THE
QUARTERLY JOURNAL
OF ECONOMICS

Vol. 127 August 2012 Issue 3

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
OREGON HEALTH STUDY GROUP

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 an 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. *JEL* Codes: H51, H75, I1.

*We are grateful to Josh Angrist, Robert Avery, David Autor, Ethan Cohen-Cole, Carlos Dobkin, Esther Duflo, Jack Fowler, Guido Imbens, Larry Katz, Jeff Kling, John McConnell, Jon Levin, Richard Levin, Ben Olken, Alan Zaslavsky, three anonymous referees, and numerous seminar participants for helpful comments and advice; to Brandi Coates, Michael Gelman, John Graves, Ahmed Jaber, Andrew Lai, Conrad Miller, Iuliana Pascu, Adam Sacarny, Nivedhitha Subramanian, Zirui Song, James Wang, and Annetta Zhou for expert research assistance; and to numerous Oregon state employees for help acquiring the necessary data and for answering our many questions about the

The Quarterly Journal of Economics (2012), 1057–1106. doi:10.1093/qje/qjs020.
Advance Access publication on May 3, 2012.

I. 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 an opportunity to study the effects of access to public insurance using the framework of a randomized controlled design.

Although the effects of health insurance on health and health care use may seem intuitive, and there have been hundreds of studies on the topic, research in this area has often been hampered by the difficulty of controlling for unobserved differences between the insured and uninsured (Levy and Meltzer 2008). Random assignment of health insurance to some but not others

administration of state programs. We gratefully acknowledge funding from the Assistant Secretary for Planning and Evaluation in the Department of Health and Human Services, the California HealthCare Foundation, the John D. and Catherine T. MacArthur Foundation, the National Institute on Aging (P30AG012810, RC2AGO36631, and R01AG0345151), the Robert Wood Johnson Foundation, the Sloan Foundation, the Smith Richardson Foundation, and the U.S. Social Security Administration (through grant 5 RRC 08098400-03-00 to the National Bureau of Economic Research as part of the SSA Retirement Research Consortium). We also gratefully acknowledge Centers for Medicare and Medicaid Services' matching funds for this evaluation. The findings and conclusions expressed are solely those of the author(s) and do not represent the views of SSA, the National Institute on Aging, the National Institutes of Health, any agency of the federal government, any of our funders, or the NBER. In addition to the individuals named as coauthors for this article, the Oregon Health Study Group includes Matt Carlson (Portland State University), Tina Edlund (Deputy Directory, Oregon Health Authority), Charles Gallia (Oregon DHS), Eric Schneider (RAND), and Jeanene Smith (Office for Oregon Health Policy and Research). Author disclosures: Finkelstein served on the CBO's Panel of Health Advisers through 2011. Wright is employed by Providence Health & Services, a nonprofit integrated health care delivery system. Gruber was a paid technical consultant to the Obama administration during the development of the Affordable Care Act and a paid consultant to the state of Oregon for modeling health insurance expansion options, and serves on the CBO's Panel of Health Advisers. Newhouse is a director of and holds equity in Aetna, which sells Medicaid policies, and serves on the CBO's Panel of Health Advisers. Allen is employed by Providence Health and Services, a nonprofit integrated health care delivery system. She formerly served as director of the Medicaid Advisory Committee and as staff to the Oregon Health Fund Board at the Office for Oregon Health Policy and Research. Baicker is a MedPAC commissioner, serves on the CBO's Panel of Health Advisers, is a director of Eli Lilly, has received honoraria from several physician groups for speaking engagements, and previously served on the Bush administration's Council of Economic Advisers.

would avoid such confounding, but such opportunities are rarely available.[1]

In this article 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 on self-reported health, financial strain, and overall well-being. By lowering the price of health care, health insurance is expected to increase health care utilization. Ultimately additional health care utilization may translate into improved health, although a one-year window might be too short a time to observe health improvements. Much less attention has been given in the literature to other potential benefits of health insurance. Because risk-spreading is arguably the primary purpose of health insurance (e.g., Zeckhauser 1970), we try to examine the impact of health insurance on consumption smoothing, which we proxy for with measures of financial strain. We also examine the impact of health insurance on overall well-being, specifically self-reported happiness; this may capture, among other things, any benefits of health insurance from reductions in stress or stigma.

The impact of Medicaid among a low-income population may be lower than that of private insurance or insurance among higher income individuals. The impact of Medicaid may be attenuated (or potentially nonexistent) if public health clinics and

1. We know of only two other randomized health insurance experiments in the United States. The RAND Health Insurance Experiment from the 1970s 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 (Manning et al. 1987; Newhouse and the Insurance Experiment Group 1993). The more recent Accelerated Benefits Demonstration project was designed to investigate the impact of health insurance for uninsured disabled adults receiving Social Security Disability Insurance during the two-year waiting period for Medicare (Michalopoulos et al. 2011).

uncompensated care allow low-income 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 claimed—Medicaid itself is not particularly "good" insurance in terms of being able to access health care providers (e.g., Medicaid Access Study Group 1994; GAO 2011; Rosenbaum 2011).

Our analysis draws on administrative data from hospital discharge, credit report, and mortality records, as well as on responses to a large mail survey 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% effective response rate, are subject to potential nonresponse bias. Our available evidence on this issue is limited but reasonably reassuring.

Prior to looking at the data on outcomes for the treatment group, virtually all of the analysis presented here was prespecified and publicly archived in a detailed analysis plan.[2] Although prespecification of hypotheses is the norm for randomized controlled medical trials, is it rare in evaluation of social policy experiments.[3] Our prespecification was designed to minimize issues of data and specification mining and to provide a record of the full set of planned analyses.

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

2. Our analysis plan was archived on December 3, 2010, at http://www.nber.org/sap/20101203/. Some of those analyses yielded little of interest, and therefore we describe them briefly, presenting the full results only in appendixes. 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.

3. In economics, within the last few years, prespecification of hypotheses has started to become more common in analyses of randomized experiments in developing countries (e.g., Alatas et al. 2010; Olken, Onishi, and Wong 2010; Schaner 2010; Casey, Glennerster, and Miguel 2011).

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%) increase in the probability of having a hospital admission, an 8.8 percentage point (15%) increase in the probability of taking any prescription drugs, and a 21 percentage point (35%) 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 confidence intervals do not allow us to rule out substantial effects in either direction. In addition, insurance is associated with 0.3 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%) decline in the probability of having an unpaid medical bill sent to a collections agency and a 20 percentage point (35%) decline in having any out-of-pocket medical expenditures. Because much medical debt is never paid, the financial incidence of expanded coverage thus appears to be not only on the newly insured but also on their medical providers (or whomever they pass the costs on to).

Finally, we find that insurance is associated with improvements across the board in measures of self-reported physical and mental health, averaging 0.2 standard deviation improvement. Two pieces of evidence suggest that the improvements in self-reported health that we detect may at least partly reflect a general sense of improved well-being. First, evidence from a separate survey we conducted very shortly after random assignment shows no impact of lottery selection on health care utilization but improvements in self-reported health that are about two-thirds the magnitude of our main survey results more than a year later. Second, we find that one year later, Medicaid is associated with about a 32% increase in self-reported overall happiness, albeit reported in the context of a survey primarily about health. Whether there are also improvements in objective, physical health is more difficult to determine with the data we now have available. More data on physical health, including biometric measures such as blood pressure and blood sugar, will be available from the in-person interviews and health exams

that we conducted about six months after the time frame in this article.

Our estimates of the impact of public health insurance apply to able-bodied uninsured adults below 100% of poverty who express interest in insurance coverage, 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% of poverty absent specific categorical requirements (Kaiser Family Foundation 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% of the federal poverty level, with no financial penalties for many individuals in this income range who do not take up coverage (Kaiser Family Foundation 2010a, 2010b; U.S. GPO 2010).

The rest of the article is structured as follows. Section II provides background on the Oregon Medicaid program and the lottery design. Section III describes the primary data sources, and Section IV presents our empirical framework. Section V presents our main results. Section VI discusses interpretation and extrapolation of our estimates. The Online Appendix provides additional details.

II. 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. OHP Plus serves the categorically eligible Medicaid population, which includes (up to specific income thresholds) children 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 public insurance 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 (Office for Oregon Health Policy and Research 2009).

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 nonemergency 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% 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 reopened 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 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. From January 28 to February 29, 2008, anyone could be added to the lottery 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 Online 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 screen for program eligibility at sign-up for the lottery. A total of 89,824 individuals were placed on the list during the five-week window it was open.

The state conducted eight lottery drawings from the list with roughly equal numbers selected from each drawing; the drawings were fairly evenly spaced 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 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.[4] About 30% of selected individuals successfully enrolled. There were two main sources of slippage: only about 60% 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 recertified their eligibility status every six months.

III. Data

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

III.A. Administrative Data on Outcomes: Hospital Discharges, Credit Reports, and Mortality

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, Dobkin, and Maestas 2008, 2009; Anderson, Dobkin, and Gross 2010). Although inpatient admissions are relatively rare (the annual admission rate for our controls is only about 5%), they are expensive, accounting for about one-quarter

4. The state reviewed applications, first examining eligibility for OHP Plus and then, if not eligible for Plus, examining eligibility for OHP Standard. Those who did not apply during this window could not apply later (so unlike those categorically eligible for Medicaid/OHP Plus, did not have "conditional coverage" if unenrolled).

of total medical expenditures for 18–64-year-olds.[5] We observe, on average, about five months of prerandomization 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 gleaned from public records, collections agencies, and "trade lines" such as credit card balances (but do not capture informal borrowing such as through relatives or pawnbrokers). 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.[6] TransUnion used the full name, full address, and date of birth reported at sign-up to match 68.5% of lottery participants to their prerandomization credit report in February 2008. The credit bureau was able to track over 97% 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 postrandomization "attrition rate" of 3%. We also observe prerandomization 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.

III.B. Survey Data on Outcomes

We supplement the outcome measures available in the administrative data with a mail survey that was sent out in seven waves over July and August 2009 to virtually all individuals selected by the lottery and a roughly equal number of unselected individuals.[7] The complete survey instrument is shown in Online Appendix Figure A4. The basic protocol involved three mail attempts. In addition, we designed a more intensive protocol, which we conducted on approximately 30% of nonrespondents. It included additional tracking efforts, mailings, and phone

5. Author calculations based on publicly available tables from the 2008 Medical Expenditure Panel Survey.

6. 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.

7. The seven survey waves sent out do not map directly to lottery drawings (of which there were eight). See Online Appendix Table A2 (on eight lottery drawings) and Table A9 (on the seven survey waves) for more detail.

contacts. The response rate to the basic protocol was 36%; about 22% of those who did not respond to the basic protocol and who received the intensive protocol responded. We calculate an effective response rate of 50%, with individuals who responded to the intensive follow-up weighted by the inverse probability of being included in the intensive follow-up subsample.

In Section V.C, 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 midway between the initial and main survey. The six-month survey was conducted on a 20% 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%, respectively.

III.C. Other Data

We obtained prerandomization demographic information that the participants provided at the time of lottery sign-up. Online Appendix Figure A2 shows the sign-up form. We use these data primarily to construct nine "lottery list variables" that we use to examine treatment and control balance on prerandomization demographics.[8] 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 these data as our primary measure of the first-stage outcome (i.e., insurance coverage). Finally we obtained state administrative records on the Food Stamp and TANF benefit history of lottery list participants from prior to the lottery through September 2009.

8. 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 metropolitan statistical area.

III.D. 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.[9] This observation period represents, on average, 16 months (std. dev. = 2 months) after individuals are notified of their selection and, on average 14 months (std. dev. = 3 months) after insurance coverage is approved for those selected by the lottery who successfully enrolled in OHP Standard. 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.[10]

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 (std. dev. = 2.7) months or 13.1 months after insurance approval (std. dev. = 2.9 months).

III.E. Sample Definition and Descriptive Statistics

Of the 89,824 individuals who were on the lottery list, we used prerandomization 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 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

9. We randomly assigned lottery draws to the control individuals as discussed in more detail in Section IV.

10. 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 affect the financial burden associated with health care utilization during that time period.

TABLE I

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 income	$39,265
Panel B: Survey responders only			
Lottery list variables			
Sex		*Language*	
% Female	0.591	% English preferred	0.917
Age		*ZIP code–level variables*	
% 50–64	0.316	% MSA	0.751
% 20–50	0.684	ZIP code median household income	$39,225
12-month mail survey variables			
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 bronchitis	0.129
% Spanish/Hispanic/Latino	0.123	Depression (screen positive)	0.557

(continued)

TABLE I

(CONTINUED)

Variable	Control mean	Variable	Control mean
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-year 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 hours per week	0.090	Any insurance?	0.325
% work 20–29 hours 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) $	13,035	# of months of last six with insurance	1.738

Notes. All statistics are reported for control individuals only. Panel A reports the control means for prerandomization demographics taken from the lottery list (from January and February 2008) for the whole sample ($N = 45{,}088$ 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 lottery list prerandomization demographics and survey questions for survey responders ($N = 11{,}933$ 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 midpoint of their bin (see Online Appendix Figure A4 for details). For the insurance questions, we code as "yes" if the respondent checked that insurance type box; because 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 checkboxes blank. We construct income relative to the federal poverty line based on self-reported income and self-reported (total) number of household members. See Online Appendix 3 for more details.

population down to 74,922 individuals (representing 66,385 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. Online 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 I 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% female; about one-quarter are 50 to 64 years old at the end of our study period (the average age is 41); 92% have a preferred language of English; and about three-quarters live in a metropolitan statistical area (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% black and 12% 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% report having ever been diagnosed with diabetes, 28% with asthma, 40% with high blood pressure; 56% screen positive for depression.[11]

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% report incomes below the eligibility cut-off of 100% of the federal poverty level; this is consistent with our finding from analyzing application data that income eligibility requirements disqualified a nontrivial share of selected individuals (Allen et al. 2010). Finally, about 30% of the controls report having insurance (which, if they had it at the time of the

11. By contrast among a general adult population, 7% report ever being diagnosed with diabetes, 14% with asthma, 24% with high blood pressure, and 28% 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.)

lottery, would have also made them ineligible for OHP Standard); 13% report having private insurance.

IV. Empirical Framework

IV.A. Intent to Treat Effects of Lottery

I. Intent-to-Treat Equation. We estimate the intent-to-treat (ITT) effect of winning the lottery (i.e., the difference between treatment and controls) by estimating the following ordinary least squares (OLS) equation:

$$y_{ihj} = \beta_0 + \beta_1 LOTTERY_h + X_{ih}\beta_2 + V_{ih}\beta_3 + \varepsilon_{ihj}, \quad (1)$$

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 will discuss, we summarize the estimates within a domain by the standardized treatment effect across outcomes in that domain; we 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 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); it is interpreted as the impact of being able to apply for OHP Standard through the Oregon 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 of 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"); as already noted, although the state randomly sampled from individuals on the list, the entire household of any selected

individual was considered selected and eligible to apply for insurance. As a result, selected (treatment) individuals are disproportionately drawn from larger households.[12] 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) because the fraction of treatment individuals varies across the seven survey waves.

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, prerandomization measures of outcomes in the credit report data and hospital discharge data, and the lottery draw to which the individual is assigned.[13] 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 prerandomization measure for the outcome analyzed in the hospital and credit report data. Online Appendix Table A15 shows that our results are not sensitive to other choices regarding the V_{ih} covariates.

In all of our ITT estimates and in our subsequent instrumental variable estimates (see Section IV.C), we estimate linear models even though a number of our outcomes are binary. Because we are interested in the difference in conditional means for the treatments and controls, linear probability models pose no concerns in the absence of covariates or in fully saturated models (Angrist 2001; Angrist and Pischke 2009). Our models are not fully saturated, however, so it is possible that the choice of models could make a difference, especially for outcomes with very low or very high mean probability. In Online Appendix Table A16 we show that the analysis of these "tail" outcomes is

12. 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).

13. In the administrative data, we measure outcomes from the notification date. For treatment individuals, notification date varies by lottery draw (which spans 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.

not sensitive to estimating nonlinear models instead (this analysis was not prespecified).

In all of our analyses we cluster the standard errors on the household identifier because the treatment is at the household level. Analyses of survey data are weighted to account for the sampling design of the survey as described already (and in Online Appendix 1.7); we show in Online Appendix Table A15 that the results are similar if we do not use survey weights or if we limit the analysis to the individuals who responded to the basic protocol (this analysis was not prespecified).

Handling many outcomes: standardized treatment effects and multiple inference. We summarize multiple findings across related outcomes within a domain J by the average standardized treatment effect:[14]

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

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 ITT estimates in equation (1), the ξ_{1j}'s correspond to the β_{1j}'s). To account for covariance in the estimates of $\frac{\xi_{1j}}{\sigma_j}$, we estimate pooled OLS for all outcomes $j \in J$.[15] 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 and 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.[16]

14. The approach in this section draws heavily on Kling and Liebman (2004) and Kling, Liebman, and Katz (2007).

15. Specifically, we stacked the data for the individual outcomes within a domain and estimated a single regression equation that allowed the coefficients on each covariate to vary flexibly across the outcomes and for correlation in the error terms across outcomes.

16. 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

The per comparison p-value may be appropriate if one is interested in the particular hypothesis of the impact of insurance on a specific outcome (such as depression). The family-wise 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). In practice, it is rare for one of our results to be statistically significant (e.g., a p-value of less than .05) in the per comparison test and not significant after adjusting for multiple inference.[17]

IV.B. 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 subsamples 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 (Online 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 Online Appendix 2.2 and Table A12).

Differences in attrition (match rates or response rates) or in the prerandomization characteristics of the treatment and control analysis samples would raise concerns about the second key assumption for causal inference. Table II therefore investigates treatment-control balance for three different samples in columns

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); see Kling and Liebman (2004) or Anderson (2008) for more detailed discussions as well as applications.

17. 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 article 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 on request.

TABLE II

TREATMENT: CONTROL BALANCE

	Control mean (std. dev.) for full sample	Difference between treatment and control		
		Full sample	Credit report subsample	Survey respondents subsample
	(1)	(2)	(3)	(4)
Panel A: Match/response rates				
Matched in September 2009 credit data	0.663		−0.0043	
	(0.473)		(0.0037)	
			[0.247]	
Responded to survey	0.506			−0.016
	(0.500)			(0.0066)
				[0.014]
Response time (in days)	53.0			1.638
	(57.8)			(1.088)
				[0.132]
Panel B: Prerandomization 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		74,922	49,980	23,741

Notes. Standard errors in parentheses; per comparison *p*-values in square brackets. 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 column. Columns (3) and (4) report estimated differences between treatments and controls for the dependent variable shown in the left 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 prerandomization 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 the preferred language for receiving materials; whether the individual signed him- or herself up for the lottery or were signed up by a household member; whether the individual gave his or her address as a PO box; whether he or she 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 given is within a census-defined MSA; and whether he or she provided a phone number on sign up. "Prerandomization 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 prerandomization outcomes, the exact regression specifications, and the results for each variable analyzed in Panel B are presented in Online Appendix 3, Table A13.

(2) through (4), respectively: the sample universe (which is the sample analyzed in the hospital discharge 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% nonresponse rate (compared to an effective match rate of over 97% in the credit report data).

Panel A shows the balance of match rates. In the credit report data, the difference in (unconditional) match rates between treatment and control groups is a statistically insignificant 0.4 percentage points. In the survey respondent sample, there is a statistically significant 1.6 percentage point (std. err. = 0.7) lower response rate for treated individuals, off of a 51% base.[18] 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%), 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% versus 63%; Newhouse and the Insurance Experiment Group 1993).

Among the matched or responding subsample, we investigated the treatment-control balance for various prerandomization characteristics. In total we examined treatment-control balance for 31 different variables in up to three different samples (54 total comparisons). In Panel B we report the summary F-statistics and p-values on the treatment-control balance of groups of variables. 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 (prerandomization outcomes) shows the balance of prerandomization outcomes that match ones that we subsequently analyze postrandomization. The selection of these variables was prespecified and is described in Online Appendix 2.2 (Appendix Table A13). Finally, the third row (both of the above) examines balance on the lottery list and prerandomization outcomes combined. In each of the three samples we are unable to reject the null of treatment-control balance on the lottery list variables, the prerandomization measures, or the combined set of variables. All p-values are above

18. Conditional on response, response time between treatment and controls is indistinguishable.

.23. For the survey respondent subsample—where a priori we were most concerned about potential imbalance—the p-values are all at least .28.[19]

The individual results for treatment-control balance for each variable that contributes to the reported F-tests are shown in Online Appendix Table A13. In only two of the comparisons (one in the full sample and one in the credit report subsample) did we obtain p-values below .05 ($p = .039$ and $p = .014$); this is consistent with what we would expect by chance with 54 comparisons.

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 subsamples do not threaten causal inference but may be important for comparing estimates across data sets as well as for extrapolating to other contexts. Survey responders are on average almost two years older and 3 percentage points more likely to be female than the full sample; they look similar to the full sample in terms of their language preference and urbanicity. There are no material observable differences between the credit report subsample and the full sample (see Online Appendix Table A1).

IV.C. Local Average Treatment Effect of Medicaid (LATE)

The ITT 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*. We model this as follows:

$$(3) \qquad y_{ihj} = \pi_0 + \pi_1 INSURANCE_{ih} + X_{ih}\pi_2 + V_{ih}\pi_3 + \upsilon_{ihj},$$

where *INSURANCE* is a measure of insurance coverage and all other variables are as defined in equation (1). We estimate

19. We use the fairly conservative procedure of Lee (2009) to bound the potential bias arising from differential response rates between treatments and controls. Our ability to reject the null of no effect of health insurance on health care use or financial strain is generally robust to this bounding exercise (although naturally the magnitudes are attenuated at the lower bound), but our ability to reject the null of no impact of health insurance on self-reported health is generally not robust to this bounding exercise. These results are presented in Online Appendix Table A14 and discussed in Online Appendix 2.2. (This analysis was not prespecified.)

equation (3) by two-stage least squares (2SLS), using the following first-stage equation:

$$(4)\quad INSURANCE_{ih} + \delta_0 + \delta_1 LOTTERY_h + X_{ih}\delta_2 + V_{ih}\delta_3 + \mu_{ihj},$$

in which the excluded instrument is the variable *LOTTERY* with the first-stage coefficient of δ_1.[20]

We interpret the coefficient on insurance from instrumental variable estimation of equation (3) as a local average treatment effect (LATE) of insurance (Imbens and Angrist 1994). In other words, our estimate of π_1 identifies the causal impact of insurance among the subset of individuals who obtain insurance on winning the lottery and who would not obtain insurance without winning the lottery (i.e., the compliers). Because the model is just identified, the LATE estimate of π_1 is given by the ratio of the ITT coefficient on LOTTERY (equation (1)) and the first-stage coefficient on LOTTERY (equation (4)), or $\frac{\beta_1}{\delta_1}$. Table III reports our first-stage estimates based on estimation of equation (4). The first row reports the results using the measure of insurance that is used in all of our LATE estimates: whether the individual was ever on Medicaid (which includes both OHP Standard and OHP Plus) during our study period, as measured in the state's Medicaid enrollment files. 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% take-up already discussed; in addition, a small percentage of the controls became eligible for OHP Plus over our study period.[21] Relative to our study population, compliers are somewhat older, more likely white, in worse health, and in lower socioeconomic status.[22]

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

21. As we discuss in more detail in Online Appendix 1.3, the way the state identified lottery participants in the Medicaid enrollment files may cause it to slightly underestimate enrollment among nonselected (control) individuals, and thus cause us to overestimate our first stage by, we estimate, 2 percentage points or less.

22. The relative likelihood of being a complier is given by the ratio of the first stage in the subgroup to the overall first stage (Angrist and Pishke 2009, p. 171). These results are reported in Online Appendix Table A26.

TABLE III

FIRST-STAGE ESTIMATES

		Full sample		Credit report subsample		Survey respondents	
		Control mean (1)	Estimated FS (2)	Control mean (3)	Estimated FS (4)	Control mean (5)	Estimated FS (6)
(1)	Ever on Medicaid	0.141	0.256	0.135	0.255	0.135	0.290
			(0.0035)		(0.0042)		(0.0066)
(2)	Ever on OHP Standard	0.027	0.264	0.028	0.264	0.026	0.302
			(0.0029)		(0.0036)		(0.0055)
(3)	# of months on Medicaid	1.408	3.355	1.352	3.366	1.509	3.943
			(0.045)		(0.055)		(0.090)
(4)	On Medicaid, end of study period	0.106	0.148	0.101	0.151	0.105	0.189
			(0.0031)		(0.0038)		(0.0061)
(5)	Currently have any insurance (self-report)					0.325	0.179
							(0.0077)
(6)	Currenty have private insurance (self-report)					0.128	−0.0076
							(0.0053)
(7)	Currently on Medicaid (self-report)					0.117	0.197
							(0.0063)
(8)	Currently on Medicaid					0.105	0.191
							(0.0060)
(9)	Ever on TANF	0.031	0.0011	0.028	0.0021	0.023	0.0019
			(0.0013)		(0.0016)		(0.0025)
(10)	TANF benefits ($)	124	−1.659	111	1.543	100	−4.991
			(5.813)		(6.571)		(10.884)
(11)	Ever on food stamps	0.606	0.017	0.594	0.018	0.622	0.023
			(0.0029)		(0.0035)		(0.0054)
(12)	Food stamp benefits ($)	1776	61.3	1787	60.0	2202	122.4
			(15.0)		(18.8)		(33.4)
	N		74,922		49,980		23,741

Notes. Standard errors in parentheses. 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 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" (rows (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 rows (1) and (2), "ever" refers to enrollment ever during the study period, as defined in the text. In row (3), "# of months" refers to number of months enrolled during the study period. In row (8), insurance is measured as being on Medicaid according to the state Medicaid enrollment data on the day the survey was returned. All outcomes are measured for the individual except that in row (10) (12) where TANF (food stamp) benefits measure total household benefits received over the study period. For purposes of measuring the first stage, the study period is defined as ending on September 30, 2009. For the second-stage outcomes that we analyze in the administrative data, the study period in the first stage begins with the notification date (which varies by lottery draw). For the second stage outcomes that we analyze in the survey data, the study period in the first stage begins on the first notification date (March 10, 2008).

The subsequent rows of Table III 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 patterns. Not surprisingly, the results in row (2) show that the 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.[23] Row (3) shows that on average the lottery is associated with an increase of 3.4 months of Medicaid coverage. Over time the difference in insurance coverage between treatment and controls attenuates as enrolled individuals had to recertify their eligibility every six months and control individuals could find other insurance, particularly OHP Plus (see Online 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)).

In our subsequent results, we use "ever on Medicaid" during our study period for all of the reported LATE estimates of the impact of Medicaid. However, these alternative definitions of insurance are also reasonable ways to scale the ITT estimates. The "number of months on Medicaid" may be more appropriate than "ever on Medicaid" where the effect of insurance on the outcome is linear in the number of months insured. The measure "on Medicaid at the end of our study" can be thought of as providing a lower bound on the first stage (and hence an upper bound on our estimate of the LATE of Medicaid), whereas our baseline measure "ever on Medicaid during our study period" provides an upper

23. When insurance is defined as "ever on OHP Standard" we can probably be comfortable interpreting the IV estimates of equation (3) as the treatment-on-treated (ToT) rather than a LATE. In practice, there are two small violations of this interpretation. First, if there were no way to get OHP Standard without winning the lottery, there would be no always-takers in the terminology of Angrist, Imbens, and Rubin (1996), but about 2% of our controls got onto OHP standard through some limited mechanisms—for example, pregnant women who are on OHP Plus can sometimes stay on OHP Standard after giving birth. Second, it is possible that some compliers were put on OHP Plus rather than Standard, because case workers are instructed to first check applicant eligibility for Plus; in practice, this number is likely to be small since the estimated first stage is very similar for "ever on Medicaid" (which includes Plus and Standard) and "ever on OHP Standard" (see rows (1) and (2) of Table III).

bound on the first stage (and hence a lower bound on our estimate of the LATE of Medicaid).

The survey data 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)). This suggests the lottery did not crowd out other forms of insurance. Consistent with this, row (6) indicates a statistically insignificant decline in self-reports of private insurance coverage of only 0.8 percentage points (std. err. = 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 LATE interpretation requires the additional identifying assumption that there is no effect on average on the outcomes studied of winning the lottery that does not operate via the lottery's impact on insurance coverage. We believe this is a reasonable approximation, but it may not be strictly true. There are (at least) two possible types of violations. First, the event of winning (or losing) the lottery may have direct effects on the outcomes we study, although it seems unlikely to us that any such effects both exist and persist a year after the lottery. Second, individuals who apply for public health insurance 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 interested applicants in applying for TANF (cash welfare) and the Supplemental Nutrition Assistance Program (more commonly known as food stamps). 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 LATE interpretation of the impact of insurance per se, as opposed to the effect of expanded access to Medicaid (the ITT analysis).

Table III indicates that selection by the lottery is not associated with any substantive or statistically significant change in TANF receipt or benefits. However, lottery selection is associated with a statistically significant but substantively trivial increase in the probability of food stamp receipt (1.7 percentage points) and in total food stamp benefits (about $60 over a 16-month period, or less than 0.5% of annual income).[24] 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%. The impact on health seems likely to be negligible as well. Thus, we are comfortable interpreting our IV estimates as the effect of insurance coverage, or more specifically, the Oregon Medicaid program, on outcomes.

V. Results

V.A. Health Care Utilization

1. Administrative Data. Table IV presents our primary estimates using the hospital discharge data; all analyses exclude admissions for childbirth.[25] The most common admission is for mental disorders (approximately 10% of admissions); skin infections, complications from diabetes, and alcohol-related disorders are also common conditions (see Online Appendix Table A4).

Table IV, Panel A reports our estimates on admissions probabilities. The LATE results suggest that insurance is associated with an increase in the probability of any hospital admission of 2.1 percentage points (std. err. = 0.7), or about 30%. The increase in hospital admissions appears to be disproportionately concentrated in the approximately 35% of admissions that do not

24. This is likely an upper estimate because the cash equivalent of food stamps may be less than one (Hoynes and Schanzenbach 2009).

25. Regardless of lottery selection, many women in our sample would become categorically eligible for OHP Plus for childbirth. However, Oregon does not have presumptive Medicaid eligibility for pregnant women (Kaiser Family Foundation 2011), so it is possible that lottery selection could have an impact on pregnancy-related hospital use. In Online Appendix Table A21, we show that our results regarding hospital utilization are quite similar if we include admissions for childbirth in the analysis or if we exclude women of childbearing age from the sample (this analysis was not prespecified).

TABLE IV
HOSPITAL UTILIZATION

	Control mean (1)	ITT (2)	LATE (3)	*p*-values (4)
Panel A: Extensive margin				
All hospital admissions	0.067	0.0054	0.021	[0.004]
	(0.250)	(0.0019)	(0.0074)	
Admissions through ER	0.048	0.0018	0.0070	[0.265]
	(0.214)	(0.0016)	(0.0062)	
Admissions not through ER	0.029	0.0041	0.016	[0.002]
	(0.167)	(0.0013)	(0.0051)	
Panel B: 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]
	(19,942)	(146)	(569)	{0.106}
Procedures	0.155	0.018	0.070	[0.031]
	(1.08)	(0.0083)	(0.032)	{0.059}
Standardized treatment effect		0.012	0.047	[0.073]
		(0.0067)	(0.026)	
Panel C: 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]
	(12,749)	(96)	(376)	{0.171}
Procedures	0.081	0.0080	0.031	[0.135]
	(0.694)	(0.0054)	(0.021)	{0.187}
Standardized treatment effect		0.011	0.044	[0.100]
		(0.0069)	(0.027)	
Panel D: Admissions not through ER				
Days	0.199	0.0033	0.013	[0.841]
	(2.38)	(0.017)	(0.065)	{0.842}
List charges	1,110	98	384	[0.281]
	(12,422)	(91)	(356)	{0.383}
Procedures	0.075	0.010	0.038	[0.080]
	(0.708)	(0.0056)	(0.022)	{0.162}
Standardized treatment effect		0.0077	0.030	[0.254]
		(0.0068)	(0.026)	

Notes. Standard errors in parentheses; per comparison *p*-values in square brackets; family-wise *p*-values in curly brackets. 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; for the IV estimates in column (3), the endogenous variable *INSURANCE* is defined as "ever on Medicaid" during our study period and the first stage is given in the first row of Table III. 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 ($N = 74{,}922$).

originate in the emergency room, suggesting that these admissions may be more price sensitive.

Table IV, Panels B–D examine three measures of total utilization commonly used in the literature (see Card, Dobkin, and Maestas 2009): number of hospital days, total list charges, and number of procedures performed.[26] 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. The LATE estimates in Panel B show substantial increases in each of the three measures of utilization—with implied proportional increases of about 20% for hospital days, 40% for list charges, and 45% for the number of procedures, although only the result for procedures is statistically significant at the 10% level. The standardized treatment effect across all three measures indicates that insurance is associated with a 0.047 standard deviation increase in overall utilization (std. err. = 0.026); the relatively small standardized effect partly reflects the large variance of the underlying variables.

We undertook several additional investigations, which are presented in detail in Online Appendix 3.2. 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 (see Online Appendix Table A18). We also explored the impact of health insurance on a measure of 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 sufficiently wide that we cannot rule out quantitatively large effects (see Online Appendix Table A19). Finally, we examined whether

26. 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, Dobkin, and Maestas 2009).

TABLE V

HEALTH CARE UTILIZATION (SURVEY DATA)

	Extensive margin (any)				Total utilization (number)			
	Control mean (1)	ITT (2)	LATE (3)	*p*-values (4)	Control mean (5)	ITT (6)	LATE (7)	*p*-values (8)
Prescription drugs currently	0.637	0.025	0.088	[0.002]	2.318	0.100	0.347	[0.049]
	(0.481)	(0.0083)	(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)	(0.0074)	(0.025)	{<0.0001}	(3.087)	(0.054)	(0.182)	{<0.0001}
ER visits last six months	0.261	0.0065	0.022	[0.335]	0.47	0.0074	0.026	[0.645]
	(0.439)	(0.0067)	(0.023)	{0.547}	(1.037)	(0.016)	(0.056)	{0.643}
Inpatient hospital admissions last six months	0.072	0.0022	0.0077	[0.572]	0.097	0.0062	0.021	[0.311]
	(0.259)	(0.0040)	(0.014)	{0.570}	(0.4)	(0.0062)	(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)	

Notes. Standard errors in parentheses; per comparison *p*-values in square brackets; family-wise *p*-values in curly brackets. 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; for the IV estimates in column (3), the endogenous variable *INSURANCE* is defined as "ever on Medicaid" during our study period and the first stage is given in the first row of Table III. 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 ($N = 23{,}741$).

[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 (age 19–64) adults below 100% 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.

insurance was associated with a change in the proportion of patients going to public versus private hospitals and were unable to detect any substantive or statistically significant differences (see Online Appendix Table A20).

2. Survey Data. The survey data allow us to examine a broader range of utilization outcomes. Table V shows the results. Once again, we present results on both the extensive margin (left panel) and on total utilization (right panel). On both margins there are substantial and (mostly) statistically significant increases in prescription drugs and outpatient use. For example, the LATE estimates suggest that insurance is associated with a 0.35 (std. err. = 0.18) increase in the number of prescription drugs currently taken (corresponding to an approximately 15% increase) and a 1.08 (std. err. = 0.18) increase in the number of outpatient visits (corresponding to an over 55% 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).

There is no discernible impact of insurance on emergency room (ER) use on either margin. OHP Standard does not impose financial penalties for individuals who go to the ER for nonemergency reasons (Kaiser Family Foundation 2011), so it is possible that our inability to detect an impact of the insurance on ER use reflects the fact that people who are used to going to the ER for nonemergency purposes continue to do so. Our finding may also simply reflect a lack of power; for example, we cannot reject the hypothesis that Medicaid is associated with an almost 10% decline in ER use or an up to 25% increase in ER use. There is also no discernible impact on inpatient hospital use on either margin. These estimates appear to reflect the time frame of the survey measure; they are consistent with our findings in the administrative data when we define hospital utilization with the same look-back period as in the survey data.[27]

27. To mimic the survey measures in our administrative discharge data, 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 (std.

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. Because the four different components of utilization have very different expected costs, in the bottom row of Table V we make a back-of-the-envelope calculation of the increase in annual spending associated with insurance by weighting each type of use by its average cost among low-income publicly insured adults in the Medical Expenditure Panel Survey (MEPS). This calculation suggests that insurance is associated with a $778 (std. err. = $371) increase in annual spending, or about a 25% 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 suspect that our one-year estimates are not capturing an initial, transitory surge of pent-up demand for health care among the uninsured. In the survey, conducted about 13 months after insurance coverage began, all of the questions ask about current utilization or utilization over the past six months, not about utilization right after insurance began. Moreover, the evidence in Table XI from the survey conducted about six months after insurance began (when a six-month look-back period likely would have captured the initial effects of insurance) shows no evidence of a larger initial utilization effect or pent-up demand.

Table VI suggests that insurance is also associated with an increase in compliance with recommended preventive care, including blood cholesterol checks, blood tests for diabetes, mammograms, and Pap tests. Overall, the results indicate a 0.3 standard deviation (std. err. = 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% increase in the probability of ever having one's blood

err. = 0.0023), compared to 0.0022 (std. err. = 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 (std. err. = 0.0034) compared to 0.0062 (std. err. = 0.0062) in the survey data. 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.

TABLE VI

COMPLIANCE WITH RECOMMENDED PREVENTIVE CARE (SURVEY DATA)

	Control mean (1)	ITT (2)	LATE (3)	*p*-values (4)
Blood cholesterol checked (ever)	0.625	0.033	0.114	[<0.0001]
	(0.484)	(0.0074)	(0.026)	{<0.0001}
Blood tested for high blood sugar/diabetes (ever)	0.604	0.026	0.090	[0.0004]
	(0.489)	(0.0074)	(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)	

Notes. Standard errors in parentheses; per comparison *p*-values in square brackets; family-wise *p*-values in curly brackets. 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; for the IV estimates in column (3), the endogenous variable *INSURANCE* is defined as "ever on Medicaid" during our study period and the first stage is given in the first row of Table III. 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 ($N = 23{,}741$).

cholesterol checked, a 15% increase in the probability of ever having one's blood tested for high blood sugar or diabetes, a 60% increase in the probability of having a mammogram within the past year (for women age 40 and over), and a 45% change in the probability of having a Pap test within the last year (for women).

V.B. Financial Strain

1. Administrative Data. Table VII, Panel A, analyzes five measures of financial strain in the credit report data: whether the individual has had a bankruptcy, lien, judgment, collection, or delinquency (any credit account with a payment that is 30 days or more late).[28] Broadly speaking, all are measures of (reasonably large) unpaid bills or outstanding obligations that are likely to have a major negative impact on access to credit, at least in a general population (Avery, Calem, and Canner 2003). Liens

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

TABLE VII

FINANCIAL STRAIN (ADMINISTRATIVE DATA)

	Control mean (1)	ITT (2)	LATE (3)	*p*-values (4)
Panel A: Overall				
Any bankruptcy	0.014	0.0022	0.0086	[0.106]
	(0.119)	(0.0014)	(0.0053)	{0.358}
Any lien	0.021	0.0012	0.0047	[0.406]
	(0.144)	(0.0014)	(0.0056)	{0.698}
Any judgment	0.064	0.0014	0.0054	[0.573]
	(0.244)	(0.0024)	(0.010)	{0.698}
Any collection	0.500	−0.012	−0.048	[0.003]
	(0.500)	(0.0041)	(0.016)	{0.013}
Any delinquency (credit accounts)	0.366	0.0016	0.0063	[0.704]
	(0.482)	(0.0042)	(0.017)	{0.698}
Standardized treatment effect		0.0022	0.0086	[0.653]
		(0.0049)	(0.019)	
Panel B: Medical debt				
Any medical collection	0.281	−0.016	−0.064	[<0.0001]
	(0.449)	(0.0040)	(0.016)	{<0.0001}
Amount owed in medical collections	1,999	−99	−390	[0.028]
	(6733)	(45)	(177)	{0.025}
Standardized treatment effect		−0.026	−0.100	[<0.0001]
		(0.0061)	(0.024)	
Panel C: Nonmedical debt				
Any nonmedical collection	0.392	−0.0046	−0.018	[0.264]
	(0.488)	(0.0041)	(0.016)	{0.455}
Amount owed in nonmedical collections	2,740	−20	−79	[0.751]
	(9,492)	(63)	(248)	{0.752}
Standardized treatment effect		−0.0058	−0.023	[0.325]
		(0.0059)	(0.023)	

Notes. Standard errors in parentheses; per comparison *p*-values in square brackets; family-wise *p*-values in curly brackets. 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; for the IV estimates in column (3), the endogenous variable *INSURANCE* is defined as "ever on Medicaid" during our study period and the first stage is given in the first row of Table III. 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 ($N = 49{,}980$).

refer to tax liens, generally taken out by governments for unpaid taxes. Judgments for unpaid bills are sought by a variety of parties, including medical providers, governments, utility companies, collections agencies, and creditors. Collections reflect unpaid bills (mostly not related to revolving credit) that have been sent to collections agencies for recovery attempts. These measures capture only a subset of unpaid bills because not all sources of borrowing are covered in the data, and some unpaid bills may not be sent to collections agencies or have judgments or

liens against them. We return to this later. As the frequencies in Table VII indicate, bankruptcies, judgments, and liens capture relatively extreme "right-tail" events (ranging from 1% to 6% frequency over our 16-month sample period), whereas collections and delinquencies are much more common (about 50% and 37% frequency, respectively).

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 (std. err. = 0.019). Four of the five measures show no significant change. However, health insurance is associated with a decline in the probability of having any unpaid bills sent to collection of 4.8 percentage points (std. err. = 0.016), or about 10% relative to the control mean. This result is highly statistically significant, even after adjusting for multiple tests. Two reasons we may see a decline in collections but not in the other measures are 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.

Another possibility is that there is more of an impact of health insurance on collections because in both our population and the general population, collections are disproportionately medical (Avery, Calem, and Canner 2003). We are able to decompose the presence and size of collections into medical (Panel B) and nonmedical (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 percentage points (std. err. = 1.6), or about 25% relative to the control mean, and a decline in the average (unconditional) amount owed in medical collections of $390 (std. err. = $177), or about 20% relative to the control mean. The corresponding estimates for nonmedical collections are substantially smaller and statistically insignificant.

These results are subject to some potential limitations, discussed in more detail in the Online Appendix. First, not all collections are reported to the credit bureaus, although our investigations did not suggest any reason to suspect reporting

TABLE VIII

FINANCIAL STRAIN (SURVEY DATA)

	Control mean (1)	ITT (2)	LATE (3)	*p*-values (4)
Any out of pocket medical expenses, last six months	0.555 (0.497)	−0.058 (0.0077)	−0.200 (0.026)	[<0.0001] {<0.0001}
Owe money for medical expenses currently	0.597 (0.491)	−0.052 (0.0076)	−0.180 (0.026)	[<0.0001] {<0.0001}
Borrowed money or skipped other bills to pay medical bills, last six months	0.364 (0.481)	−0.045 (0.0073)	−0.154 (0.025)	[<0.0001] {<0.0001}
Refused treatment because of medical debt, last six months	0.081 (0.273)	−0.011 (0.0041)	−0.036 (0.014)	[0.01] {0.01}
Standardized treatment effect		−0.089 (0.010)	−0.305 (0.035)	[<0.0001]

Notes. Standard errors in parentheses; per comparison *p*-values in square brackets; family-wise *p*-values in curly brackets. 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; for the IV estimates in column (3), the endogenous variable *INSURANCE* is defined as "ever on Medicaid" during our study period and the first stage is given in the first row of Table III. 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 ($N = 23{,}741$).

to be correlated with insurance status (see Online 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 (Online Appendix 3.3). 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 "nontraditional" credit sources—which may be particularly important in a low-income population. Given these limitations, the survey measures of financial strain are a useful complement to the credit report measures.

2. Survey Data. Table VIII reports results for four measures of financial strain: whether the respondent has any out-of-pocket medical expenditures in the past six months, currently owes money for medical expenses, had to borrow money (or skip paying other bills or pay them late) to pay medical expenses in the past six months, and has been refused medical treatment

because of medical debt in the past six months. We find a statistically significant decline in all four survey measures of financial strain, including, for example, a 20 percentage point (35%) decline in the probability of having out-of-pocket expenses and a 15 percentage point (40%) decline in the probability of having to borrow money or skip paying other bills to pay medical expenses.[29] The average standardized treatment effect indicates that insurance is associated with a 0.3 standard deviation (std. err. = 0.035) decline in these measures of financial strain.

The results suggest that financial benefits from Medicaid coverage accrue not just to the newly insured. The declines in out-of-pocket expenses and in reported difficulty paying nonmedical bills point to direct financial benefits to the newly insured. At the same time, both the survey data and the credit report data indicate a reduction in outstanding medical bills, the vast majority of which are ultimately never paid. This suggests that some of the financial benefits from Medicaid coverage accrue to medical providers or to whomever they pass on the costs of unpaid care.

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, Figure I, Panels A and B indicate that selection by the lottery is associated with declines in out-of-pocket

29. These results imply that about 35% 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 LATE estimate of the impact of insurance on these expenses is –122 (std. err. = 43), implying that those covered by OHP average $185 in semi-annual out-of-pocket medical expenses. Our impression from focus groups is that these reflect some combination of continued scheduled payments on prior debts, 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) or for uncovered services, such as dental care (even though the survey question explicitly said to exclude dental). Gross and Notowidigdo (2011) similarly find evidence of reported out-of-pocket spending from Medicaid recipients in the MEPS.

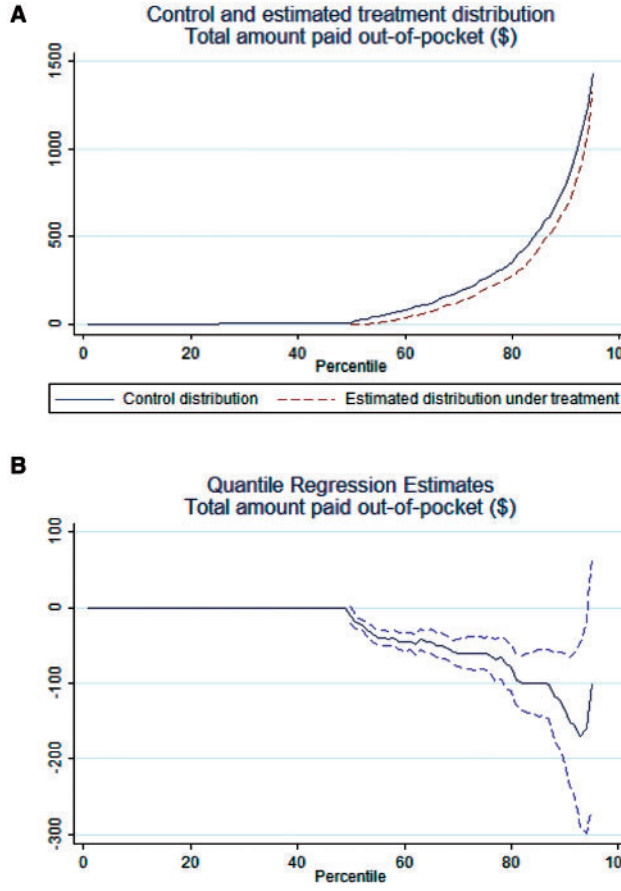

FIGURE I

Distribution of Out-of-Pocket Medical Expenses, Last Six Months (Survey Data)

Panel A 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). Panel B plots the quantile estimates from equation (1) (along with their 95% confidence interval) starting from the smallest quantile that is nonzero 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.

TABLE IX
HEALTH

	Control mean (1)	ITT (2)	LATE (3)	p-values (4)
Panel A: Administrative data				
Alive	0.992	0.00032	0.0013	[0.638]
	(0.092)	(0.00068)	(0.0027)	
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)	(0.0076)	(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)	(0.0051)	(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)	(0.0067)	(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	0.023	0.078	[0.001]
	(0.470)	(0.0071)	(0.025)	{0.003}
Standardized treatment effect		0.059	0.203	[<0.0001]
		(0.011)	(0.039)	

Notes. Standard errors in parentheses; per comparison p-values in square brackets; family-wise p-values in curly brackets. 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; for the IV estimates in column (3), the endogenous variable *INSURANCE* is defined as "ever on Medicaid" during our study period and the first stage is given in the first row of Table III. 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 ($N = 74{,}922$). 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 ($N = 23{,}741$).

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

spending at all the (nonzero) quantiles of the distribution. In the Online Appendix we show similar effects for self-reported medical liabilities and medical collections (Online Appendix Figures A5a–A6b).

V.C. Health

Table IX shows our estimates of the impact of health insurance on health. We have one measure of health from administrative data, mortality, which we measure from the notification date through September 30, 2009. Mortality—although important and objectively measured—is very low in our population; only about 0.8% of the controls died 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% of the population), and (2) to examine "tail" behavior, self-reported health not poor (86% of the population). The other measures are (3) whether self-reported health status is about the same or has gotten better over past six months (versus gotten worse), (4) the number of days in good physical health in past month (0–30), (5) the number of days not impaired by physical or mental health in the past month (0–30), (6) the number of days in good mental health in the past 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 capture only mental health.

The results in Table IX, 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 (std. err. = 0.04) improvement in self-reported health. This includes, among other things, an increase in the probability of screening negative for depression of 7.8 percentage points (std. err. = 2.5) or about 10% 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 (std. err. = 2.6), or about 25% 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 discuss shortly, their interactions with the health care system. Given the subjective nature of the responses, however, it is difficult to judge the extent to which these results reflect improvements in physical health. For mental health, the self-reported and subjective nature of the questions is less of an issue, because diagnosis of depression, by its nature, relies on such self-reports; the depression screen we use correlates highly with clinical diagnoses of depression (Kroenke, Spitzer, and Williams 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, however, 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 X, Panel A) and perceived quality of care (conditional on receipt) (Table X, Panel B). However, there is also evidence that a substantial part of the estimated improvements may reflect a general sense of improved wellbeing. Table XI 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 (this analysis was not prespecified). 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 survey (all those in the main survey except the depression screen). Given the limited time after coverage approval and

TABLE X

POTENTIAL MECHANISMS FOR IMPROVED HEALTH (SURVEY DATA)

	Control mean (1)	ITT (2)	LATE (3)	*p*-values (4)
Panel A: Access to care				
Have usual place of clinic-based care	0.499	0.099	0.339	[<0.0001]
	(0.500)	(0.0080)	(0.027)	{<0.0001}
Have personal doctor	0.490	0.081	0.280	[<0.0001]
	(0.500)	(0.0077)	(0.026)	{<0.0001}
Got all needed medical care, last six months	0.684	0.069	0.239	[<0.0001]
	(0.465)	(0.0063)	(0.022)	{<0.0001}
Got all needed drugs, last six months	0.765	0.056	0.195	[<0.0001]
	(0.424)	(0.0055)	(0.019)	{<0.0001}
Didn't use ER for nonemergency, last six months	0.916	−0.0011	−0.0037	[0.804]
	(0.278)	(0.0043)	(0.015)	{0.804}
Standardized treatment effect		0.128	0.440	[<0.0001]
		(0.0084)	(0.029)	
Panel B: Quality of care				
Quality of care received last six months good/very good/excellent (conditional on any)	0.708	0.043	0.142	[<0.0001]
	(0.455)	(0.0081)	(0.027)	
Panel C: Happiness				
Very happy or pretty happy (vs. not too happy)	0.594	0.056	0.191	[<0.0001]
	(0.491)	(0.0074)	(0.026)	

Notes. Standard errors in parentheses; per comparison *p*-values in square brackets, family-wise *p*-values in curly brackets. 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; for the IV estimates in column (3), the endogenous variable *INSURANCE* is defined as "ever on Medicaid" during our study period and the first stage is given in the first row of Table III. 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 ($N = 23{,}741$).

TABLE XI

ESTIMATED EFFECTS OF LOTTERY AT DIFFERENT TIMES

	ITT estimates from survey			*p*-value of difference between		
	Initial (1)	Six month (2)	Main (3)	Initial and six month (4)	Six month and main (5)	Initial and main (6)
Utilization (extensive margin)	0.0038	0.047	0.050	0.035	0.867	<0.0001
	(0.0084)	(0.020)	(0.011)			
	[0.656]	[0.02]	[<0.0001]			
Utilization (total)	−0.00023	0.027	0.040	0.188	0.556	0.001
	(0.0086)	(0.020)	(0.011)			
	[0.978]	[0.187]	[0.0003]			
Financial strain	−0.035	−0.099	−0.089	0.002	0.613	<0.0001
	(0.0089)	(0.020)	(0.010)			
	[<0.0001]	[<0.0001]	[<0.0001]			
Health	0.042	0.097	0.061	0.014	0.121	0.112
	(0.010)	(0.023)	(0.011)			
	[<0.0001]	[<0.0001]	[<0.0001]			
Access	0.047	0.075	0.119	0.163	0.026	<0.0001
	(0.0078)	(0.019)	(0.0086)			
	[<0.0001]	[<0.0001]	[<0.0001]			

Notes. The analysis in this table was not prespecified. Standard errors in parentheses; per comparison *p*-values in square brackets. 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 (6 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 20percent sample of the other surveys (see Section III.D and Online 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) report 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).

the lack of any increase in health care utilization, it seems likely that this immediate improvement does not reflect changes in objective physical health.[30] 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% the magnitude of the later survey, which again seems more likely to reflect an improved outlook. It is not clear, however, that the immediate effects are directly comparable to those from one year later. Some of the immediate improvements may reflect "winning" effects that are less likely to be picked up in the estimates one year later, and the later estimates of self-reported improvements in health may be biased downward by the impact of health insurance on receiving health care and hence poor health diagnoses, an effect that would not arise in the initial survey because health care use does not appear to go up immediately.

Consistent with an improved overall sense of well-being, there is evidence in the later survey of a substantial (32%) increase in self-reported overall happiness (Table X, Panel C). To put this in perspective, if we compare our estimates to the literature on the impact of income on happiness, the impact of insurance roughly equivalent to the impact of a doubling of income.[31] Of course, it is 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.

30. There is also evidence of a decline in financial strain in the initial survey that is about 40% the magnitude of the analogous measures in the later survey. This is consistent with the fact that coverage was applied retroactively for lottery winners to about one month prior to the approval date. Individuals in the initial survey had therefore been covered retroactively for about two months, or about one-third of the six-month look-back period. Thus, although they may not have changed their health care utilization in the first month of coverage, the retroactive coverage for two months had an impact on their finances.

31. Our IV or LATE estimates in Table X indicate that Medicaid increases happiness by about 0.4 standard deviations. The literature on the relationship between happiness and income tends to find that a log-point increase in income is associated with an increase in happiness of about this magnitude (see, e.g., Sacks, Stevenson, and Wolfers 2010, table 1).

VI. Discussion: Interpreting and Extrapolating the Results

VI.A. Comparison with Other Estimates

We compare our experimental estimates with those that would be obtained by constructing observational estimates in our own data (ignoring the role of the lottery in driving insurance coverage) or in national survey data (see Online Appendix 3.6 and Online Appendix Tables A27 and A28). Compared with our experimental estimates, observational analogs suggest larger impacts of health insurance on health care utilization and 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 who select health insurance coverage are in poorer health (and therefore demand more medical care) than those who are uninsured, just as standard adverse selection theory would predict.

We also compare our estimates to existing experimental estimates. Our estimates of the impact of moving low-income adults from no insurance to Medicaid are not directly comparable to the RAND experimental estimates of varying the extent of coverage among the insured for a representative, nonelderly population. Nevertheless, with this important caveat in mind, our estimates seem slightly smaller than those found in RAND, whereas the improvements in self-reported health are larger.[32] The RAND Experiment did not analyze the impact of health insurance on financial risk exposure.[33] Slightly more comparable to our

32. For example, RAND found that moving from the least comprehensive insurance plan—which still offered considerable insurance coverage to full insurance was associated with a 45% increase in annual spending, whereas our back-of-the-envelope calculation suggested that relative to being uninsured, Medicaid was associated with a 25% increase in six-month spending. The same insurance variation in RAND also produced about a 75% increase in the number of annual outpatient visits, compared with the 55% increase we estimated (Newhouse and the Insurance Experiment Group 1993, p. 41). RAND found no evidence of an impact of insurance generosity on adult self-reported general health or adult mental health (Newhouse and the Insurance Experimenta Group 1993, p. 209).

33. In the United States, we know of only three quasi-experimental studies of the impact of health insurance on risk exposure. Like us, these studies find that health insurance reduces the distribution of out-of-pocket medical expenditures (Finkelstein and McKnight 2008; Englehardt and Gruber 2011) although, unlike us, evidence that Medicaid expansions (at a slightly higher income level) are associated with a decline in personal bankruptcies (Gross and Notowidigdo 2011).

estimates are the results from a contemporaneous randomized trial of the one-year effects of extending public health insurance coverage to uninsured nonelderly adults receiving Social Security Disability Insurance (Michalopoulos et al. 2011). This population is similar in age to the one in this study but had higher income and was in much worse health (with not surprisingly substantially higher health care use); it is also had a much higher fraction nonwhite. Qualitatively, the findings are very similar; in both experiments, insurance is associated with increased use of prescription drugs and primary care (although no discernible impact on ER or inpatient hospital use in survey data), decreased financial strain (such as declines in the right tail of out-of-pocket medical expenditures), and improvements in self-reported health.

VI.B. Extrapolation to Other Contexts

Our results should be interpreted in light of the particular characteristics of the study population and the nature and timing of the policy intervention. Although it is a natural impulse to try to generalize these experimental estimates to other contexts, including the planned 2014 Medicaid expansions, it is important to bear in mind several caveats.

First, our results speak only to the approximately one-year impact of expanding Medicaid access. As discussed, our reading of the available evidence is that the results do not reflect a transitorily large use response stemming from initial pent-up demand. However, the increase in utilization may be larger than what we would find if individuals were not at risk of losing their insurance (and being unable to reapply for it) if they fail to continue to meet the eligibility requirements. The effects of health insurance on the outcomes studied may also change over longer time horizons than we are able to analyze. For example, because health is a stock rather than a flow, the long-run health impacts of health insurance may exceed our one-year estimates. On the other hand, there is evidence from the Moving to Opportunity experiment, which provided housing vouchers to low-income families living in public housing in high-poverty areas, suggesting that short-run improvements in self-reported general health may fade out after a few years, although self-reported improvements in mental health outcomes may persist for longer (Katz, Kling, and Liebman 2001; Kling, Liebman and Katz 2007).

Second, our findings speak to the partial equilibrium effects of covering a small number of people, holding constant the rest of the health care system; the effects of much larger health insurance expansions might differ because of supply-side responses by the health care sector (Finkelstein 2007).

Third, our results are specific to a particular population in a particular health care environment at a particular time period. Our population 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 U.S. population. It is also somewhat (four to five years) older and on some measures appears to be in somewhat worse self-reported health (Allen et al. 2010). These differences are amplified when focusing on compliers, who, relative to the overall lottery population, are somewhat older, more likely to be white, in worse health, and of lower socioeconomic status (as proxied by education and having revolving credit at the time of the lottery). Although we examined heterogeneity in treatment effects by these and other observables, we lacked power to draw precise inferences (see Online Appendix 3.5 and Online Table A26 for details).

In addition to observable demographic differences, our study population voluntarily signed up for the lottery and enrolled in health insurance. The impact of health insurance on, for example, health care utilization may well be larger for those who select into it than for those enrolling when coverage is mandatory; Einav et al. (2011) provide evidence of such "selection on moral hazard" in one specific context. Consistent with this idea, in our setting we find some evidence that the approximately 10% of the sample who signed up the first day the lottery list was open have larger utilization responses than the rest of the lottery participants (see Online Appendix Table A26) (this analysis was not prespecified). The impact of a public health insurance expansion may also vary with the business cycle; our estimates come from a time of substantial economic downturn.

Fourth, the impact of Medicaid may also vary with the specifics of the Medicaid program and the nature of the safety net available to the uninsured. There is a great deal of heterogeneity across states in these dimensions. Our reading of the data, however, suggests that Oregon is not particularly atypical on the

dimensions we can measure. For example, like Oregon, almost all states operate Medicaid managed care programs. Medicaid fees in Oregon are around the median of U.S. states. OHP Standard is fairly generous in its lack of any copayments and its coverage of nonmandated services, although some states cover dental and vision, which OHP Standard does not (Ramirez de Arellano and Wolfe 2007; Kaiser Family Foundation 2011). It appears the safety net may be slightly more generous in Oregon, although other measures of the health care system like the share of admissions in public hospitals and the physician-to-population ratio are quite similar to the national average (Allen et al. 2010).

Finally, the insurance we study was free or heavily subsidized, and as a result 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—were estimated to be about $3,000 in 2001–2004 (Wallace et al. 2008), 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.

VI.C. Conclusion

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. We also found evidence of improvement in self-reported mental and physical health measures, perceived access to and quality of care, and overall well-being. Our results suggest that Medicaid provides benefits to this population above and beyond the non-Medicaid alternatives that exist through various safety-net options. These results are important inputs into a careful cost-benefit analysis of this expansion in Medicaid, although such an analysis would require a number of additional assumptions that go beyond the data that this experiment can provide.

SUPPLEMENTARY MATERIAL

An Online Appendix for this article can be found at QJE online (qje.oxfordjournals.org).

NATIONAL BUREAU OF ECONOMIC RESEARCH AND
MASSACHUSETTS INSTITUTE OF TECHNOLOGY
NATIONAL BUREAU OF ECONOMIC RESEARCH
CENTER FOR OUTCOMES RESEARCH AND EDUCATION AT
PROVIDENCE HEALTH AND SERVICES
NATIONAL BUREAU OF ECONOMIC RESEARCH
NATIONAL BUREAU OF ECONOMIC RESEARCH AND
MASSACHUSETTS INSTITUTE OF TECHNOLOGY
NATIONAL BUREAU OF ECONOMIC RESEARCH AND HARVARD
UNIVERSITY
CENTER FOR OUTCOMES RESEARCH AND EDUCATION AT
PROVIDENCE HEALTH AND SERVICES
NATIONAL BUREAU OF ECONOMIC RESEARCH AND
HARVARD UNIVERSITY

REFERENCES

Alatas, Vivi, Abhijit Banerjee, Rema Hanna, Benjamin A. Olken, and Julia Tobias, "Targeting the Poor: Evidence from a Field Experiment in Indonesia," NBER Working Paper No. 15980, 2010.

Allen, Heidi, Katherine Baicker, Amy Finkelstein, Sarah Taubman, and Bill Wright, and the Oregon Health Study Group, "What the Oregon Health Study Can Tell Us about Expanding Medicaid," *Health Affairs*, 29 (2010), 1498–1506.

Anderson, Michael, "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 (2008), 1481–1495.

Anderson, Michael, Carlos Dobkin, and Tal Gross, "The Effect of Health Insurance Coverage on the Use of Medical Services," NBER Working Paper No. 15823, 2010.

Angrist, Joshua, "Estimation of Limited-Dependent Variable Models with Dummy Endogenous Regressors: Simple Strategies for Empirical Practice," *Journal of Business Economics and Statistics*, 19 (2001), 2–16.

Angrist, Joshua, Guido Imbens, and Donald Rubin, "Identification and Causal Effects Using Instrumental Variables," *Journal of the American Statistical Association*, 91 (1996), 444–472.

Angrist, Joshua, and Steve Pischke. *Mostly Harmless Econometrics: An Empiricist's Companion* (Princeton, NJ: Princeton University Press, 2009).

Avery, Robert, Paul Calem, and Glenn Canner, "An Overview of Consumer Data and Credit Reporting," *Federal Reserve Bulletin*, 89 (2003), 47–78.

Card, David, Carlos Dobkin, and Nicole Maestas, "The Impact of Nearly Universal Coverage on Health Care Utilization: Evidence from Medicare," *American Economic Review*, 98 (2008), 2242–2258.

———, "Does Medicare Save Lives?" *Quarterly Journal of Economics*, 124 (2009), 597–636.

Casey, Katherine, Rachel Glennerster, and Edward Miguel, "Reshaping Institutions: Evidence on Aid Impacts Using a Pre-Analysis Plan." Mimeo, 2011, available online at http://elsa.berkeley.edu/~emiguel/pdfs/miguel_gbf.pdf.

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

Einav, Liran, Amy Finkelstein, Stephen Ryan, Paul Schrimpf, and Mark Cullen, "Selection on Moral Hazard in Health Insurance," NBER Working Paper No. w16969, 2011.

Engelhardt, Gary V., and Jonathan Gruber, "Medicare Part D and the Financial Protection of the Elderly," *American Economic Journal: Economic Policy*, 3 (2011), 77–102.

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

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

General Accountability Office (GAO), "Medicaid and CHIP: Most Physicians Serve Covered Children but Have Difficulty Referring them to Specialists," GAO-11-624, 2011; available online at http://www.gao.gov/new.items/d11624.pdf.

Getzen, Thomas, "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 (2000), 259–270.

Gross, Tal., and Matthew J. Notowidigdo, "Health Insurance and the Consumer Bankruptcy Decision: Evidence from the Expansion of Medicaid," *Journal of Public Economics*, 95 (2011), 767–778.

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

Imbens, Guido W., and Joshua Angrist, "Identification and Estimation of Local Average Treatment Effects," *Econometrica*, 62 (1994), 467–475.

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

———, "Focus on Health Reform: Summary of New Health Reform Law," Publication 8061, 2010b; available online at http://www.kff.org/healthreform/8061.cfm (accessed February 22, 2011).

———, "State Health Facts," (2011); available online at http://www.statehealthfacts.org.

Katz, Lawrence F., Jeffrey R. Kling, and Jeffrey B. Liebman, "Moving to Opportunity in Boston: Early Results of a Randomized Mobility Experiment," *Quarterly Journal Of Economics*, 116 (2001), 607–654.

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

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

Kroenke, K., R. L. Spitzer, and J. B. Williams, "The Patient Health Questionnaire-2: Validity of a Two-Item Depression Screener," *Medical Care*, 41 (2003), 1284–1292.

Lee, David S., "Training, Wages, and Sample Selection: Estimating Sharp Bounds on Treatment Effects," *Review of Economic Studies*, 76 (2009), 1071–1102.

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

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*, 77 (1987), 251–277.

Medicaid Access Study Group, "Access of Medicaid Recipients to Outpatient Care," *New England Journal of Medicine*, 330 (1994), 1426–1430.

Michalopoulos, Charles, David Wittenburg, Dina Israel, Jennifer Schore, Anne Warren, Aparajita Zutshi, Stephen Freedman, and Lisa Schwartz, "The Accelerated Benefits Demonstration and Evaluation Project: Impacts on Health and Employment at 12 months," MDRC, 2011; available online at http://www.mdrc.org/publications/597/full.pdf.
Newhouse, Joseph P. and the Insurance Experiment Group. *Free for All?: Lessons from the RAND Health Insurance Experiment* (Cambridge, MA: Harvard University Press, 1993).
Olken, Benjamin A., Junko Onishi, and Susan Wong, "Indonesia's PNPM Generasi Program: Interim Impact Evaluation Report," (World Bank, 2010).
Office for Oregon Health Policy and Research, "Trends in Oregon's HealthCare Market and the Oregon Health Plan: A Report to the 75th Legislative Assembly," (February 2009).
Ramirez de Arellano, Annette, and Sidney Wolfe, "Unsettling Scores: A Ranking of State Medicaid Programs," Public Citizen Health Research Group, Public Citizen Press, 2007; available online at http://www.citizen.org/medicaid.
Rosenbaum, Sara, "NEJM Perspective: Equal Access for Medicaid Beneficiaries—The Supreme Court and the *Douglas* Cases," *New England Journal of Medicine*, 365 (2011), 2245–2247; available online at http://www.nejm.org/doi/full/10.1056/NEJMp1111428.
Sacks, Daniel, Betsey Stevenson, and Justin Wolfers, "Subjective Well-Being, Income, Economic Development, and Growth," NBER Working Paper No. w16441, 2010.
Schaner, Simone, "Intrahousehold Preference Heterogeneity, Commitment and Strategic Savings: Theory and Evidence from Kenya," MIT Working Paper, 2010; available online at http://econ-www.mit.edu/files/6221.
U.S. GPO, "Patient Protection and Affordability Care Act (HR 3590)," (2010); available online at http://www.gpo.gov/fdsys/pkg/PLAW-111publ148/pdf/PLAW-111publ148.pdf.
Wallace, Neal, John McConnell, Charles Gallia, and Jeanene Smith, "How Effective are Co-Payments for Reducing Expenditures for Low-Income Adult Medicaid Beneficiaries? Experience from the Oregon Health Plan," *Health Services Research*, 43 (2008), 515–530.
Westfall, Peter H., and Stanley S. Young *Resampling-Based Multiple Testing: Examples and Methods for P-value Adjustment* (New York: Wiley, 1993).
Zeckhauser, Richard, "Medical Insurance: A Case Study of the Tradeoff between Risk Spreading and Appropriate Incentives," *Journal of Economic Theory*, 2 (1970), 10–26.