and controls attenuates as enrolled individuals had to recertify their eligibility every 6 months and control individuals could find other insurance (particularly OHP Plus); this can be seen clearly in Appendix Figure A3. As a result, when the dependent variable is defined as "on Medicaid at the end of the study period" rather than our primary measure of "ever on Medicaid" the estimated first stage declines from 0.26 (row 1) to 0.15 (row 4).

The survey data (column 4) provide a broader measure of insurance coverage than available in the Medicaid administrative data. The results in row 5 indicate that the estimated increase in self-reported "any" insurance coverage is very similar (within two percentage points) to the estimated increase in self-reported Medicaid coverage (row 7). Consistent with this, row 6 indicates a decline in self reports of private insurance coverage of only 0.8 percentage points (standard error =0.5). The estimated increase in Medicaid coverage as reported in the survey and as measured in the administrative data are quite similar when measured over the same time horizon (compare rows 7 and 8).

The identifying assumption behind the 2SLS estimates is that there is no effect on the outcomes studied, on average, of winning the lottery that does not operate via the lottery's impact on insurance coverage. We believe this is a reasonable approximation, but may not be strictly true. There are (at least) two possible types of violations. First, it is possible that the event of winning (or losing) the lottery may have direct effects on the outcomes we study, although it seems unlikely to us that such effects both exist and would persist for a year after the lottery. [30] Second, individuals who apply for public health insurance

---

[30] In much of the literature on public health insurance, there is a notion of a potential "option value" of public health insurance among those who are eligible but not covered since they may choose to take up that coverage if and when it becomes needed (see e.g. Cutler and Gruber (1996) who refer to such individuals as "conditionally covered").

may also be encouraged to apply for other public programs for which they are eligible, such as food stamps or cash welfare. In particular, if the individual applied for OHP in person (rather than by mail) case workers were instructed to offer assistance to the applicant (if they were interested) in also applying for TANF (cash welfare), the Supplemental Nutrition Assistance Program (more commonly known as food stamps), and employment related child care assistance. These other cash (or cash equivalent) transfer programs could have direct effects on the outcomes we study. This is not an idiosyncratic feature of our setting but a more general feature of the application process for public programs; as such, it may be a relevant component of the impact of attempts to expand Medicaid more generally. However, any direct impact of winning the lottery on receipt of other benefits is a violation of the exclusion restriction for the IV interpretation of the impact of insurance per se, as opposed to the effect of expanded access to Medicaid (the reduced form).

The results in Table 3 suggest that there is a statistically significant but substantively trivial impact of the treatment (lottery selection) on other program receipt. Rows 9 and 10 of Table 3 indicate that selection by the lottery is associated with a statistically significant increase in the probability of food stamp receipt of 1.7 percentage points and of total food stamp benefits over about $60 over a 16 month period, or less than 0.5 percent of annual income.[31] Estimates of the income elasticity of health care use range from a low end of about 0 to a high end of about 1.5 (Getzen 2000, Table 1), suggesting that the income effect of food stamp receipt on health care use would be considerably less than 1 percent. The impact on health seems likely to be negligible as well.[32]

However in our context this is not relevant since individuals who won the lottery were only eligible for Medicaid coverage if they successfully submitted an application within 45 days of receiving it.

[31] This is likely an upper estimate since the cash equivalent of food stamps may be less than one (Hoynes and Schanzenbach 2009). Moreover, some of the increase in food stamp benefits could reflect an impact of insurance, rather than of insurance application. Consistent with this, the estimated impact of lottery selection on food stamp receipt grows over time; for example, in the first three months after lottery selection (which is more plausibly a time frame during which one would see direct effects of selection on other program receipt) the impact of lottery on food stamp receipt and food stamp benefits is only 0.4 percentage points (standard error =0.3) and $3.6 (standard error = 2.5) respectively (not shown).

[32] We have not yet been able to obtain data on receipt of TANF or employment related child care assistance. However TANF receipt qualifies the individual for OHP Plus and our estimates suggest that the lottery is associated with about a 0.3 percentage point increase in OHP Plus (statistically significant in some samples) which suggests the

5. RESULTS[33]

5.1 Health care utilization

5.1.1 Administrative data

Table 4 presents our primary estimates using the hospital discharge data; all analyses exclude admissions for childbirth.[34] Table 4a reports results on the extensive margin (i.e. admission probabilities). The two stage least squares results suggest that insurance is associated with an increase in the probability of any hospital admission of 2.1 percentage points (standard error = 0.7), or about 30 percent. The increase in hospital admissions appears to be disproportionately concentrated in the approximately 35 percent of admissions that do not originate in the emergency room, suggesting that these admissions may be more price sensitive.

Table 4b examines three measures of total utilization commonly used in the literature (see e.g. Card et al., 2009): number of hospital days, total list charges, and number of procedures performed.[35] Although total utilization is arguably of greater interest (particularly for estimating the impact on total costs), not surprisingly – given the skewed right tail of total utilization – we have less statistical precision here than on the extensive margin. [36] The 2SLS estimates in Panel A show substantial increases in each of the three measures of utilization – with implied proportional increases of about 20 percent for hospital days,[37] 40

---

lottery did not have a substantive impact on TANF receipt. The child care assistance program is considerably smaller than the other two.

[33] More detail on variable definitions and descriptive statistics for the outcomes analyzed here can be found in Appendix 1.

[34] Regardless of lottery selection, many women in our sample would become categorically eligible for OHP Plus for childbirth.

[35] List charges are accounting charges for rooms and procedures and do not reflect transacted prices. They are perhaps best viewed as a price-weighted summary of treatment, albeit at somewhat artificial prices (Card et al., 2009).

[36] Given the skewed nature of these outcomes (see Appendix Table A3) we also estimated proportional models which we report in Appendix Table A15. The results based on a proportional model are qualitatively similar to the linear estimates in suggesting increases in all three measures and the implied proportional effects of the linear reduced form are roughly similar in magnitude to the proportional estimates. The proportional estimates also suffer from a lack of precision.

[37] The estimated increase in total days (0.1), while imprecise, is consistent with the estimated increase on the extensive margin. Given an average number of days (conditional on any admission) of 7.4 over our study period (see Table A3), the results on the extensive margin imply that, if the marginal admit had this average expected number of days, we would expect to see an increase in days of 0.16 (=0.021*7.4). This is well within the 95 percent confidence

percent for list charges, and 45 percent for the number of procedures, although only the result for procedures is statistically significant at the 10 percent level. The standardized treatment effect across all three measures indicates that insurance is associated with a 0.047 of a standard deviation increase in overall utilization (standard error = 0.026); the relatively small standardized effect in part reflects the large variance of the underlying variables.

We undertook several additional investigations, which are presented in detail in Appendix 3.1. First, we examined hospital utilization for seven conditions of interest and of reasonably high prevalence in our population: heart disease, diabetes, skin infections, mental disorders, alcohol and substance abuse, back problems, and pneumonia. We found a statistically significant increase in utilization (both extensive and total) only for heart disease. We also explored the impact of health insurance on the quality of outpatient care (admissions for ambulatory care sensitive conditions) and three measures of quality of care for inpatient care (not having an adverse patient safety event, not being readmitted within 30 days of discharge, and quality of hospital). We were unable to reject the null of no effects on either outpatient or inpatient quality, although our confidence intervals are extremely wide and do not allow us to rule out quantitatively large effects. Finally, we examined whether insurance was associated with a change in the proportion of patients going to public vs. private hospitals and were unable to detect any substantive or statistically significant differences.

5.1.2 Survey data

The survey data allow us to examine a broader range of utilization outcomes. Table 5 shows the results. Once again, we present results on both the extensive margin (left-hand panel) and on total utilization (right-hand panel). On both margins there are substantial and (mostly) statistically significant increases in prescription drugs and outpatient use. For example, the two stage least squares estimates suggest that insurance is associated with a 0.35 (standard error = 0.18) increase in the number of prescription drugs currently taken (corresponding to an approximately 15 percent increase) and a 1.08

interval of the point estimate of a 0.1 increase (standard error = 0.1) in the number of days; moreover, the marginal admit might be expected to have a shorter length of stay. Calculations are similar for admissions that do and do not originate through the emergency room.

(standard error = 0.18) increase in the number of outpatient visits (corresponding to an over 55 percent increase). The responses on the extensive margin may account for a large share of the increase in total utilization, although some of the increase in outpatient utilization – and perhaps in total drug utilization – likely reflects increased use among existing users (i.e. on the intensive margin).[38] There is no discernible impact of insurance on emergency room or inpatient hospital use on either margin, although the confidence intervals do not allow us to rule out potentially large effects and are consistent with the findings on hospital utilization from the administrative data.[39]

Overall, across the four utilization measures, we estimate that insurance is associated with a statistically significant increase in total utilization of 0.14 standard deviation, and in any utilization of 0.17 standard deviation. Since the four different components of utilization have very different expected costs, in the bottom row of Table 5 we make a back-of-the-envelope calculation of the increase in annual spending associated with insurance by weighting each type of utilization by its average cost among low-income publicly insured adults in the MEPS. This back-of-the-envelope calculation suggests that insurance is associated with a $778 (standard error = $371) increase in annual spending, or about a 25 percent increase relative to the implied control mean annual spending.

Although the longer-run impact of health insurance on health care utilization may differ from the one-year effects, we do not believe our one year estimates are capturing an initial, highly transitory surge of "pent up demand" for health care among the uninsured. In the survey, conducted about 13 months after

---

[38] The average number of drugs taken among control individuals who use drugs is 3.6, and the average number of outpatient visits among those with any is 3.2 (see Appendix 1.7, Table A11). If the marginal user on the extensive margin has utilization equal to the average among those who use, the response on the extensive margin would imply a total increase in drug use that is about 90 percent of what we estimate, and a total increase in outpatient visits that is about 60 percent of what we estimate. In fact, the marginal user probably uses less than average.

[39] To mimic the survey variables, we constructed measures in the discharge data of "any hospital visit in the last six months" and "number of hospital visits in the last six months", each defined for the six months prior to the individual's survey response date. For "any visit in the last six months", the reduced form estimate in the discharge data is 0.0015 (standard error = 0.0023), compared to 0.0022 (standard error = 0.0040) in the survey data; for "number of visits in the last six months" the reduced form estimate in the discharge data is 0.0025 (standard error = 0.0034) compared to 0.0062 (standard error = 0.0062) in the survey data. Interestingly, self reports do tend to overstate inpatient hospital use on average. The average for the controls of "any visit in the last six months" is 0.07 in the survey compared to 0.03 in the discharge data; for "number of visits in the last six months" these numbers are 0.10 and 0.03 respectively. There is not however, any difference in "reporting error" between treatments and controls.

insurance coverage began, all of the questions ask about current utilization or utilization over the last six months, not about utilization right after insurance began. Moreover, the evidence in Table 11 below from the survey conducted about 6 months after insurance began (so that a six month look back period likely captures the initial effects of insurance) shows no evidence of a larger initial utilization effect, suggesting that such "pent up" demand effects may not in fact be present.

Table 6 suggests insurance is also associated with an increase in compliance with recommended preventive care. We look at four different measures of preventive care: blood cholesterol checks, blood tests for diabetes, mammograms, and pap tests. [40] Overall, the results indicate a 0.3 standard deviation (standard error = 0.04)) increase in the probability of getting recommended preventive care. This reflects statistically significant increases in all four of the measures examined, including a 20 percent increase in the probability of ever having one's blood cholesterol checked, a 15 percent increase in the probability of ever having one's blood tested for high blood sugar or diabetes, a 60 percent increase in the probability of having a mammogram within the last year (for women 40 and over), and a 45 percent change in the probability of having a pap test within the last year (for women).

5.2 Financial strain

5.2.1 Administrative data

Table 7, Panel A, analyzes five measures of financial strain in the credit report data: whether the individual has had a bankruptcy, a lien, a judgment, a collection, or any credit account with a payment that is 30 days or more late ("a delinquency").[41] Broadly speaking, all are measures of unpaid bills or outstanding obligations that are likely to have a major negative impact on one's access to credit, at least in a general population (Avery et al, 2003). As the frequencies in Table 7 indicate, bankruptcies, judgments

---

[40] Our survey only distinguishes between preventive care received within the last year and that received more than a year ago. See Appendix 1.7 (and Figure A4) for detail. For blood cholesterol and diabetes checks we look at whether one has "ever" had these because the recommendation is to do it every 3 to 5 years (and our analysis is over an approximately one year horizon). For pap tests and mammograms we look within the last year since the recommendation at the time of the study was for these to occur annually, and we look only within the relevant populations (women, and women over 40 respectively).

[41] Delinquencies are mechanically zero for the one quarter of our sample who has no open credit over our study period.

and liens capture relatively extreme "right tail" events (ranging from 1 to 6 percent frequency over our 16-month sample period), while collections and delinquencies are much more common (about 50 percent and 37 percent frequency, respectively).[42]

The average standardized treatment effect suggests no evidence of a decline in financial strain across all these measures; the point estimate is of a statistically insignificant increase in financial strain associated with health insurance of 0.009 standard deviations (standard error = 0.019). Four of the five measures show no significant change. However there is evidence that health insurance is associated with a decline in the probably of having any unpaid bills sent to collection of 4.8 percentage points (standard error = 0.016), or about 10 percent relative to the control mean. This result is highly statistically significant, even after adjusting for multiple tests. One reason that we may see a decline in collections but not in the other measures is that a collection is less of a "right tail" event and occurs with less of a lag following an unpaid bill than a judgment, lien or bankruptcy.[43] Another possibility is that there is more of an impact of health insurance on collections since, both in our population and in the general population, collections are disproportionately medical.[44]

We are able to decompose the presence and size of collections into medical (Panel B) and non-medical (Panel C) components; this decomposition is not feasible for the other measures. The decline in overall collections shown in Panel A is primarily (or perhaps entirely) driven by a decline in medical collections. We find declines in both the existence and magnitude of medical collections. For example, health insurance is associated with a decline in the probability of having medical collections of 6.4

---

[42] With the exception of the bankruptcy rate which is about the same in the general Oregon population, all of the measures of financial strain are about two to four times more common in our study population than in the general Oregon population (see Appendix Table A8).

[43] For example, discussions with hospitals in Oregon suggest that it will take about 3 to 4 months before an unpaid hospital bill is sent to a collection agency. Similarly it will take time to decide to seek a judgment and then to win a judgment. Therefore while we analyze data through September 2009, our measures may not fully capture the financial strain incurred through that end date.

[44] In a general population, collections for medical bills are the single most common kind of collection, followed by utility bills (Avery et al., 2003). In our study population over our study period, about 28 percent of our sample has a medical collection, and about 39 percent has a non-medical collection; on average in our sample, the unconditional amount owed in collections is about $2,000 for medical collections and $2,700 for non-medical collections. It is possible that some "non-medical" collections may also reflect medical debt (for example, if one charges one's medical bills to a credit card and then does not pay the credit card bill and the credit card company tries to collect).

percentage points (standard error = 1.6), or about 25 percent relative to the control mean, and a decline in

the average (unconditional) amount owed in medical collections of $390 (standard error = $177), or about

20 percent relative to the control mean. The corresponding estimates for non-medical collections are

substantially smaller and statistically insignificant.

These results are subject to some potential limitations, discussed in more detail in the Appendices.

First, not all collections are reported to the credit bureaus, although our investigations did not suggest any

reason to suspect reporting to be correlated with insurance status (See Appendix 1.6). Second, in theory

health insurance might affect access to revolving credit itself, which could complicate interpretation of

measures of financial strain based on late payments for revolving credit (i.e. delinquencies); however we

found no evidence of effects on access to credit (Appendix 3.2). Third, many of the measures capture

only right tail events and with a substantial lag. Last, credit reports do not capture the use of informal or

"non-traditional" credit sources – which may be particularly important in a low-income population.[45]

Given these limitations, the survey measures of financial strain are a useful complement to the credit

report measures.

5.2.2 Survey data

Table 8 reports results for four measures of financial strain: whether the respondent has any out-of-

pocket medical expenditures in the last six months, whether the respondent currently owes money for

medical expenses, whether the respondent had to borrow money (or skip paying other bills or pay them

late) to pay medical expenses in the last six months, and whether the respondent has been refused medical

treatment because of medical debt in the last six months. We find a statistically significant decline in all

four survey measures of financial strain, including, for example, a 20 percentage point (35 percent)

decline in the probability of having out of pocket expenses and a 15 percentage point (40 percent) decline

---

[45] For example, in the 2007 Survey of Consumer Finances, of individuals who report having any debt (which is about three-quarters of the total population but only about half of those below the poverty level), about 10 percent of the debt of those below the poverty level is owed to "individual lenders" (e.g. relatives and friends) which would not be captured in credit reports, compared to only 3 percent of the debt in the overall population (authors' calculations).

in the probability of having to borrow money or skip paying other bills to pay medical expenses. [46] The average standardized treatment effect indicates that insurance is associated with a 0.3 standard deviation (standard error = 0.035) decline in these measures of financial strain.

The results suggest that the financial incidence of Medicaid coverage is on both the newly insured and their medical providers. The declines in out of pocket expenses and in reported difficulty paying non medical bills point to direct financial benefits to the newly insured. At the same time, both the survey data indicating a reduction in unpaid medical bills and the credit report data indicating a reduction in medical collections point to increased revenue for providers, since a large share of medical debt is never paid.[47]

For risk averse consumers, the largest welfare gains from any consumption-smoothing effects of insurance come from reducing extreme negative shocks to consumption. Although we cannot measure consumption directly, following Finkelstein and McKnight (2008) we estimate quantile regression models of the reduced form equation (1) to examine the impact of health insurance on the quantiles of the distribution of out-of-pocket medical expenditures. As expected given the comprehensive nature of Medicaid coverage, Figures 1a and 1b indicate that selection by the lottery is associated with declines in out-of-pocket spending at all the (non-zero) quantiles of the distribution. The same appears true for self-reported medical liabilities (Figures 2a and 2b), and the decline in medical collections is particularly concentrated in the right tails of the distribution (Figures 3a and 3b).

---

[46] These results imply that about 35 percent of those covered by OHP still have out-of-pocket medical expenses. The control group reports, on average, $307 in semi-annual out of pocket medical expenses; the IV estimate of the impact of insurance on these expenses is -122 (standard error = 43), implying that those covered by OHP average $185 in semi-annual out of pocket medical expenses. Our casual impression from focus groups is that these reflect some combination of continued scheduled payments on prior debts, uncovered services (primary dental), reporting of monthly premiums as out-of-pocket medical expenses, and perhaps including travel costs to the medical provider. It is also possible that some individuals report out of pocket medical spending for other family members (even though the question directed individuals to report only expenditures on themselves). Gross and Notowidigdo (2011) similarly find evidence of reported out of pocket spending on Medicaid recipients in the Medical Expenditure Panel Survey.

[47] In our population we estimate that less than 2 percent of medical collections are ultimately paid. Because collections may be paid with as substantial lag, we computed this statistic by looking at collections that originated between 2005 and 2007 and tracking whether they were paid by the end of September 2009. Across all types of collections, we similarly estimate that about 3 percent are paid. Avery et al (2003) estimate, for a general population, that about 11 percent of medical collections are paid off.

5.3 Health

Table 9 shows our estimates of the impact of health insurance on health. We have one measure of health from administrative data, namely mortality, which we measure from the notification date through September 30th, 2009. Mortality – although important and objectively measured – is very low in our population; only about 0.8 percent of the controls die over the 16 month study period. Not surprisingly, Panel A shows that we do not detect any statistically significant improvement in survival probability.

Panel B analyzes seven different measures of self-reported health from the survey data. The first two use the question about self-reported health (fair, poor, good, very good, or excellent) to construct two binary measures: (1) self-reported health good, very good or excellent (55 percent of the population) and, (2) to examine "tail" behavior, self-reported health not poor (i.e. 86 percent of the population). The other measures are: (3) whether self-reported health status is about the same or gotten better over last six months (vs. gotten worse), (4) the number of days in good physical health in last month (0-30), (5) the number of days not impaired by physical or mental health in the last month (0-30), (6) the number of days in good mental health in the last month (0-30), and (7) whether the respondent screened negative for depression. Many of these measures capture both physical and mental health; the last two, however, capture only mental health.

The results in Panel B indicate that insurance is associated with statistically significant improvements in each of the seven measures. On average, our results suggest that health insurance is associated with a 0.2 standard deviation improvement in self-reported health (standard error = 0.04). This includes, among other things, an increase in the probability of screening negative for depression of 7.8 percentage points (standard error = 2.5) or about 10 percent relative to the control mean, and an increase in the probability of reporting one's health as good, very good, or excellent of 13 percentage points (standard error = 2.6), or about 25 percent relative to the control mean.

There is thus an overwhelming sense from the survey outcomes that individuals feel better about their health and, as we come to below, their interactions with the health care system. Given the subjective nature of the responses, it is more difficult to judge with the available data whether the results reflect

27

improvements in actual, physical health. For mental health, the self-reported and subjective nature of the questions is less of an issue, since diagnosis of depression, by its nature, relies on such self-reports; the depression screen we use correlates highly with clinical diagnoses of depression (Kroenke et al., 2003). However, the self-reported physical health measures could reflect a more general sense of improved well-being rather than actual improvements in objective health. (A priori we were concerned that by increasing contact with the health care system, health insurance would cause individuals to learn more about their health problems (e.g. a doctor would tell a person who had not known it that they diabetes or high blood pressure) and thus could cause them to report themselves to be in worse health; to the extent this happens, it does not outweigh effects in the opposite direction.)

There is evidence of several mechanisms by which health insurance could have improved objective, physical health. Besides the previously documented increase in health care utilization and compliance with recommended preventive care, the survey data also indicate that insurance is associated with statistically significant increased self-reported access to care (Table 10, Panel A) and perceived quality of care (conditional on receipt) (Table 10, Panel B). [48]

However, there is also evidence that a substantial part of the estimated improvements in self-reported physical health may reflect a general sense of improved well-being. In particular, Table 11 compares reduced form estimates from our main survey to reduced form estimates from the initial survey which we fielded on average about 2.6 months after random assignment and about 1 month after coverage was approved. [^] As we would expect given this timing, there is no evidence of an increase in health care utilization in this earlier survey. However, there is evidence of an improvement in self-reported health of about two-thirds the magnitude of our main survey estimates from more than a year later; this reflects statistically significant improvements in all of the individual health measures available in the initial

---

[48] We also examined the impact of health insurance on two "health behaviors" (smoking and a possible measure of exercise). We present these results in Appendix 3.3; their interpretation is not obvious. We find no evidence of a decline in the probability of smoking and while we find a substantial and statistically significant increase in the probability of reporting that one is more physically active compared to others one's age, it is difficult to know whether to interpret this as an increase in health-promoting behavior or in health.
[^] This analysis was not pre-specified.

survey (all those in the main survey except the depression screen). Given the limited time after coverage approval and the lack of any increase in health care utilization, it seems likely that this immediate improvement does not reflect changes in objective physical health.[49] Likewise, self-reported access to care also shows a statistically significant improvement in the initial survey (despite no evidence of an increase in utilization) of about 40 percent the magnitude of the later survey, which again seems more likely to reflect an improved outlook. Of course, it is also possible that some of these immediate improvements reflect "winning" effects that are less likely to be picked up in the estimates one year later

Consistent with an improved overall sense of well-being, there is evidence in the later survey of a substantial (32 percent) increase in self-reported overall happiness (Table 10, Panel C). It is, of course, difficult to know how much of the self-reported happiness improvement reflects feeling better about one's health, just as it is difficult to know how much of the self-reported health improvement reflects more general improvement in a sense of well being. Overall, the evidence suggests that people feel better off due to insurance, but with the current data it is difficult to determine the fundamental drivers of this improvement.

6. DISCUSSION: INTERPRETING AND EXTRAPOLATING THE RESULTS

Using a randomized controlled experiment design, we examined the approximately one year impact of extending access to Medicaid among a low-income, uninsured adult population. We found evidence of increases in hospital, outpatient, and drug utilization, increases in compliance with recommended preventive care, and declines in exposure to substantial out-of-pocket medical expenses and medical debts. There is also evidence of improvement in self-reported mental and physical health measures, perceived access to and quality of care, and overall wellbeing.

---

[49] There is also evidence of a decline in financial strain in the initial survey that is about 40 percent the magnitude of the analogous measures in the later survey. This seems sensible when one recalls (see Section 3.4) that for an individual who successfully obtained insurance through the lottery, coverage was applied retroactively to about one month prior to the approval date. Individuals in the initial survey had therefore been covered retroactively by about 2 months, or about one-third of the six-month look back period on the financial strain questions. In other words, individuals may not have changed their health care utilization behavior in any detectable way after only 1 months of coverage (out of the last 6), but the retroactive coverage for 2 months (of the last six) has had an impact on their finances.

These results speak to the approximately one-year impact of expanding Medicaid access. As discussed earlier, we do not believe the results reflect a transitorily large utilization response stemming from initial "pent up demand". However the results may be larger than what we would find if individuals were not at risk of losing their insurance (and being unable to re-apply for it) if they do not continue to meet the eligibility requirements. The effects of health insurance may also change over longer horizons. At approximately two years after random assignment, we conducted in-person interviews and health examinations on a subsample of our study population located in the Portland metropolitan area. Results from those data should help shed light both on the longer-run impacts of insurance coverage, and on the impact of insurance on more objective measures of physical health, including biometric measures.

One natural way to try to interpret the current results is via a cost-benefit analysis. One could compare the cost of public funds from Medicaid expenditures on the newly insured as well as the moral hazard cost of increased utilization to the benefits from reduced financial strain and from improved self-reported health. However monetizing the costs, and especially the benefits, would require – and likely would be quite sensitive to – a number of assumptions; we consider this beyond the scope of the current paper.[50] Instead, to help put our results – particularly their magnitudes – in context, in the remainder of this section we compare the results to what we would have obtained from observational estimates in this setting, and to the existing experimental and quasi-experimental estimates from other settings. We also comment on some issues in trying to extrapolate from our results to the likely impact of health insurance in other contexts.

6.1 Comparison to observational estimates

We take several complementary approaches to compare our experimental estimates to what we might have estimated using observational data. In Table 12 we compare differences in outcomes within our study population across individuals who differ in their insurance coverage. The first column replicates our previous IV estimates using the lottery as an instrument for insurance coverage, and the remaining

---

[50] An interesting issue that any cost-benefit analysis will have to grapple with is how to monetize any generalized improvements in overall well-being that are not directly tied to mental or physical health improvements but that our results suggest may be among the most important benefits of health insurance.

columns present various OLS comparisons of people with and without insurance within our full study population (column 2), our control group (column 3), and our treatment group (column 4); unlike our IV estimates, these OLS comparisons will capture the effect of endogenous take-up of Medicaid, which can include a medical provider's qualifying a person for coverage at the time they seek care. To provide observational estimates of the impact of insurance coverage from outside of our data, Table 13 analyzes survey data from the Behavioral Risk Factors Surveillance Study (BRFSS) and the National Health Interview Study (NHIS) for outcome measures that are identical or very comparable to those collected by our survey. We compare the outcomes for insured vs. uninsured adults aged 19 – 64 with income below 100 percent of poverty in the national surveys both without controls (column 2) and with a rich set of demographic controls (column 3).

We find marked differences between our experimental estimates and our attempts to approximate their observational analogs in Tables 12 and 13. There is a general pattern of larger estimated impacts of health insurance on health care utilization in the observational estimates relative to the experimental ones, and of opposite signed impacts of health insurance on health, with health insurance appearing to worsen health in the observational estimates. These differences suggest that, at least within a low-income population, individuals select health insurance coverage are in poorer health (and therefore demand more medical care), than those who are uninsured, as standard adverse selection theory would predict. The results also suggest that the OLS estimates of the reduction in financial strain caused by health insurance are lower than the IV estimates, suggesting that, within a low-income population, the insured are generally poorer than the uninsured; that may reflect the income eligibility ceilings for public health insurance.

6.2 Comparisons to existing experimental and quasi-experimental estimates

We also try to place our estimates broadly in the context of the existing experimental and quasi-experimental estimates, although naturally there are no strictly comparable estimates. The only other randomized experiment (the RAND) took place over 30 years ago and involved variation in generosity of coverage among the insured (rather than variation on the extensive margin) in a representative population;

due to the available variation, most of the quasi-experimental studies have focused on the elderly or on the young rather than on prime age adults (as our study does).

With these caveats in mind, our loose sense is that the magnitude of our experimental estimates of the impact of moving low-income adults from no insurance to Medicaid coverage on health care utilization, while large in absolute terms, may, if anything, be slightly smaller than existing estimates might have suggested. For example, the RAND experiment from the 1970s – which experimentally varied the generosity of coverage in a representative population under age 65 – found that moving from the least comprehensive insurance plan – which still offered considerable insurance coverage (with an average out-of-pocket expenditure share of about 30 percent and a frequently binding out-of-pocket maximum) – to full insurance ("free care") was associated with a 45 percent increase in annual spending, while our back-of-the-envelope calculation suggested that Medicaid was associated with a 25 percent increase in six month spending. The same insurance variation in RAND, which is much smaller than the variation we observe, also produced about a 75 percent increase in the number of annual outpatient visits, while we estimate that Medicaid is associated with about a 55 percent increase in the number of semi-annual outpatient visits. For hospital admissions, both studies find about a 30 percent increase in the probability of a hospital admission over approximately one year.[51] More recent quasi-experimental work suggests that Medicare coverage is associated with about a 10 percent increase in the annual probability of admission for 65 year olds; this appears to reflect some combination of the 10 percentage point increase in insurance coverage associated with Medicare as well as the increase in total insurance coverage among the previously insured (Card et al., 2008).[52]

Most quasi-experimental estimates of the impact of health insurance on health have focused on mortality as an objective and readily available – if incomplete – measure of health. Yet mortality is a particularly poor measure for a prime age adult population, given the very low baseline mortality rate; as

---

[51] RAND estimates are from Newhouse et al. (1993) page 41.

[52] Like us, recent regression discontinuity estimates of the impact of "aging" onto insurance at 65 (Card et al., 2008) or off of it at 19 (Anderson et al., 2010) find that the impact of insurance on hospital admissions is disproportionately concentrated in admissions that do not come through the emergency room.

noted, most quasi-experimental estimates focus on the young or the old, rather than prime age adults. Like us, the RAND Experiment found no evidence of an impact of insurance on adult mortality (Newhouse et al, 1993 page 211).

Our findings regarding the impact of Medicaid on self-reported health may be larger than what might have been expected based on the (few) existing experimental and quasi-experimental estimates. RAND found no evidence of an impact of insurance generosity on adult self-reported general health or adult mental health (Newhouse et al., page 209); by contrast, our estimates suggested that the move from no insurance to Medicaid produced large (approximately two tenths of a standard deviation) improvements in adult self-reported general health, including an approximately 10 percent decline in the probability of a positive depression screen (see Table 9). For a somewhat older population – those aging onto Medicare at age 65 –quasi-experimental evidence suggests that health insurance is associated with mortality reductions for a severely ill subset of the population but does not have a discernible effect in the general population (Card et al., 2008, 2009), and that it has positive –but small and imprecisely estimated – effects on self-reported health (Card et al., 2004).

Compared to the voluminous observational literature and the sizable quasi-experimental literature on the impact of health insurance on health care utilization and health outcomes, there is a remarkable dearth of analysis of the impact of health insurance on risk smoothing, arguably its primary purpose. The RAND health insurance experiment did not analyze the impact of health insurance on exposure to out-of-pocket medical expenditure risk. The only existing experimental evidence we know of is a 2005 experiment in Mexico that randomized the provision of health insurance and of improvements in health facilities in selected communities; 10 months later, the study found that households in the insured communities had lower average out-of-pocket expenditures and were less likely to have out-of-pocket costs of more than 30 percent of their post-subsistence income (King et al., 2009). In the U.S., we know of only three quasi-experimental studies of the impact of health insurance on risk exposure. Finkelstein and McKnight (2008) and Englehardt and Gruber (2010) examine the impact of the introduction of Medicare (Parts A and B) and of Medicare Part D, respectively, on the distribution of out-of-pocket medical expenditures. Like us,

both studies find that insurance coverage is associated with a pronounced decline in the right tail of the out-of-pocket expenditure distribution. Unlike us, Gross and Notowidigdo (2011) estimate that Medicaid expansions are associated with a decline in personal bankruptcies, although their expansions cover individuals at slightly higher income (around 100 to 200 percent of FPL).

6.3 Extrapolation to other contexts

A natural instinct is to try to generalize our experimental estimates to other contexts, including the planned 2014 Medicaid expansions. Any such attempt comes with important caveats. First, by their nature, our findings speak to the partial equilibrium effects of covering a small number of people, holding constant the rest of the health care system. In particular, the lottery we studied covered about 10,000 low-income uninsured adults, relative to a total Oregon population of about 3.8 million, including about 650,000 uninsured and about 200,000 low-income adult uninsured.[53] Our estimates are therefore difficult to extrapolate to the likely effects of much larger health insurance expansions, in which there may well be supply side responses from the health care sector (Finkelstein, 2007).

Second, our results are specific to a population of low-income, uninsured adults in Oregon who expressed interest in obtaining health insurance (by signing up for the lottery). This group is not representative of the low-income uninsured adults in the rest of the United States on a number of observable (and presumably unobservable) dimensions. One striking difference is that our study population has more whites and fewer African-Americans (by about 15 percentage points each) than the general low-income, uninsured adult, US population. It is also somewhat (4 to 5 years) older and on some measures appears to be in somewhat worse self-reported health (Allen et al. 2010).[54] These differences are amplified when focusing on compliers, who, relative to the overall lottery population, are somewhat older, more white, in worse health, and of lower socio-economic status (as proxied by education and having

---

[53] Estimates are from various public sources from 2008 and 2009 and can be found here: www.statehealthfacts.org (last accessed February 2011). "Low-income adults" refers to adults aged 19-64 below 100 percent of the federal poverty level.

[54] Allen et al (2010) also note that the overall health care environment in Oregon is similar in many ways to the nation as a whole, including the market share of public hospitals, the magnitude of uncompensated care as a fraction of hospital charges, and physicians per capita.

revolving credit at the time of the lottery); although we examined heterogeneity in treatment effect by these and other observables, we lacked power to draw precise inferences (see Appendix 3.4 and Table A23 for details).

Although our setting shares some features with the planned 2014 Medicaid expansions, it might not with other insurance expansions. The insurance offered in this setting was free or heavily subsidized, so our estimates capture the combined effect of insurance at actuarially fair prices and the wealth effect from the large premium subsidy; average annual OHP Standard expenditures – and hence an actuarially fair premium – are about $3,000, which is quite high relative to the actual annual premium of $0 to $240. Presumably, however, most health insurance coverage for this type of low income population would also be heavily or completely subsidized. Our results suggest that Medicaid provides benefits to this population above and beyond the non-Medicaid alternatives that exist through various safety net options.

REFERENCES

Agency for Healthcare Research and Quality, US Department of Health and Human Services. 2001. *Annual Report of the National CAHPS Benchmarking Database 2000*.

Alatas, Vivi, Abhijit Banerjee, Rema Hanna, Benjamin A. Olken, and Julia Tobias. 2010. "Targeting the Poor: Evidence from a Field Experiment in Indonesia." NBER Working Paper 15980.

Allen, Heidi, Katherine Baicker, Amy Finkelstein, Sarah Taubman, Bill Wright, and the Oregon Health Study Group. 2010. "What the Oregon Health Study Can Tell Us About Expanding Medicaid." *Health Affairs* Volume 29 No 8: 1498-1506.

Anderson, Michael.2008. "Multiple Inference and Gender Differences in the Effects of Early Intervention: A Reevaluation of the Abecedarian, Perry Preschool, and Early Training Projects." *Journal of the American Statistical Association.* 103(484): 1481-1495.

Anderson, Michael, Carlos Dobkin, and Tal Gross. 2010. "The Effect of Health Insurance Coverage on the Use of Medical Services." NBER Working Paper 15823.

Avery, Robert, Paul Calem and Glenn Canner. 2003. "An Overview of Consumer Data and Credit Reporting." *Federal Reserve Bulletin.*

Beebe, T. J., M. E. Davern, et al. (2005). "Increasing response rates in a survey of Medicaid enrollees: the effect of a prepaid monetary incentive and mixed modes (mail and telephone)." Med Care 43(4): 411-4.

Brown, J. A., S. E. Nederend, et al. (1999). "Special issues in assessing care of Medicaid recipients." Med Care 37(3 Suppl): MS79-88.

Bodenheimer, Thomas. 1997. "The Oregon Health Plan – Lessons for the Nation." *New England Journal of Medicine,* 337: 651-656.

Card, David, Carlos Dobkin, and Nicole Maestas. 2004. "The Impact of Nearly Universal Health Care Coverage on Health Care Utilization and Health: Evidence from Medicare." NBER Working Paper 10365.

Card, David, Carlos Dobkin, and Nicole Maestas. 2008. "The Impact of Nearly Universal Coverage on Health Care Utilization: Evidence from Medicare." *American Economic Review* Vol 98, Issue 5, pp 2242-58.

Card, David, Carlos Dobkin and Nicole Maestas. 2009. "Does Medicare Save Lives?" *Quarterly Journal of Economics* Vol 124, No 2, pp. 597-636.

Carlson, M. J., J. DeVoe, et al. (2006). "Short-term impacts of coverage loss in a Medicaid population: early results from a prospective cohort study of the Oregon Health Plan." Ann Fam Med 4(5): 391-8.

Cutler, David and Jonathan Gruber. 1996. "Does public insurance crowd out private insurance?" *Quarterly Journal of Economics.* 1996; 111(2): 391-430.

Doyle, Joesph. 2005. "Health Insurance, Treatment, and Outcomes: Using Auto Accidents as Health Shocks." *Review of Economics and Statistics* 87(2): 256-270.]

Engelhardt, G. and Gruber, J. (2010) "Medicare Part D and the Financial Protection of the Elderly." *NBER Working Paper 16155.*

Finkelstein, Amy. 2007. "The Aggregate Effects of Health Insurance: Evidence from the Introduction of Medicare." *Quarterly Journal of Economics* 122(1): 1-37.

Finkelstein, Amy and Robin McKnight. 2008. "What Did Medicare Do?" The Initial Impact of Medicare on Mortality and Out-of-pocket Medical Spending." *Journal of Public Economics* 92: 1644-1669.

Fowler, F. J., Jr., P. M. Gallagher, et al. (1999). "Comparing telephone and mail responses to the CAHPS survey instrument. Consumer Assessment of Health Plans Study." Med Care 37(3 Suppl): MS41-9.

Gallagher, P. M., F. J. Fowler, Jr., et al. (2005). "The Nature of Nonresponse in a Medicaid Survey: Causes and Consequences." Journal of Official Statistics 21(1): 73-87.

Getzen, Thomas. 2000. "Health Care is an individual necessity and a national luxury: applying multilevel decision models to the analysis of health care expenditures." *Journal of Health Economics.* 19: 259-270.

Gross, Tal and Matthew J. Notowidigdo. 2011. "Health Insurance and the Consumer Bankruptcy Decision: Evidence from the Expansion of Medicaid." *Journal of Public Economics* Volumne 95, Issues 7-8, August, pages 767-778.

Hartz, A. J., R. Noyes, et al. (2000). "Unexplained symptoms in primary care: perspectives of doctors and patients." Gen Hosp Psychiatry 22(3): 144-52.

Hoynes, Hilary and Diane Schanzenbach. 2009. "Consumption Responses to In-Kind Transfers: Evidence from the Introduction of Food Stamps." *American Economic Journal – Applied Economics.* 1(4): 109-139.

Imbens, G. W. and J. D. Angrist (1994): "Identification and Estimation of Local Average Treatment Effects," *Econometrica*, 62, 467-475.

Institute of Medicine. 2003. Hidden Costs, Value Lost: Uninsurance in America. Washington, DC. The National Academies Press.

Kaiser Family Foundation (2011). "Medicaid and State Funded Coverage Income Eligibility Limits for Low-Income Adults, 2011". http://www.statehealthfacts.org/comparereport.jsp?rep=54&cat=4

Kaiser Family Foundation (2010a). "Implementation Timeline". Accessed October 27, 2010 at http://healthreform.kff.org/timeline.aspx.

Kaiser Family Foundation (2010b) "Focus on Health Reform: Summary of New Health Reform Law", Publication 8061. Accessed February 22, 2011at http://www.kff.org/healthreform/8061.cfm

King, Gary, Emmanuela Gakidou, Kosuke Imai, Jason Lakin, Ryan T Moore, Clayton Nall, Nirmala Ravishankar, Manett Vargas, Matha Maria Tellez-Rojo, Juan Eugenio Hernandez Avila, Mauricio-Hernandez-Avila, Hector Hernandez Llamas. 2009. "Public policy for the poor? A randomized assessment of the Mexican universal health insurance program." *The Lancet* Vol 373, April 25, 1447 – 1454.

Kling, Jeffrey and Jeffrey Liebman. 2004. "Experimental Analysis of Neighborhood Effects on Youth." Unpublished Mimeo. Available at: http://www.nber.org/mtopublic/483.pdf

Kling, Jeffrey, Jeffrey Liebman and Lawrence Katz. 2007. "Experimental Analysis of Neighborhood Effects." *Econometrica* 75(1): 83-119.

Kroenke K, Spitzer RL, Williams JB. 2003. The Patient Health Questionnaire-2: validity of a two-item depression screener. Med Care; 41:1284-92

Manning, Willard, Joseph Newhouse, Naihua Duan, Emmett Keeler, Arleen Leibowitz, and Susan Marquis, "Health Insurance and the Demand for Medical Care: Evidence from a Randomized Experiment." *American Economic Review* Vol 77, No 3 (1987): 251-277.

Newhouse, Joseph P, and the Insurance Experiment Group. (1993). *Free for All?:   Lessons from the RAND Health Insurance Experiment*. Harvard University Press, Cambridge, MA.

Oregon Department of Human Services (2008) OHP Standard Reservation List: Application Progress Report.

Oregon Office for Health Policy and Research (2009) Trends in Oregon's HealthCare Market and the Oregon Health Plan: A Report to the 75th Legislative Assembly. February.

Olken, Benjamin A., Junko Onishi, and Susan Wong, "Indonesia's PNPM Generasi Program: Interim Impact Evaluation Report," World Bank, 2010

Schaner, Simone. 2010. Intrahousehold Preference Heterogeneity, Commitment and Strategic Savings: Theory and Evidence from Kenya.". MIT Working Paper. http://econ-www.mit.edu/files/6221
U.S. GPO. 2010. "Patient Protection and Affordability Care Act (HR 3590). Washington, DC. Available at www.gpo.gov/fdsys/pkg/PLAW-111publ148/pdf/PLAW-111publ148.pdf.

Wallace, Neal, John McConnell, Charles Gallia and Jeanene Smith. 2008. "How effective are co-payments for reducing expenditures for low-income adult Medicaid beneficiaries? Experience from the Oregon Helath Plan." *Health Research and Educational Trust.*

Westfall, Peter H. and Young, Stanley S. 1993. Resampling-Based Multiple Testing: Examples and Methods for P-value Adjustment. New York: Wiley.

Zeckhauser, Richard. 1970. "Medical Insurance: A Case Study of the Tradeoff Between Risk Spreading and Appropriate Incentives." *Journal of Economic Theory*, Vol 2: 10-26.

Figure 1a and 1b: Distribution of out-of-pocket medical expenses, last six months (survey data)
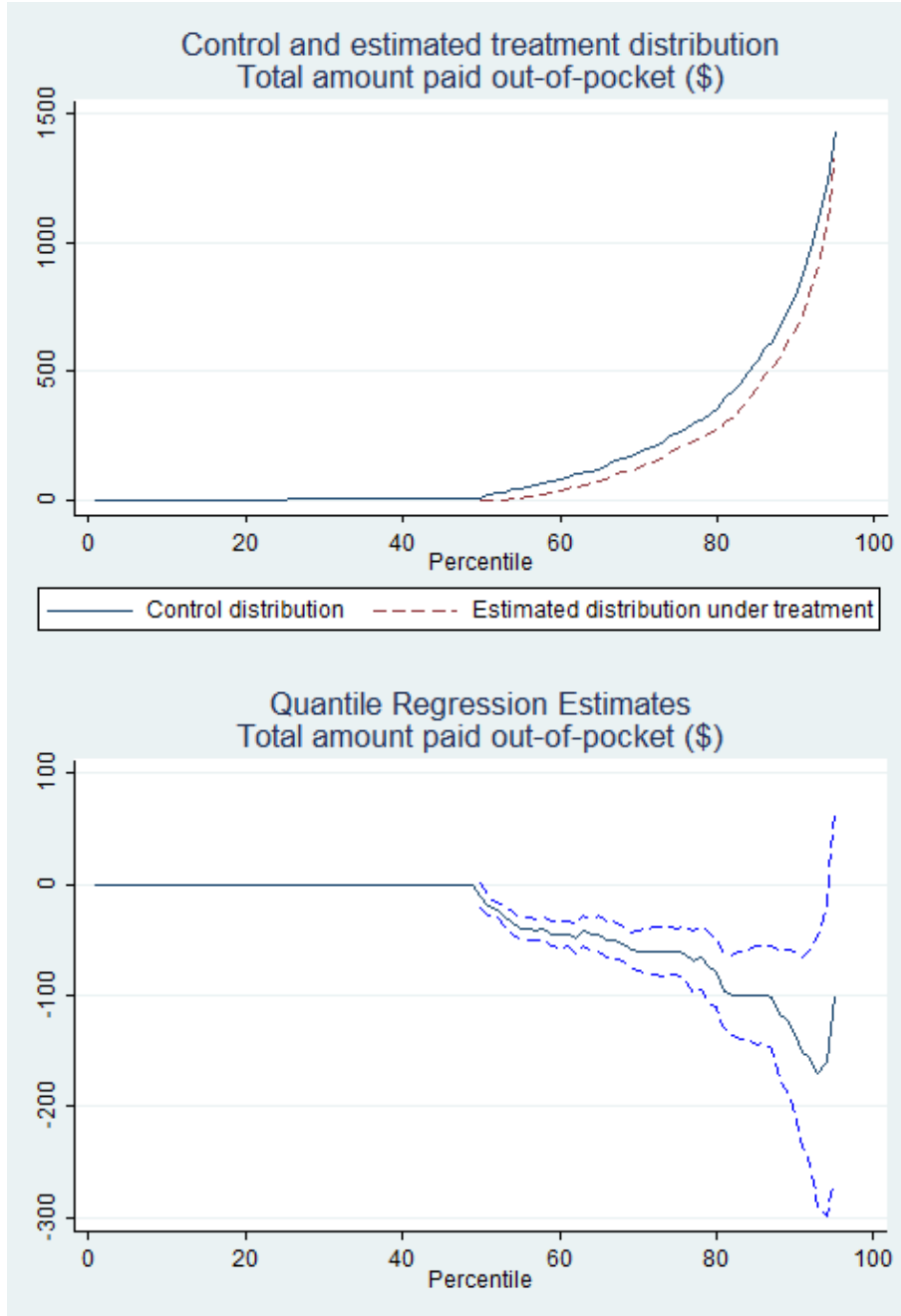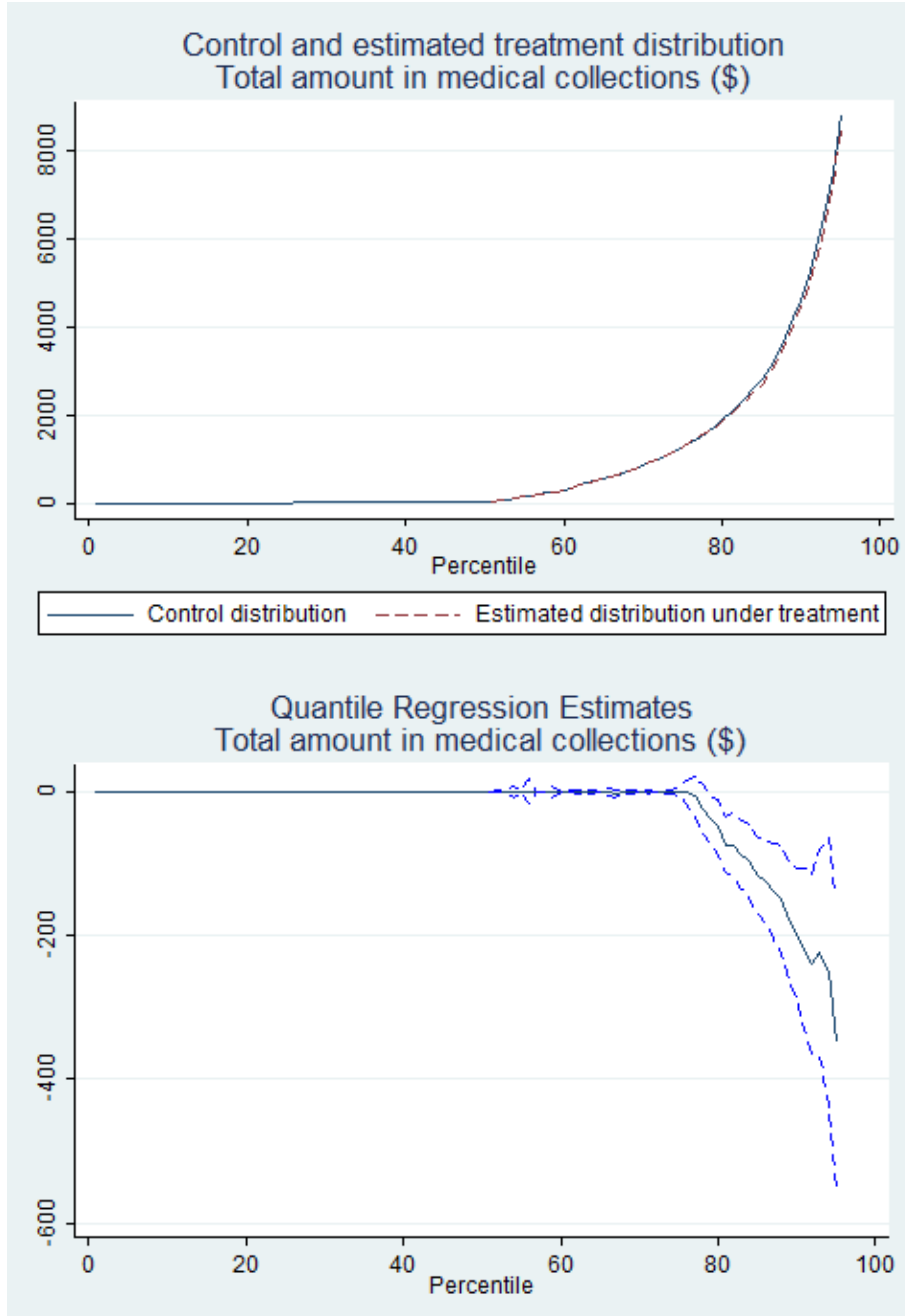

Notes: Figure 1a shows the distribution of out-of-pocket medical spending for controls, and the estimated distribution for treatments through the 95th quantile; the estimated distribution for treatments is the control distribution added to the beta on LOTTERY from the quantile estimation of the reduced form equation (1). Figure 1b plots the quantile estimates from equation (1) (along with their 95 percent confidence interval) starting from the smallest quantile that is non-zero in either the treatment or control distribution through the 95th quantile. The confidence intervals are calculated based on 500 bootstraps clustered on household. Data are from the sample of survey responders (N=24,012); all results use survey weights. Quantile estimation of equation (1) includes household size dummies, survey wave dummies, and the interaction of the two.

Figure 2a and 2b: Distribution of amount owed in medical debt (survey data)
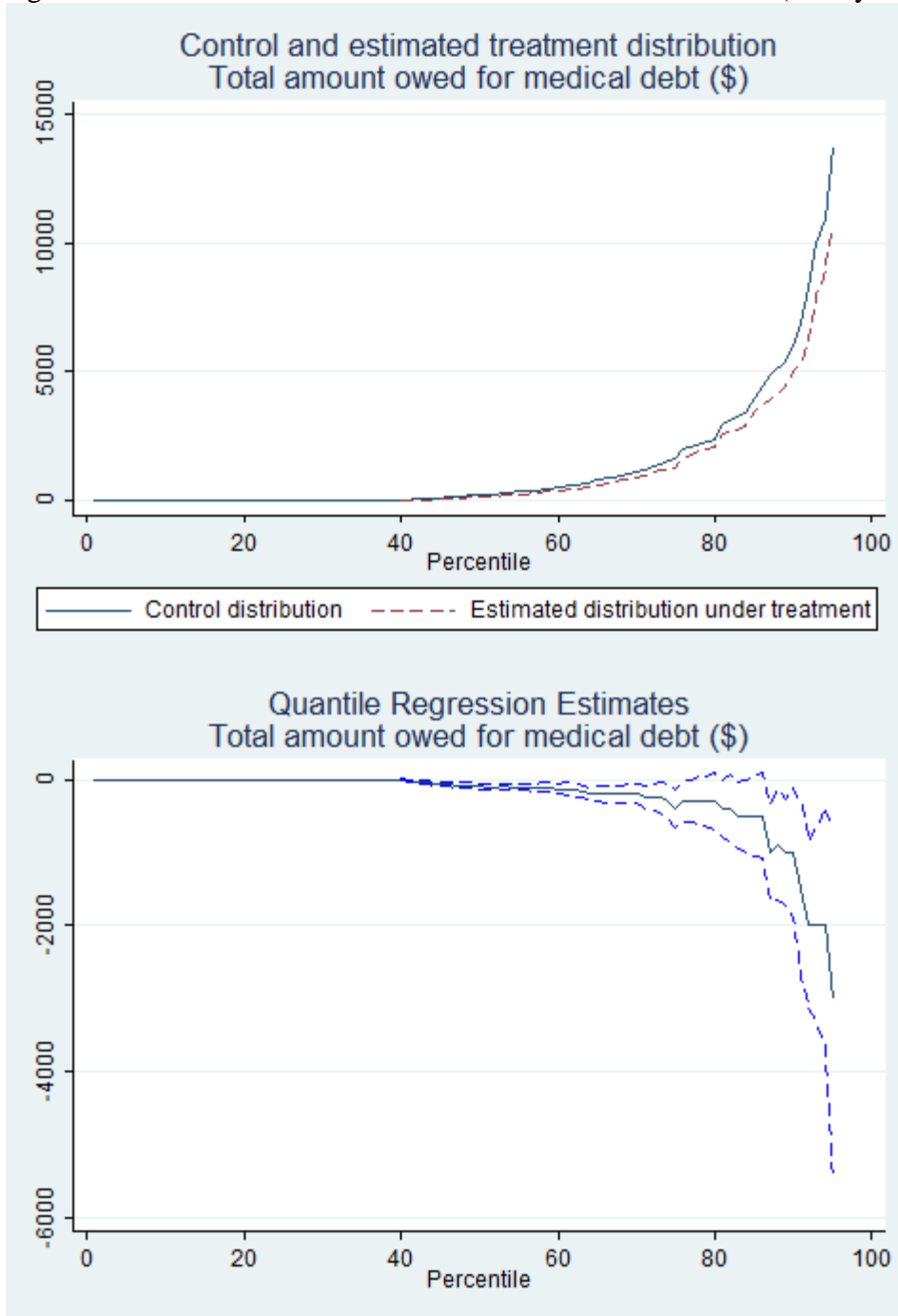
Notes: Figure 2a shows the distribution of medical debt for controls, and the estimated distribution for treatments through the 95th quantile; the estimated distribution for treatments is the control distribution added to the beta on LOTTERY from the quantile estimation of the reduced form equation (1). Figure 2b plots the quantile estimates from equation (1) (along with their 95 percent confidence interval) starting from the smallest quantile that is non-zero in either the treatment or control distribution through the 95th quantile. The confidence intervals are calculated based on 500 bootstraps clustered on household. Data are from the sample of survey responders (N=24,012); all results use survey weights. Quantile estimation of equation (1) includes household size dummies, survey wave dummies, and the interaction of the two.

Figure 3a and 3b: Distribution of medical collection amount owed (credit report data)

Notes: Figure 3a shows the distribution of amount owed in medical collections for controls, and the estimated distribution for treatments through the 95th quantile; the estimated distribution for treatments is the control distribution added to the beta on LOTTERY from the quantile estimation of the reduced form equation (1). Figure 3b plots the quantile estimates from equation (1) (along with their 95 percent confidence interval) starting from the smallest quantile that is non-zero in either the treatment or control distribution through the 95th quantile. The confidence intervals are calculated based on 500 bootstraps clustered on household. Data are from the sample of survey responders (N=49,545). Quantile estimation of equation (1) includes household size dummies, lottery draw dommies, and the individual's pre-lottery medical collection amount owed.

Table 1: Demographic Characteristics of Study Population (Control Group)

Variable	Control mean	Variable	Control mean
Panel A: Full sample			
Sex		*Language*	
% Female	0.557	% English preferred	0.922
Age		*Zip code level variables*	
% 50-64	0.267	% MSA	0.773
% 20-50	0.733	Zip code median household incom	$39,265
Panel B: Survey responders only			
Race		*Health Status*	
% White	0.820	Ever diagnosed with:	
% Black	0.038	Diabetes	0.175
		Asthma	0.276
		High Blood Pressure	0.399
Ethnicity		Emphysema or Chronic Bronchiti	0.129
% Spanish/Hispanic/Latino	0.123	Depression (screen positive)	0.557
Education		*Income (% federal poverty line)*	
% Less than High School	0.177	<50%	0.406
% High school disploma or GED	0.491	50-75%	0.138
% Vocational Training or 2-yr degree	0.220	75-100%	0.140
% 4 year college degree or more	0.112	100-150%	0.177
		Above 150%	0.139
Employment			
% don't currently work	0.551	*Insurance Coverage*	
% work <20 hrs per week	0.090	Any Insurance?	0.325
% work 20-29 hrs per week	0.099	OHP / Medicaid	0.117
% work 30+ hrs per week	0.259	Private Insurance	0.128
		Other	0.102
Average household income (2008) $	13,053	# of months of last 6 with insuranc	1.738

Notes: All statistics are reported for control individuals only. Panel A reports the control means for pre-randomization demographics taken from the lottery list (from January and February 2008) for the whole sample (N=45088 for controls). Age refers to age at the end of the study period. "English as preferred language" indicates whether the individual did not check a box requesting materials in a language other than English. Panel B reports control means of survey questions for survey responders (N = 11933 for controls), weighted using survey weights. "Household income" is gross household income (in $) for 2008 (before taxes and deductions but including any cash assistance or unemployment assistance received); it is reported in bins and we assign individuals the income at the mid- point of their bin (see Appendix Figure A4 for details). For the insurance questions, we code as "yes" if the respondent checked that insurance type box; since the survey allows one to check multiple boxes for types of insurance, the subgroups (OHP/Mediciad, private, and other) won't necessarily add up to "any". Private insurance includes employer and privately paid insurance; "Other" insurance includes "Medicare and other." We treat responses for insurance as missing if the responder checked "I don't know" or left all check boxes blank. We construct income relative to the federal poverty line based on self-reported income and self reported (total) number of household members. See Appendix 3 for more details.

Table 2: Treatment - Control Balance

	Control Mean (std dev) for full sample (1)	Difference between treatment and control		
		Full sample (2)	Credit report subsample (3)	Survey respondents subsample (4)
Panel A: Match / response rates				
Matched in September 09 credit data	0.663 (0.473)		-0.004 (0.004) [0.247]	
Responded to survey	0.506 (0.500)			-0.016 (0.007) [0.014]
Response time (in days)	53.0 (57.8)			1.638 (1.088) [0.132]
Panel B: Pre-randomization characteristics				
Lottery list variables				
F-statistic		1.286	0.553	0.574
[p-value]		[0.239]	[0.836]	[0.820]
Pre-randomization outcomes				
F-statistic		0.543	0.921	1.266
[p-value]		[0.844]	[0.518]	[0.281]
Both of the above				
F-statistic		0.915	0.793	0.782
[p-value]		[0.56]	[0.726]	[0.680]
N		74922	49980	23741

(Standard errors in parentheses)
[Per comparison p-values in square brackets]

Notes: In Panel A, we analyze match and response rates. The first column reports the mean and standard deviation for the control sample of the outcome shown in the left hand column. Columns 3 and 4 report estimated differences between treatments and controls for the dependent variable shown in the left hand column and the sample indicated in the column heading. Specifically they report the coefficient on LOTTERY based on estimating equation (1). All regressions include household fixed effects and cluster on household. In addition, in column (4) we include survey wave fixed effects and the interaction of survey wave fixed effects and household fixed effects and use survey weights. The full sample (i.e the sample used in the hospital discharge and mortality data) is used in column 3 and the entire survey sample is used in column 4. In Panel B we report the F-statistic and p-value from regressing multiple pre-randomization characteristics on LOTTERY in equation (1). "Lottery list variables" are common across all three samples and consist of nine demographic variables derived from information provided at the time of lottery sign up: year of birth; sex; whether English is their preferred language for receiving materials; whether the individual signed themselves up for the lottery or were signed up by a household member; whether the individual gave their address as a PO box; whether they signed up the first day the lottery list was open; the median household income in the 2000 census from their zip code; whether the zip code they gave is within a census-defined MSA; and whether they provided a phone number on sign up. "Pre randomization outcomes" are specific to the sample (we look at the hospital outcomes that we subsequently analyze for column 2, the credit report outcomes we subsequently analyze for column 3, and a few measures from each that approximate survey questions we subsequently analyze for column 4). More detail on the pre-randomization outcomes, the exact regression specifications, and the results for each variable analyzed in Panel B are presented in Appendix 3, Table A13.

Table 3: First Stage Estimates

		Full sample		Credit report subsample		Survey respondents	
		Control mean	Estimated FS	Control mean	Estimated FS	Control mean	Estimated FS
		(1)	(2)	(3)	(4)	(5)	(6)
(1)	Ever on Medicaid	0.141	0.256 (0.004)	0.135	0.255 (0.004)	0.135	0.290 (0.007)
(2)	Ever on OHP Standard	0.027	0.264 (0.003)	0.028	0.264 (0.004)	0.026	0.302 (0.005)
(3)	# of Months on Medicaid	1.408	3.355 (0.045)	1.352	3.366 (0.055)	1.509	3.943 (0.090)
(4)	On Medicaid, end of study period	0.106	0.148 (0.003)	0.101	0.151 (0.004)	0.105	0.189 (0.006)
(5)	Currently have any insurance (self report)					0.325	0.179 (0.008)
(6)	Currenty have private insurance (self report)					0.128	-0.008 (0.005)
(7)	Currently on Medicaid (self report)					0.117	0.197 (0.006)
(8)	Currently on Medicaid					0.093	0.177 (0.006)
(9)	Ever on Food Stamps	0.606	0.017 (0.003)	0.594	0.018 (0.004)	0.606	0.014 (0.005)
(10)	Food Stamp Benefits ($)	1694	58.5 (14.8)	1716	52.2 (18.5)	1776	63.7 (23.9)
	N		74,922		49,980		23,741

(Standard errors in parentheses.)

Notes: Even numbered columns report the coefficient and standard error on "LOTTERY" from estimating the first stage equation (4) with the dependent variable "INSURANCE" defined in the left hand column; odd numbered columns report the control mean for that measure of "INSURANCE". "Full sample" is the sample analyzed in the hospital discharge and mortality data. All regressions include dummies for household size and adjust standard errors for household clusters. The regressions in columns 2 and 4 also include lottery draw dummies; the regressions in column 6 also include dummies for survey wave and survey wave interacted with household size dummies, and use survey weights. The insurance measures are taken from the Medicaid enrollment administrative data except for those labeled "self report" (row 5 through 7) which are taken from the survey. In the survey, respondents could report various types of insurance; we define "private insurance" as employer or private insurance and "any insurance" as Medicaid, Medicare, employer, private or other insurance. In row 8 insurance is measured as being on Medicaid according to the state Medicaid enrollment data on the day the survey was returned. In Row 10 Food Stamp Benefits measure total household benefits received over the study period.

Table 4a: Hospital Utilization: Admission Probabilities (Administrative Data)

	Extensive Margin (Any)			
	Control Mean	Reduced Form	2SLS	p-values
	(1)	(2)	(3)	(4)
Panel A: All hospital admissions	0.067	0.005	0.021	[0.004]
	(0.250)	(0.002)	(0.007)	
Panel B: Admissions through ER	0.048	0.002	0.007	[0.265]
	(0.214)	(0.002)	(0.006)	
Panel C: Admissions not through ER	0.029	0.004	0.016	[0.002]
	(0.167)	(0.001)	(0.005)	

(Standard errors in parentheses)
[Per comparison p-values in square brackets]
{Family wise p-values in curly brackets}

Notes: Table investigates non-childbirth-related hospitalizations during the time period from notification date to August 31, 2009. Column 2 reports the coefficient and standard error on LOTTERY from estimating equation (1) by OLS. Column 3 reports the coefficient and standard error on INSURANCE from estimating equation (3) by IV. Column 4 reports the per-comparison p value and (where applicable) the family wise p-value across the three different measures of utilization used to create the standardized treatment effect. Standardized treatment effect reports results based on equation (2). All regressions include household size fixed effect, lottery draw fixed effects and the analogous outcome measure for the time period from January 1, 2008 through the notification date. All standard errors are clustered on the household. Sample consists of entire sample universe (N = 74,922).

Table 4b: Total Hospital Utilization (Administrative Data)

	Control Mean	Reduced Form	2SLS	p-values
	(1)	(2)	(3)	(4)
Panel A: All hospital admissions				
Days	0.498	0.026	0.101	[0.329]
	(3.795)	(0.027)	(0.104)	{0.328}
List Charges	2,613	258	1,009	[0.077]
	(19942)	(146)	(569)	{0.106}
Procedures	0.155	0.018	0.070	[0.031]
	(1.08)	(0.008)	(0.032)	{0.059}
Standardized Treatment effect		0.012	0.047	[0.073]
		(0.007)	(0.026)	
Panel B: Admissions through ER				
Days	0.299	0.023	0.089	[0.183]
	(2.326)	(0.017)	(0.067)	{0.187}
List Charges	1,502	163	636	[0.091]
	(12749)	(96)	(376)	{0.171}
Procedures	0.081	0.008	0.031	[0.135]
	(0.694)	(0.005)	(0.021)	{0.187}
Standardized Treatment effect		0.011	0.044	[0.100]
		(0.007)	(0.027)	
Panel C: Admissions not through ER				
Days	0.199	0.003	0.013	[0.841]
	(2.38)	(0.017)	(0.065)	{0.842}
List Charges	1,110	98	384	[0.281]
	(12422)	(91)	(356)	{0.383}
Procedures	0.075	0.010	0.038	[0.080]
	(0.708)	(0.006)	(0.022)	{0.162}
Standardized Treatment effect		0.008	0.030	[0.254]
		(0.007)	(0.026)	

(Standard errors in parentheses)
[Per comparison p-values in square brackets]
{Family wise p-values in curly brackets}

Notes: Table investigates non-childbirth-related hospitalizations during the time period from notification date to August 31, 2009. All outcomes are measured unconditionally (i.e. are not conditional on admission). Column 2 reports the coefficient and standard error on LOTTERY from estimating equation (1) by OLS. Column 3 reports the coefficient and standard error on INSURANCE from estimating equation (3) by IV. Column 4 reports the per-comparison p value and (where applicable) the family wise p-value across the three different measures of utilization used to create the standardized treatment effect. Standardized treatment effect reports results based on equation (2). All regressions include household size fixed effect, lottery draw fixed effects and the analogous outcome measure for the time period from January 1, 2008 through the notification date. All standard errors are clustered on the household. Sample consists of entire sample universe (N = 74922).

Table 5: Health Care Utilization (Survey Data)

	Extensive Margin (Any)				Total Utilization (Number)			
	Control Mean	Reduced Form	2SLS	p-values	Control Mean	Reduced Form	2SLS	p-values
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Rx drugs currently	0.637	0.025	0.088	[0.002]	2.318	0.100	0.347	[0.049]
	(0.481)	(0.008)	(0.029)	{0.005}	(2.878)	(0.051)	(0.176)	{0.137}
Outpatient visits last six months	0.574	0.062	0.212	[<0.0001]	1.914	0.314	1.083	[<0.0001]
	(0.494)	(0.007)	(0.025)	{<0.0001}	(3.087)	(0.054)	(0.182)	{<0.0001}
ER visits last six months	0.261	0.006	0.022	[0.335]	0.47	0.007	0.026	[0.645]
	(0.439)	(0.007)	(0.023)	{0.547}	(1.037)	(0.016)	(0.056)	{0.643}
Inpatient Hospital admissions last six months	0.072	0.002	0.008	[0.572]	0.097	0.006	0.021	[0.311]
	(0.259)	(0.004)	(0.014)	{0.570}	(0.4)	(0.006)	(0.021)	{0.510}
Standardized treatment effect		0.050	0.173	[<0.0001]		0.040	0.137	[0.0003]
		(0.011)	(0.036)			(0.011)	(0.038)	
Annual spending [a]					3,156	226	778	[0.037]
						(108)	(371)	

(Standard errors in parentheses)
[Per comparison p-values in square brackets]
{Family wise p-values in curly brackets}

Notes: Hospital admissions exclude childbirth. Columns 2 and 6 report the coefficient and standard error on LOTTERY from estimating equation (1) by OLS. Columns 3 and 7 report the coefficient and standard error on INSURANCE from estimating equation (3) by IV. Columns 4 and 8 report the per-comparison p value and the family wise p-value across the four different measures of utilization used to create the standardized treatment effect. Standardized treatment effect reports results based on equation (2). All regressions include household size fixed effects, survey wave fixed effects, and the interaction between the two. All standard errors are clustered on the household and all regressions are weighted using survey weights. Sample consists of survey responders (N = 23741).

[a] To calculate the implied spending effects associated with the estimated utilization effects we use data from the 2002-2007 (pooled) Medical Expenditure Panel Survey (MEPS) on expenditures of all nonelderly (19-64) adults below 100 percent of poverty who are publicly insured. This gives us a total sample of over 7,500 individuals. We use their expenditures (all inflated with the CPI-U to 2007 dollars) to calculate average expenditures per outpatient visit, average expenditures per ER visit, average expenditures per inpatient visit (for visits not related to childbirth), and average semi-annual (six month) spending on prescription drug. All spending is total expenditures (i.e. not just insured) expenditures. The underlying costs are $150 per outpatient visit, $435 per ER visit, $7,523 per inpatient visit, and $156 six month expenditure per current prescription drug; we multiply these all by two to get annual costs.

Table 6: Compliance with Recommended Preventive Care (Survey Data)

	Control Mean	Reduced Form	2SLS	p-values
	(1)	(2)	(3)	(4)
Blood cholesterol checked (ever)	0.625	0.033	0.114	[<0.0001]
	(0.484)	(0.007)	(0.026)	{<0.0001}
Blood tested for high blood sugar / diabetes (ever)	0.604	0.026	0.090	[0.0004]
	(0.489)	(0.007)	(0.026)	{<0.0001}
Mammogram within last 12 months (women >=40)	0.298	0.055	0.187	[<0.0001]
	(0.457)	(0.012)	(0.04)	{<0.0001}
Pap test within last 12 months (women)	0.406	0.051	0.183	[<0.0001]
	(0.491)	(0.01)	(0.034)	{<0.0001}
Standardized treatment effect		0.087	0.300	[<0.0001]
		(0.012)	(0.041)	

(Standard errors in parentheses)
[Per comparison p-values in square brackets]
{Family wise p-values in curly brackets}

Notes: Column 2 reports the coefficient and standard error on LOTTERY from estimating equation (1) by OLS. Column 2 reports the coefficient and standard error on INSURANCE from estimating equation (3) by IV. Column 4 reports the per-comparison p value and the family wise p-value across the four different preventive care measures used to create the standardized treatment effect. Standardized treatment effect reports results based on equation (2). All regressions include household size fixed effects, survey wave fixed effects, and the interaction between the two. All standard errors are clustered on the household and all regressions are weighted using survey weights. Sample consists of survey responders (N = 23741).

Table 7: Financial Strain (Administrative Data)

	Control Mean (1)	Reduced Form (2)	2SLS (3)	p-values (4)
Panel A: Overall				
Any bankruptcy	0.014 (0.119)	0.002 (0.001)	0.009 (0.005)	[0.106] {0.358}
Any lien	0.021 (0.144)	0.001 (0.001)	0.005 (0.006)	[0.406] {0.698}
Any judgment	0.064 (0.244)	0.001 (0.002)	0.005 (0.010)	[0.573] {0.698}
Any collection	0.500 (0.500)	-0.012 (0.004)	-0.048 (0.016)	[0.003] {0.013}
Any delinquency (credit accounts)	0.366 (0.482)	0.002 (0.004)	0.006 (0.017)	[0.704] {0.698}
Standardized treatment effect		0.002 (0.005)	0.009 (0.019)	[0.653]
Panel B: Medical Debt				
Any medical collection	0.281 (0.449)	-0.016 (0.004)	-0.064 (0.016)	[<0.0001] {<0.0001}
Amount owed in medical collections	1,999 (6733)	-99 (45)	-390 (177)	[0.028] {0.025}
Standardized treatment effect		-0.026 (0.006)	-0.100 (0.024)	[<0.0001]
Panel C: Non-Medical Debt				
Any non-medical collection	0.392 (0.488)	-0.005 (0.004)	-0.018 (0.016)	[0.264] {0.455}
Amount owed in non-medical collections	2,740 (9492)	-20 (63)	-79 (248)	[0.751] {0.752}
Standardized treatment effect		-0.006 (0.006)	-0.023 (0.023)	[0.325]

(Standard errors in parentheses)
[Per comparison p-values in square brackets]

{Family wise p-values in curly brackets}

Notes: All outcomes are measured since notification date through September 2009. Column 2 reports the coefficient and standard error on LOTTERY from estimating equation (1) by OLS. Column 2 reports the coefficient and standard error on INSURANCE from estimating equation (3) by IV. Column 4 reports the per-comparison p value and the family wise p-value across the different measures used to create the standardized treatment effect. Standardized treatment effect reports results based on equation (2). All regressions include household size fixed effects, lottery draw fixed effects, and the analogous outcome measure from the February 2008 credit report data. All standard errors are clustered on the household. Sample consists of all those matched to credit report data (N =49980).

Table 8: Financial Strain (Survey Data)

	Control Mean	Reduced Form	2SLS	p-values
	(1)	(2)	(3)	(4)
Any out of pocket medical expenses, last six months	0.555 (0.497)	-0.058 (0.008)	-0.200 (0.026)	[<0.0001] {<0.0001}
Owe money for medical expenses currently	0.597 (0.491)	-0.052 (0.008)	-0.180 (0.026)	[<0.0001] {<0.0001}
Borrowed money or skipped other bills to pay medical bills, last six	0.364 (0.481)	-0.045 (0.007)	-0.154 (0.025)	[<0.0001] {<0.0001}
Refused treatment bc of medical debt, last six months	0.081 (0.273)	-0.011 (0.004)	-0.036 (0.014)	[0.01] {0.01}
Standardized treatment effect		-0.089 (0.010)	-0.305 (0.035)	[<0.0001]

(Standard errors in parentheses)
[Per comparison p-values in square brackets]
{Family wise p-values in curly brackets}

Notes: Column 2 reports the coefficient and standard error on LOTTERY from estimating equation (1) by OLS. Column 3 reports the coefficient and standard error on INSURANCE from estimating equation (3) by IV. Column 4 reports the per-comparison p value and the family wise p-value across the four different measures of financial strain used to create the standardized treatment effect. Standardized treatment effect reports results based on equation (2). All regressions include household size fixed effects, survey wave fixed effects, and the interaction between the two. All standard errors are clustered on the household and all regressions are weighted using survey weights. Sample consists of survey responders (N =23741).

Table 9: Health

	Control Mean (1)	Reduced Form (2)	2SLS (3)	p-values (4)
Panel A: Administrative data				
Alive	0.992	0.00032	0.001	[0.638]
	(0.092)	(0.001)	(0.003)	
Panel B: Survey Data				
Self reported health good / very good / excellent (not fair or poor)	0.548	0.039	0.133	[<0.0001]
	(0.498)	(0.008)	(0.026)	{<0.0001}
Self reported health not poor (fair, good, very good, or excellent)	0.86	0.029	0.099	[<0.0001]
	(0.347)	(0.005)	(0.018)	{<0.0001}
Health about the same or gotten better over last six months	0.714	0.033	0.113	[<0.0001]
	(0.452)	(0.007)	(0.023)	{<0.0001}
# of days physical health good, past 30 days*	21.862	0.381	1.317	[0.019]
	(10.384)	(0.162)	(0.563)	{0.018}
# days poor physical or mental health did not impair usual activity, past 30 days*	20.329	0.459	1.585	[0.009]
	(10.939)	(0.175)	(0.606)	{0.015}
# of days mental health good, past 30 days*	18.738	0.603	2.082	[0.001]
	(11.445)	(0.184)	(0.64)	{0.003}
Did not screen positive for depression, last two weeks	0.671	0.0230	0.078	[0.001]
	(0.470)	(0.007)	(0.025)	{0.003}
Standardized treatment effect		0.0590	0.203	[<0.0001]
		(0.011)	(0.039)	

(Standard errors in parentheses)
[Per comparison p-values in square brackets]
{Family wise p-values in curly brackets}

Notes: Column 2 reports the coefficient and standard error on LOTTERY from estimating equation (1) by OLS. Column 3 reports the coefficient and standard error on INSURANCE from estimating equation (3) by IV. Column 4 reports the per-comparison p value and the family wise p-value across the different measures used to create the standardized treatment effect. Standardized treatment effect reports results based on equation (2). All regressions include household size fixed effects and standard errors are clustered on the household. The regressions in panel A include lottery draw fixed effects, and the dependent variable "alive" is measured from the notification date through September 2009 (N=74922). The regressions in panel B include survey wave fixed effects, and the interaction of survey wave fixed effects with household size fixed effects, and are weighted using the survey weights (N = 23741).

*These questions were worded to ask about # of days health "not good" or "impaired"; we switched the sign for consistency with the other measures. See Appendix Figure A4 for the exact survey wording

Table 10: Potential Mechanisms for Improved Health (Survey Data)

	Control Mean	Reduced Form	2SLS	p-values
	(1)	(2)	(3)	(4)
Panel A: Access to care				
Have usual place of clinic-based care	0.499	0.099	0.339	[<0.0001]
	(0.500)	(0.008)	(0.027)	{<0.0001}
Have personal doctor	0.490	0.081	0.280	[<0.0001]
	(0.500)	(0.008)	(0.026)	{<0.0001}
Got all needed medical care, last six months	0.684	0.069	0.239	[<0.0001]
	(0.465)	(0.006)	(0.022)	{<0.0001}
Got all needed drugs, last six months	0.765	0.056	0.195	[<0.0001]
	(0.424)	(0.006)	(0.019)	{<0.0001}
Didn't use ER for non-emergency, last six months	0.916	-0.001	-0.004	[0.804]
	(0.278)	(0.004)	(0.015)	{0.804}
Standardized treatment effect		0.128	0.440	[<0.0001]
		(0.008)	(0.029)	
Panel B: Quality of care				
Quality of care received last six months good/vg/exc (condl on any)	0.708	0.043	0.142	[<0.0001]
	(0.455)	(0.008)	(0.027)	
Panel C: Happiness				
Very happy or pretty happy (vs. not too happy)	0.594	0.056	0.191	[<0.0001]
	(0.491)	(0.007)	(0.026)	

(Standard errors in parentheses)
[Per comparison p-values in square brackets]
{Family wise p-values in curly brackets}

Notes: Column 2 reports the coefficient and standard error on LOTTERY from estimating equation (1) by OLS. Column 3 reports the coefficient and standard error on INSURANCE from estimating equation (3) by IV. Column 4 reports the per-comparison p value and the family wise p-value across the four different measures of financial strain used to create the standardized treatment effect. Standardized treatment effect reports results based on equation (2). All regressions include household size fixed effects, survey wave fixed effects, and the interaction between the two. All standard errors are clustered on the household and all regressions are weighted using survey weights. Sample consists of survey responders (N = 23741).

Table 11: Estimated Effects of Lottery at Different Times^

	Reduced Form Estimates From Survey:			P-value of Difference Between:		
	Initial (0m)	6m	Main (12m)	0m and 6m	6m and 12m	0m and 12m
	(1)	(2)	(3)	(4)	(5)	(6)
Utilization (extensive margin)	0.004	0.047	0.050	0.035	0.867	<.0001
	(0.008)	(0.020)	(0.011)			
	[0.656]	[0.02]	[<0.0001]			
Utilization (total)	0.000	0.027	0.040	0.188	0.556	0.001
	(0.009)	(0.020)	(0.011)			
	[0.978]	[0.187]	[0.0003]			
Financial strain	-0.035	-0.099	-0.089	0.002	0.613	<.0001
	(0.009)	(0.020)	(0.010)			
	[<0.0001]	[<0.0001]	[<0.0001]			
Health	0.042	0.097	0.061	0.014	0.121	0.112
	(0.01)	(0.023)	(0.011)			
	[<0.0001]	[<0.0001]	[<0.0001]			
Access	0.047	0.075	0.119	0.163	0.026	<.0001
	(0.008)	(0.019)	(0.009)			
	[<0.0001]	[<0.0001]	[<0.0001]			

^ This analysis was not pre-specified

(Standard errors in parentheses)
[Per comparison p-values in square brackets]

Notes: Table reports standardized treatment effects from estimating equation (1) in three different surveys. The first column shows the results for our so-called "initial survey" which occurred about 2.6 months after lottery notification; the second column shows results for our "six month survey" which occured about 8 months after lottery notification (six months after insurance coverage began); the third column shows results for our "main survey" which occured about 15 months after lottery notification; this is the survey that has been analyzed in prior tables. The surveys had similar response rates and questionnaires; the "six month" survey was conducted on an approximately 20 percent sample of the other surveys (see Section 3.4 and Appendix 1 for details). Columns 1, 2, and 3 report the coefficient, standard error and p-value on LOTTERY from estimating equation (1) by OLS. Columns 3, 4, and 5 reports the p-value of the difference between various estimates. All regressions include household size fixed effects, survey wave fixed effects, and the interaction of survey wave and household size fixed effects. Regressions based on the "six month" and "main" surveys (columns 2 and 3) are weighted using the survey weights for those surveys. Standard errors are clustered on the household. N = 26,423 for the initial survey, N=6,359 for the six month survey, and N=23,741 for the main survey. The individual components of the standardized treatment effects are the same as in the earlier tables except that "health" excludes the depression screen question and "access" excludes the personal doctor question (as these questions were not asked in the initial survey). The reference period for the underlying questions is usually either "currently" or "in the last six months" (see prior tables for details)

Table 12: Observational Estimates of Effect of Insurance in Study Population

	Random assignment	Any Medicaid vs. No Medicaid	Any medicaid vs. No Medicaid (controls only)	OHP Standard vs. No Medicaid (treatment only)
	(1)	(2)	(3)	(4)
Sample size	74,922	74,922	45,088	26,437
Percent insured	26	26	17	32
Panel A: Administrative Data				
Total hospital utilization	0.047	0.241	0.352	0.133
	(0.026)	(0.015)	(0.030)	(0.017)
	[0.073]	[<0.0001]	[<0.0001]	[<0.0001]
Financial strain	0.009	0.008	0.024	-0.005
	(0.019)	(0.005)	(0.008)	(0.008)
	[0.653]	[0.118]	[0.002]	[0.534]
Panel B: Survey Data				
Total health care utilization	0.137	0.291	0.460	0.193
	(0.038)	(0.014)	(0.030)	(0.018)
	[0.0003]	[<0.0001]	[<0.0001]	[<0.0001]
Financial strain	-0.305	-0.165	-0.091	-0.199
	(0.035)	(0.011)	(0.020)	(0.015)
	[<0.0001]	[<0.0001]	[<0.0001]	[<0.0001]
Self reported health	0.203	-0.074	-0.127	-0.086
	(0.039)	(0.013)	(0.024)	(0.017)
	[<0.0001]	[<0.0001]	[<0.0001]	[<0.0001]

(Standard errors in parentheses)
[Per comparison p-values in square brackets]

Notes: All estimates are standardized treatment effects (see equation 2). Column 1 reports standardized treatment effects calculated based on IV estimation of equation (3). All other columns are based on OLS estimation of equation (1), but with the variable LOTTERY substituted with an "Any Medicaid" indicator in columns 2 and 3, and for an "OHP Standard vs no Medicaid" indicator in Column 4. Column 2 compares all those with any Medicaid coverage during our study period to those without Medicaid (regardless of lottery status); this represents the "as treated" analysis sometimes done in clinical trials. To avoid having much of the variation in insurance coming from the lottery, the third column perform the same analysis within the control group; here, most of the insurance coverage is OHP Plus which covers a somewhat different population than OHP Standard. The fourth column therefore performs the analysis within the treatment group, comparing those on OHP Standard to those with no Medicaid (and dropping the small percentage of treatment individuals on Plus). Regressions using the administrative data (panel A), include household size fixed effects, lottery draw fixed effects, and a pre-period measure of the dependent variable. Regressions using the survey data (pane B) include household size fixed effects survey wave fixed effects, and the interaction of the two, and are weighted using the survey weights. All standard errors are clustered on the household. For each standardized treatment effect we report the estimate, standard error, and per comparison p-value. The components of the standardized treatment effects are identical to those given in the relevant tables. Specifically, in Panel A hospital utilization is from Table 4b (all admissions), and financial strain is from Table 8, Panel A. In Panel B, total health care utilization is based on Table 6 (total utilization), financial strain is from Table 9, and self reported health is from Table 10, Panel B. The top two rows report the sample size and percent insured for the full sample universe.

Table 13: Observational Estimates of Effect of Insurance in BRFSS and NHIS

	IV Estimates (Study Population)	OLS Estimates	
		Any Insurnace vs. Not Insurance	Any Insurance vs. No Insurance (covariate-adjusted)
	(1)	(2)	(3)
Panel A: Health Care Use (NHIS)			
Number of outpatient visit last six months	1.850	1.593	1.453
	(0.235)	(0.044)	(0.06)
	[<0.0001]	[<0.0001]	[<0.0001]
Number of ER visits last six months	0.068	0.124	0.089
	(0.117)	(0.013)	(0.018)
	[<0.0001]	[<0.0001]	[<0.0001]
Number of inpatient hospital admissions last six months	0.043	0.125	0.051
	(0.042)	(0.042)	(0.073)
	[0.311]	[0.003]	[0.482]
Panel B: Health (BRFSS)			
Self reported health good/ very good / excellent	0.133	-0.020	0.004
	(0.026)	(0.006)	(0.006)
	[<0.0001]	[0.001]	[0.45]
# days poor physical or mental health did not impair usual activity, past 30 days*	1.317	-2.406	-0.931
	(0.563)	(0.219)	(0.208)
	[0.019]	[<0.0001]	[<0.0001]
# of days physical health good, past 30 days*	1.585	-2.265	-0.936
	(0.606)	(0.189)	(0.175)
	[0.009]	[<0.0001]	[<0.0001]
# of days mental health good, past 30 days*	2.082	-0.606	0.252
	(0.64)	(0.19)	(0.184)
	[0.001]	[0.001]	[0.172]
Did not screen positive for depression, last two weeks	0.078	-0.046	-0.002
	(0.025)	(0.031)	(0.031)
	[0.001]	[0.14]	[0.95]

(Standard errors in parentheses)

[Per comparison p-values in square brackets]

Notes: Table explores comparability of the randomized results to observational estimates. Column (1) reports the IV estimates for our study population reported in previous tables for specific outcomes that we observe in national, observational data. Columns 2 and 3 report the observational estimates using 2004-2009 national data on adults aged 19-64 below 100 percent of the federal poverty line from the NHIS (N=15,528) and the BRFSS (N=144,829) as indicated. The fraction reporting insurance coverage is 65 percent in the NHIS and 56 percent in the BRFSS In column 2 we report results from a bivariate OLS regression of the outcome on an insurance dummy. In column 3 we add demographic controls for age, race, gender, education, marital status, number of children in the household, employment status, and bins of annual income. In both columns 2 and 3 we report robust standard errors. The outcomes analyzed are based on virtually identical questions in our survey (column 1) and the national data (columns 2 and 3) except that the utilization questions in the NHIS are asked with a 12 month rather than a 6 month look-back period; we therefore divided the responses by two to try to make them comparable with our survey outcomes. In addition, the hospital utilization question in the NHIS does not explicitly exclude child birth while ours does.

* These questions were worded to ask about days "not good" or "impaired"; we switched the sign for consistency w the other measures. See Appendix Figure A4 for the exact survey wording