Supplementary Appendix

This appendix has been provided by the authors to give readers additional information about their work.

Supplement to: Baicker K, Taubman SL, Allen HL, et al. The Oregon experiment — effects of Medicaid on clinical outcomes. N Engl J Med 2013;368:1713-22. DOI: 10.1056/NEJMsa1212321

Supplemental Appendix

This Appendix provides additional information on the study population and sample design, our empirical strategy, details on variable construction, and results from additional specifications. Many of these details on design have been previously published (Finkelstein et al. 2012). Most of the analyses presented here were pre-specified in a publicly-archived analysis plan (Baicker et al.), with the exception of a few ex-post specifications denoted by a caret (^) in the tables.

Authorship

This study was conducted by the Oregon Health Study Group. In addition to the named authors, The Oregon Health Study Group includes Matt Carlson (Portland State University), Tina Edlund (Deputy Director, Oregon Health Authority), Charles Gallia (Oregon DHS), and Jeanene Smith (Office for Oregon Health Policy and Research). All authors contributed to study design. Baicker, Finkelstein, and Taubman prepared the first draft of the manuscript and vouch for the analysis. Allen and Wright vouch for the data collection.

Oregon's Medicaid Lottery[1]

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. (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 de- pending 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.

[1] This section is reproduced from Finkelstein et al (2012).

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 aware- ness campaign about the lottery opportunity. To keep barriers to sign-up low, the sign-up form 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 house- hold 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.[2] 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, pro- vided that they recertified their eligibility status every six months.

Study population

Our study population is drawn from individuals included on Oregon's Medicaid lottery list. Figure A1 shows the evolution of the study population from submitting names to the lottery to survey response. Of the 89,824 individuals on the lottery list, 74,922 were left after the exclusions shown (such as giving an address outside Oregon or having died before the lottery drawing). The in-person study sample was limited to the Portland area for logistical reasons and included 10,405 individuals selected in the lottery and 10,340 individuals not selected. Of those sampled for inclusion, a total of 12,229 individuals (6,387 in the treatment group and 5,842 in the control group) responded to the survey by our end date of October 13, 2010.

[2] 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).

Analytic specifications

All of our regression specifications leverage the random assignment from the lottery to make unbiased comparisons between the treatment and control group. The lottery randomly assigned the ability to apply for Medicaid. We can estimate the *effect of lottery selection* by fitting ordinary least squares regressions and comparing the average outcome for all individuals selected in the lottery to the average outcome for all control individuals. This is an intent-to-treat estimate. We can estimate the *effect of insurance* by fitting two-stage least squares regressions (with lottery selection as an instrument for insurance coverage) and estimating the local average treatment effect of insurance. Under standard assumptions discussed below, both the estimates are unbiased. In our main tables (Tables 2-5), we report the estimates of the effect of insurance (referred to in the main tables as "Change with Medicaid Coverage"). Tables S1-S4 show the estimates of the effect of lottery selection as well as the estimates of the effect of insurance.

The difference between the two estimates in our setting can be seen in the following example. Suppose 100 people are drawn to be able to apply for Medicaid and 100 cannot. 25 of those drawn end up enrolling in Medicaid. None of those not drawn enroll in Medicaid. Suppose when we give them the SF8 mental health questionnaire, the average score in the control group is 44.4, but the score among those drawn in the lottery (the treatment group) is 0.5 higher, or 44.9. This 0.5 increase is the impact of being able to apply for Medicaid – or the intent-to-treat. But if we think the only reason scores are higher is because of actual enrollment in Medicaid, then the changes come from the 25 people in the treatment group who ended up enrolled. The average for the whole group drawn in the lottery of 0.5 is the average of a 0 effect for the 75 people who didn't enroll and 2.0 for the 25 who did. [(25% of the sample)* (effect X) + (75% of the sample) *(no treatment effect) = 0.5, so X=2.0.] This 2.0 increase is the effect of enrolling in Medicaid – or the local average treatment effect. The effect of enrolling in Medicaid is 4 times the effect of being able to apply for Medicaid – based on the fact that the effect is coming from the 25% of people drawn in the lottery who actually gain coverage and thus drive any health effects.

The two-stage least squares estimates of the effect of insurance are unbiased (Angrist, Imbens and Rubin 1996), but are specific to those individuals who obtain insurance when selected in the lottery (sometimes called the "compliers"). Although we cannot observe whether any given individual is a complier, we can estimate characteristics of the compliers as a group (Angrist and Pischke 2009). Table S5 shows these estimated characteristics of the compliers. The compliers are, on average, older than the study population in general, more likely to be non-Hispanic whites, and more likely to have been interviewed in English.

Effect of Lottery Selection (Intent to Treat)

Our treatment group are those selected in the lottery and our controls are those who were not. We estimate the intent-to-treat (ITT) effect of winning the lottery (i.e. the difference between treatment and controls) by fitting the following ordinary least squares equation:

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

where *i* denotes an individual and *h* denotes a household. For example y_{ij} might be the self-reported health of individual *i*.

LOTTERY is an indicator variable for whether or not household *h* was selected by the lottery. The coefficient on *LOTTERY* (β_1) is the main coefficient of interest, and gives the average difference in (adjusted) means between the treatment group (the lottery winners) and the control group (those not selected by the lottery); 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"); 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 households of larger household size.[3]

We denote by V_{ih} a second set of covariates that can be included to potentially improve power by accounting for chance differences between treatment and control groups in variables that may be important determinants of outcomes. These covariates are not needed for β_1 to give an unbiased estimate of the relationship between winning the lottery and the outcome, however, as they are not related to treatment status. Our primary analysis does not control for any V_{ih} covariates; the exception to this is in the analysis of the blood pressure measures, where we add adjustment for age (in decile bins) and sex.[4] As a secondary analysis, we will explore whether our results are sensitive to inclusion of V_{ih} covariates, and are reassured that the results look very similar (see Table S16a-d below).

In all of our ITT estimates and in our subsequent instrumental variable estimates (see below), we fit 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 would 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 results could be affected by this functional form choice, especially for outcomes with very low or very high mean probability. We therefore explore the sensitivity of our results to an alternate specification using logistic regression and calculating average marginal effects for all binary outcomes, and are reassured that the results look very similar (see Table S15a-d below).

[3] The proportion of treated individuals who come from households of size 1 is 71.40% (78.40% for controls), from households of size 2 is 28.36% (21.57% for controls) and from households of size 3 is 0.25% (0.03% for controls).

[4] To decide ex ante whether to control for age and sex, we used data from the controls only and regressed each clinical outcome on age and sex. In general, the explanatory power of these controls was low (explaining less than 5% of the variance). The exception was for the blood pressure measures, where age and sex explained as much as 20% of the variance in the outcomes. Based on these results, we included the age and sex adjustment in the primary specification for blood pressure in an effort to increase our statistical power. We present robustness checks using the alternative specifications below.

In all of our analyses we cluster the standard errors on the household identifier since the treatment is at the household level. All analyses are weighted to account for the sampling design of the survey as described below.

Effect of Insurance (Local Average Treatment Effect)

Under the assumption that lottery status is randomly assigned, the intent-to-treat estimates from equation (1) provide an unbiased estimate of the causal effect of winning the lottery (i.e. winning the opportunity 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* itself. We model this as follows:

$$y_{ih} = \pi_0 + \pi_1 INSURANCE_{ih} + X_{ih}\pi_2 + V_{ih}\pi_3 + \nu_{ih} \quad (2)$$

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

$$INSURANCE_{ih} = \delta_0 + \delta_1 LOTTERY_{ih} + X_{ih}\delta_2 + V_{ih}\delta_3 + \mu_{ih} \quad (3)$$

in which the excluded instrument is the variable *LOTTERY*.

We interpret the coefficient on insurance from instrumental variable estimation of equation (2) as the local average treatment effect of insurance, or LATE (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 upon winning the lottery but who would not obtain insurance without winning the lottery (i.e. the compliers).[5]

The LATE interpretation requires the additional identifying assumption that the only mechanism through which winning the lottery affects the outcomes studied is the lottery's impact on insurance coverage. We believe this is a reasonable approximation; in earlier work we discussed potential violations; where we could explore them we did not find cause for concern (Finkelstein et al. 2012).

[5] If 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 Rubens (1996), but about 2 percent of our controls got onto OHP standard through some limited alternative 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, since 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 S9).

Survey weights

We use weights to adjust for several aspects of our survey fielding in all of our summary statistics and regression analyses. Our weighting method allowed us to devote resources to an intensive follow up of non-respondents and to continue fielding in the face of a new lottery that affected a non-random subsample of our population while preserving balance between treatment and control study participants. We first give an overview of the principle used to construct the weights and then give details on each type of weight.

Rationale for weights

We recruited study participants in groups (called "sample releases"). These rolling releases facilitated efficient use of recruitment, tracking, and interviewing resources. There were 44 such releases, roughly weekly, typically of about 450 individuals. Over the course of the fielding, we routinely dropped individuals from our active recruitment sample. There were two reasons for doing so. First, to promote a high response rate, we regularly took a random subsample of the participants who had been released but had not yet responded and instructed the fielding staff to continue active recruitment only on the selected group while dropping the others from the sample ("intensive follow-up drops"). This allowed our staff to devote additional time and effort to potential participants who were difficult to locate or recruit, without diverting too many resources away from the rest of the potential participants. Second, during our fielding the state was conducting a new lottery for OHP Standard (discussed more below). Following each new lottery drawing, we excluded from our active sample individuals selected in the new lottery who had not yet responded to our survey ("new lottery drops"). Between the intensive follow-up drops and the new lottery drops, the active sample was restricted roughly every 2 to 4 weeks. A typical drop removed a few hundred individuals from the active sample.

We adjust for both types of drops using weights constructed on the following principle: within any (even non-random) subset of the original sample base, a randomly selected group can be weighted to stand in for the non-selected remainder, based on the probability of that random selection, without introducing bias. For each of the drops, we can thus construct a weight that corrects for that drop. The final analytic weight is simply the product of all the weights introduced over the course of the study.

Our weighting is roughly analogous to weighting done for censoring or attrition in longitudinal data (Cole and Hernán 2008, Kalton 1986). As in those settings, we weight each observation at each time point by the inverse probability of being in the sample, and we generate overall weights as the product of the weights across all time points. In our setting, the time points correspond to changes in the active sample. We do not need to model the probability of being in the sample since we randomly assigned active sample status and know the probabilities.

Table S6 summarizes the distribution of the weights for the entire sampling base, the recruitment base, and the survey responders. Over the entire sampling base, the weights have a mean close to 1, and there are relatively few extreme weights. For the survey respondents, which comprise the sample we analyze, the weights exhibit even less variance. The average weight is 1.24, with the 5^{th} to 95^{th} percentile range of 0.96 to 2.076. The controls are impacted more by the weights than the treatment group, as they were more likely to sign up for the new lottery (see below).

The primary cost of the weighting is an increase in variance. One way to quantify the cost in variance is to calculate an "effective sample size," which is the unweighted sample size that would have equivalent precision to our weighted sample. Our effective sample size is 4,786 controls and 5,406 treatments (compared to our actual sample of 5,842 controls and 6,387 treatments). The following sections give more detail on the construction of the weights.

Continuous intensive follow-up of non-respondents

For each "intensive follow-up drop" we construct weights as follows. Let N_t be the set of individuals in a specified sample release (or a group of releases) who have not yet completed an interview at time *t*. We select a random subsample F_t from N_t with sampling probability $p_t=|F_t|/|N_t|$. The weights are defined for each individual *i* in the sampling base as:

$$w_t(i) = \begin{cases} \frac{1}{p_t} \; if \; i \; in \; F_t \\ 0 \; if \; i \; in \; N_t \; and \; i \; not \; in \; F_t \\ 1 \; if \; i \; not \; in \; N_t \end{cases} \quad (4)$$

This weighting does not impact any individuals not yet released for fielding or already having completed interviews; it up-weights the subsample selected for intensive follow-up by the inverse of the probability of being selected, so that they stand in for those dropped (who are assigned weight zero).

New state lottery

Early in our fielding period, the state of Oregon began conducting a new lottery for OHP Standard. The state mailed postcards to those on the original list who were not selected (our controls) asking if they would like to be included in this second lottery. Those who returned the postcard were added to the new waiting list and an initial draw was done just from that group. Following that initial draw, the state opened the new waiting list to the general public (including both our controls and our treatments as well people not on our original list); drawings from this list were conducted approximately monthly. Unlike the original 2008 waiting list, the new waiting list remained continuously open: individuals could sign up at any point. As with the original lottery, draws were done on individuals, but the opportunity to apply for OHP (treatment) was extended to the whole household. After each drawing, we probabilistically matched[6] the new waiting list to our study population to identify individuals who were eligible for selection by the state (called "opt-ins") and those who were actually selected in a given drawing (called "selected opt-ins"). By December 6, 2010 the state had selected everyone in our original sample who signed up for the new lottery; we limit our analysis to data collected by October 13, 2010 to avoid having extreme weights.

Given the difficulty in interpreting the "treatment" received by those who were drawn in the new lottery, we chose to drop the selected opt-ins from our recruitment sample. Additional weights are needed to correct for this. For each lottery drawing, the set of opt-ins is not a random sample of our study population: signing up for the new list was optional, and thus subject to the

[6] The matching was done using LinkPlus software.

influence of factors such as underlying health. However, the set of selected opt-ins *is* a random sample of the opt-ins. We were therefore able to use weights to adjust for the sample dropped because of the second lottery using the same principle as above: within any (even non-random) subset of the original study population, a randomly selected group can be weighted to stand in for the non-selected remainder based on the probability of that random selection.

Let O_t be the set of opt-ins in our study population eligible for new lottery drawing on date t. Let S_t be the set of opt-ins selected in drawing on date t. For those released into active fielding and having already completed an interview, the new lottery does not pose any problems. This whole set is assigned weight 1. For who have not yet completed an interview, we define the weight for individual *i* to be:

$$w_t(i) = \begin{cases} \frac{1}{1-p_t} \; if \; i \; in \; O_t \; and \; not \; in \; S_t \\ 0 \; if \; i \; in \; S_t \\ 1 \; if \; i \; not \; in \; O_t \end{cases} \quad (5)$$

where p_t is the probability of an opt-in being selected.

Selection probabilities varied by the number of household members on the new list, so in all cases, we estimated the selection probability separately by strata of "tickets" (household members on the new waiting list at time *t*). Additionally, because of complexity in the release process, the probability p_t depended on whether and when an individual was released. Thus we actually assign these weights in groups of releases where p_t was constant.[7]

Final analytic weights

Each weight variable w_t is designed to adjust for the sampling event at time *t* (whether an intensive follow-up drop or a new lottery drop). We define the cumulative weight variable W_T as the product of all w_t for $t \leq T$. Weighting by W_T ensures that the recruitment sample is representative of the full sampling base. Whenever there is a sampling event, W_T changes appropriately: multiplication by w_t is precisely what is necessary for the recruitment sample to remain representative.

Treatment-control balance

Response rates and balance of respondents

In previous work (Finkelstein et al. 2012), we discuss the random assignment of treatment and control groups. Here we examine treatment and control differences in the subset of the study population who completed interviews. For selected characteristics, the treatment and control differences are reported in Table 1. In that table, we present the control mean, the regression-adjusted treatment mean, and the per comparison p-value. We estimate the regression-adjusted treatment mean by fitting a regression with the characteristic as the dependent variable, an

[7] Due to a technical complication in our sampling, releases 4-21 were stratified on opt-in status with different sampling probabilities for opt-ins and non-opt-ins. We use an additional set of sampling weights to correct for this. The net effect of this is small; the range of these corrective weights is 0.85-2.59.

indicator of lottery selection as the independent variable and controls for household size. The regression-adjusted treatment mean is the sum of the control mean and the regression estimate of lottery selection (the estimated difference). Table S7 reports the results on treatment and control differences for a wide range of characteristics, reporting the control mean, estimated difference and per comparison p-values for each.

Panel A reports the balance on response rates to the survey. Our weighted effective response rate for the controls was 73% and the treatments did not respond at a significantly different rate (0.28 percentage points; 95% CI -2.77 to 3.34; P=0.86). We obtained valid anthropometric and blood pressure data on 98% of respondents and valid blood assays (total cholesterol, HDL cholesterol, and Hemoglobin A1c) on 99% of respondents. Over 98% of respondents either provided medications to be catalogued or reported no medication use (although 8% indicated that this catalog was incomplete). There is no evidence of differential response rates between treatments and controls on any of these components.

Given any response rate of less than 100%, however, there is the potential for bias even if the overall response rate for treatment and control groups is the same: the controls that respond could have systematically different characteristics from the treatments that respond. In Panel B, we examine respondents' balance on characteristics that cannot be affected by lottery selection. Some are measured pre-randomization, taken from the information they provided when signing up for the lottery.[8] Some are measured in the survey but are immutable, such as age or race. Others are characteristics of the data collection effort, such as response date (including season, weekend vs. weekday, etc.), response time (days between start of recruitment and completion of the interview), location of the interview, and language of the interview (English, Spanish, or interpreter of another language). All these variables are intended to help identify potential response bias by capturing characteristics of the responders that may be related to outcomes (men may differ from women, those who chose to come in on the weekend may differ from those who chose to come in during the week, and so on) but are not likely to be affected by the lottery itself. The overall F-stat for differences in all the characteristics pooled has a p-value of 0.84.

In Panel C, we test whether there is any evidence of differential sorting across our interviewers or equipment on the basis of treatment status. We do not expect that there will be differences here, as assignment to interviewer or equipment should not be related to treatment status. As such, we do not want to include all these additional tests in our global test of response bias in Panel B because it could mask real differences between respondents in the characteristics in Panel B. However, because the interviewer or equipment used has such a direct effect on the outcome measurement, we might be concerned about differences even arising from chance.[9] The F-stats for the tests on the three pieces of equipment have p-values of 0.29, 0.23, and 0.73 respectively. The F-stat for the test of sorting across our 48 interviewers has a p-value of 0.150.

[8] Specifically, from the demographic information that the participants provided at the time of lottery sign-up, we construct six pre-randomization variables: 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.

[9] It is worth noting that the division between Panel B and Panel C is not completely clean. For example, interviewers and equipment were assigned to specific clinics.

As a final check of imbalance between treatment and control respondents, we examined differences in pre-randomization characteristics measured in other, administrative, datasets. We examine whether treatments and controls differed in having any hospital visits or the number of hospital visits in the pre-randomization period (as measured in hospital discharge data) or in having any medical or non-medical collections (as measured in credit report data). These datasets are described by Finkelstein et al. (2012), and this analysis follows Table A13 from that paper. There is no evidence of any difference; the F-stat for the test on these four variables combined has a p-value of 0.19 (not shown).

Although these results are not conclusive—there is still the possibility of differences on other unobserved variables—they are reassuring. To the extent that we are able to examine it, we find no evidence of differential selection into responding between treatment and control groups.

Pre-lottery diagnoses

In some of our analyses (see penultimate row of Table 2, and also Tables S14b and S14c), we limit the sample to individuals who report having pre-randomization diagnoses of specific health conditions. Table S8 examines the balance of treatment and control respondents on reports of pre-randomization diagnoses for ten conditions. Participants are considered to have a pre-randomization diagnosis if they reported in their interview having a specific diagnosis first made before March 2008. The multivariate F-statistic for differences in all these conditions pooled has a p-value of 0.30; the standardized treatment effect for change in diagnosis of all these conditions is -0.0026 standard deviations. This suggests that there is no differential reporting of pre-randomization diagnoses by treatment or controls. In addition to observing balance on the individual conditions, we also see no evidence of imbalance on a composite measure for having a pre-randomization diagnosis of diabetes, hypertension, high cholesterol, heart attack or congestive heart failure (estimated average difference is -0.26 (standard error =0.9; p value is 0.77). We use this composite measure to identify a subset of our population that is at increased risk of adverse cardiovascular outcomes. This subset does not include those also at increased risk who have not been diagnosed pre-randomization because we have no way to identify them.

Insurance coverage

For the purposes of estimating the local average treatment effect of insurance, we define our insurance measure as "ever on Medicaid during the study period." Table S9 reports the control means and effects of lottery selection for this and alternative definitions of insurance coverage.

The primary source of data on insurance coverage (including our baseline measure "ever on Medicaid during study period") come from data the state provided us on Medicaid enrollment. These data are described in more detail elsewhere (Finkelstein et al. 2012). Table S9 also reports results from some alternative measures of insurance coverage, including self-reported insurance coverage from our in-person survey.

There are two distinct Oregon Medicaid programs: the program for the traditional Medicaid population (OHP Plus) and the program for the expansion population (OHP Standard). We define someone as ever on "Medicaid" if they are on either Medicaid program, including both Plus and Standard.

Since the lottery was for the OHP Standard program, that is where we would expect to find increases in coverage, and this is borne out in the data. In fact, the increase in OHP Standard is

slightly greater than the increase in any Medicaid (26.49 percentage points compared to 24.14), suggesting that a few percentage points of the increase in OHP Standard may have come from individuals who would have been on another Medicaid program at some point during the study period.

The effect of the lottery on Medicaid coverage attenuates over time: using "current" enrollment (measured on the date of interview) reduces the lottery effect on insurance coverage from 24.14 (row 1) to 11.35 (row 4). There are two reasons for this. First, those who successfully enroll in OHP (through the lottery or other means) are required to recertify eligibility every six months, leading to attrition in coverage. Additionally over time, those not selected in the lottery may obtain Medicaid coverage through the OHP Plus program. Figure A2 shows the time path of enrollment in OHP Standard and all Medicaid over time for both treatments and controls.

Because the initial take-up of Medicaid was relatively low, lottery selection is associated with an average increase of 4.16 months on Medicaid (row 3) – both because only a subset of those selected in the lottery obtained coverage and because those who obtained coverage were not necessarily covered for the entire study period. For those who did obtain coverage through the lottery, there is an increase of 17.24 months on Medicaid (95% CI: 16.49 to 17.99; P<0.001). This is less than the average of 25 months of the study period[10] for several reasons. Lottery selection occurred in 8 draws between March and October 2008, initial enrollment in OHP took 1-2 months after lottery selection, and some of those enrolled in Medicaid through the lottery lost coverage by failing to recertify.

Unlike the administrative data that capture only Medicaid coverage, the interview data capture all sources of insurance (including private coverage). The difference in Medicaid coverage associated with the lottery as measured in the interviews is similar to the difference in Medicaid coverage as measured in the administrative data on the same date. The increase in any insurance coverage is similar to the increase in Medicaid coverage, suggesting that the lottery had little impact on non-Medicaid insurance coverage. Specifically, we see no evidence of crowd-out of private insurance; private insurance rates are unchanged by the lottery.

Outcome measures

The outcomes in this analysis are drawn from the physiological measures and in-person questionnaire (see the protocol document available with this article at nejm.org). We developed the questionnaire for this study, drawing on existing survey instruments whenever possible.

[10] For the purposes of this paper, we define the study period as beginning on March 10, 2008, which is the first date that anyone was notified of being selected in the lottery. In Finkelstein et al. 2012, we used a slightly different definition of the study period based on individual notification dates (which vary across the 8 lottery draws from March to October). Using the same definition as in Finkelstein et al., our average survey response occurs 22 months after notification (standard deviation = 4 months) or 20 months after insurance approval (standard deviation = 4 months). By contrast, in the Finkelstein et al. earlier study, those time periods were, on average, 15 months (std. dev. = 3 months) after notification and 13 months after insurance approval (std. dev. = 3 months) for the mail survey data and 16 months after notification (std. dev. = 2 months) and 14 after insurance approval months (std. dev. = 3 months) for the administrative data. This current study is therefore based on data that is, on average, about 6 to 7 months after the data in Finkelstein et al. (2012).

Table S10 provides a summary of the outcome variables and Table S11 provides additional detail on the distribution of some variables. The outcomes fall into several broad domains.

Clinical measures of health

Our clinical health measurements were modeled on those done by the National Health and Nutrition Examination Survey (NHANES); we worked with consultants from the National Center for Health Statistics to develop them. For each we examine a continuous measure and define binary indicators for higher health risk (worse outcomes) based on clinical guidelines.

We measured blood pressure using the OMRON IntelliSense unit, model HEM-907XL, which automatically inflates the cuff to the desired level and does not require adjustment by the interviewer. Our blood pressure measure is the average of three readings taken 30 seconds apart, following a period of sitting quietly for 5 minutes. We examine continuously measured **systolic blood pressure** and continuously measured **diastolic blood pressure**. We define **pre-hypertension or hypertension** using the standard clinical cut-points of systolic blood pressure of at least 120 or diastolic blood pressure of at least 80 and **hypertension** as systolic blood pressure of at least 140 or diastolic blood pressure of at least 90 (Chobanian et al. 2003).

We collected up to 5 drops of whole blood from a finger stick. Samples were collected on Whatman 903 specimen-collection paper and dried and stored following established protocols (McDade, Williams and Snodgrass 2007). The University of Washington Department of Laboratory Medicine performed the assays from the stored blood. Dried blood spot measurements were converted to clinical values using protocols and formulas developed by the National Center for Health Statistics and Thomas McDade for the Moving to Opportunity Study.

We examine continuously measured **total cholesterol**. We define **elevated cholesterol** as total cholesterol greater than or equal to 200 mg/dL and **high cholesterol** as total cholesterol greater than or equal to 240 mg/dL (Expert Panel on Detection Evaluation And Treatment of High Blood Cholesterol In Adults 2001). We also examine continuously measured ("good") **HDL cholesterol**. We define **low HDL cholesterol** as HDL cholesterol below 40 mg/dL. We do not have a separate measurement of LDL cholesterol.

We examine continuously measured **glycated hemoglobin**. We defined **elevated risk of diabetes** as glycated hemoglobin of at least 5.7% and **diabetes** as glycated hemoglobin of at least 6.5% (International Expert Committee 2009, American Diabetes Association 2010).

Table S12 compares the control means for our measures of blood pressure, cholesterol, and glycosylated hemoglobin to data from the NHANES from 1999-2008. We limit to four subgroups of the NHANES data: adults with income less than 200% of the federal poverty level, adults with income less than 200% of the federal poverty level and no health insurance, adults with income less than 100% of the federal poverty level, and adults with income less than 100% of the federal poverty level and no health insurance. We see that our study population looks roughly similar to a national sample of low-income adults, regardless of insurance status, on these physiologic measures of health.

We examine whether the participant **screened positive for depression** based on the 8-question version of the Patient Health Questionnaire (PHQ-8). The Patient Health Questionnaire is a standard scale for measuring depression (Kroenke et al. 2009) and is used for measuring depression prevalence in the US population in both the NHANES (Shim et al. 2011) the BRFSS (Kroenke et al. 2009). The PHQ-8 asks about the frequency of eight depression symptoms. The summary score is calculated by assigning a score of 0 – 3 for each question of the questionnaire

(0 for not at all; 3 for nearly every day) and then summing those scores, so higher scores indicate more severe depression symptoms. The positive depression screen is based on a cut-point of PHQ-8 summary score of 10 or above. Using a cut-point of 10 or above for depression in a 9-question version of the PHQ has been shown to correlate highly with clinician diagnosis of major depressive disorder (Kroenke, Spitzer and Williams 2001). The PHQ-8 is a modified version of the 9-question version differing only in excluding the question about suicidal ideation (which is rarely answered in the affirmative, and thus makes little substantive difference in scores (Huang et al. 2006)).

Diagnosis and medication

We construct our measures of pre- and post-lottery diagnosis using survey data. We asked respondents if they have ever been diagnosed with each of hypertension, high cholesterol, diabetes, and depression. For those answering in the affirmative, we asked for detail on when they first received that diagnosis. We consider participants to have post-lottery first diagnosis if they reported having a specific diagnosis first made in March 2008 or after. Those never having received a diagnosis or having received it prior to March 2008 were considered not to have a post-lottery first diagnosis.

We construct our information on medications for hypertension, high cholesterol, diabetes, and depression using data from our medication cataloging. If an individual reported any medication use in the last 4 weeks, we took a detailed inventory of the actual medications. Participants were asked to bring all current medications (prescription and over-the-counter) to their interview. For each medication, interviewers asked whether the medication was prescribed or over the counter. If the participant had not brought all current medications to the interview, a phone follow-up was attempted to obtain any remaining medications. Of the 68% of participants who reported any medication use, 12% said that they did not provide all medications at the interview and did not complete a phone follow-up.

The interviewers entered information (including medication name, dosage, frequency and route) on each medication through an interface that looked up records in a drug database obtained from First DataBank. This drug database codes medications into classes, with drugs with multiple uses having multiple class codes. We use these classes to identify indications with input from a physician. For example, we considered anyone taking a medication classified as an antidepressant to be taking medication for depression (even though that drug may have been prescribed for a different indication). Table S13 lists the names of medications considered as treating each of hypertension, high cholesterol, diabetes, and depression

Framingham risk score

We use a sex-specific multivariable point-mapping system to calculate the probability of specific atherosclerotic cardiovascular disease (CVD) events, i.e., coronary heart disease, cerebrovascular disease, peripheral vascular disease, and heart failure (D'Agostino et al. 2008). This system, derived using data from the Framingham Heart Study, incorporates age, total and high-density lipoprotein (HDL) cholesterol, systolic blood pressure, treatment for hypertension, smoking, and diabetes status. For each of these variables, a number of points between -3 and 12 is allotted. Total points are then aggregated across all variables and mapped to a probabilistic risk of CVD events in the next 10 years. To calculate the CVD risk score, we used gender and age variables from survey responses, as well as smoking status (see *health behaviors* below).

The measures of total and HDL cholesterol, systolic blood pressure and diabetes are described in *clinical measures of health* above. A person is considered "treated for hypertension" if one or more medication from her medication survey was classified as a hypertension medication (more detail in *diagnosis and medication* above). The Framingham CVD risk score is only defined in those aged 30 or older.

Health-related quality of life and happiness

We use survey data to code whether **health status has stayed the same or gotten better over last twelve months** (vs. gotten worse).

Our survey included the Medical Outcomes Survey Short Form (SF-8) and we examine both the **SF-8 physical component score** and **SF-8 mental component score**. The SF-8 is a short form (8-item version) of the Medical Outcomes Survey designed to measure health-related quality of life (Ware et al. 2001). The eight questions ask about general health, work, physical and social limitations, pain, energy levels and emotional problems. Each response is assigned a score, and the physical component score and mental component summary are both sums of those scores using different weightings. The scoring is designed so that the summary scores will be comparable to scores obtained using the validated SF-36 (McHorney, Ware and Raczek 1993). The scores range from 0 to 100, with higher scores indicating better health, and are normalized to a mean of 50 and standard deviation of 10 in a general population sample. We use one of the SF-8 questions separately to capture whether you **had no or only very mild pain** in the past 4 weeks.

We use a question about self-reported health to construct two binary measures: self-reported **health good, very good or excellent** (vs. fair, poor, or very poor) and self-reported **health fair, good, very good, or excellent** (vs. poor or very poor). These differ in the handling of the 26 percent of participants reporting fair health. These two measures are not reported in the main text, but are included in the comparison to previous results in Table S18a.

We also asked about how individuals were feeling in general, and we construct a measure of being "**very happy" or "pretty happy"** as compared too "not so happy."

Financial hardship

We consider several measures of financial strain based on out-of-pocket spending, medical debt, and borrowing money or skipping paying bills because of medical debt.

Our module on health care use and costs was based on the Health and Retirement Study questionnaires (Health and Retirement Study 2000). We asked survey participants about out-of-pocket spending in the last year for their own doctor visits, ED visits, outpatient surgeries, hospital visits, dental care, and "other" medical care not included in the first five categories. The survey also asked about monthly out-of-pocket prescription medication costs, which we converted to estimated yearly costs by multiplying by 12. Participants were only asked about a given category of spending if they reported use of that category of medical care.

In cases where participants could not give a close estimate of how much they spent in a given category, they were asked follow-up probes that broke spending into nine possible intervals. We incorporate answers to probes into total spending estimates using the midpoint of each probe interval, except for top-coded intervals, which we coded as their lower bounds.

We define **any out-of-pocket spending** as occurring when the individual reported non-zero spending in at least one of the following: doctor visits, ED visits, outpatient surgery, hospital

visits, prescription drugs, or other medical care. We did not include out-of-pocket spending on dental care because such care is not covered by OHP. We define the **amount of out-of-pocket spending** as the sum of reported spending for the same categories (doctor visits, ED visits, outpatient surgery, hospital visits, prescription medications, and other medical care). We treat the sum as missing if any of the component measures was missing. We truncate the amount of out-of-pocket spending at $2*99^{th}$ percentile, recoding outliers to missing.

We define **catastrophic expenditures** as occurring when the amount of an individual's reported out-of-pocket spending on himself exceeded 30 percent of reported household income. Household income was reported in brackets; for this calculation we used the midpoint of each bracket and the lower bound of the top bracket ($50,000). Different studies use different definitions for catastrophic expenditures based on the share of total income or post-subsistence income (Xu et al. 2003, King et al. 2009). We use a cut-point of 30 percent of income following King et al., but we use total income (whereas they used post-subsistence income), as there is no clear way to separate subsistence and post-subsistence income in our data.

The questions on medical debt were taken from our 12-month mail survey (Finkelstein et al. 2012). We define **any medical debt** by the individual's response to the question "Do you currently owe money to a health care provider, credit card company, or anyone else for medical expenses?" We define **borrowed money or skipped paying other bills** by the individual's response to the question "In the last 12 months, have you had to borrow money, skip paying other bills, or pay other bills late in order to pay health care bills?"

Health care utilization

We consider five utilization categories: prescription drugs, doctor's office visits, outpatient surgery, emergency department visits, and hospital visits. Our survey module on utilization was based on the Health and Retirement Study questionnaires (Health and Retirement Study 2000). We asked about each kind of health care visit separately. In cases where the participant could not give a close estimate of how many visits, we asked for the best guess between zero, one, and more than one visit. If the answer to the probe was "more than once," we code it as if the individual had 2 visits in the last 12 months. Less that 0.2% of answers are imputed from probes for each variable. We truncate each of the number of visits measures (office, outpatient surgery, emergency room, and hospital) at $2*99^{th}$ percentile, recoding outliers to missing. The cut-points for truncation and percent truncated are shown in Table S11.

If an individual reported any medication use in the last 4 weeks, we took a detailed inventory of the actual medications (as described in the *diagnosis and medication conditions* section above). For **number of current prescription drugs**, we counted up all medication records that could be identified as prescription drugs from the medication survey, after removing duplicates. We note that the number of prescription drugs is likely an underestimate because for 8% of respondents the medication catalog was incomplete.

We define **number of office visits in the last 12 months** by the individual's response to the question "In the last 12 months, about how many times have you seen a doctor or other health care professional at a doctor's office, a clinic, or at home?"

We define **number of outpatient surgery visits in the last 12 months** by the individual's response to the question "In the last 12 months, how many times have you had outpatient surgery?" Almost everyone who reports an outpatient surgery visit also reports an office visit, as

we would expect since such surgery would likely require associated office visits (for diagnosis, pre-operative consultation or post-operative follow-up).

We define **number of ED visits in the last 12 months** by the individual's response to the question "In the last 12 months, about how many times have you gone to an emergency room or urgent care clinic?"

We define **number of hospital visits in the last 12 months** by the individual's response to the question "In the last 12 months, how many times have you had to stay in a hospital at least overnight?"

Annual spending estimation

To calculate the implied annual spending effects associated with the estimated utilization effects we use data from the 2002-2007 (pooled) Medical Expenditure Panel Survey (MEPS) on expenditures of all nonelderly (19-64) adults below 100 percent of poverty who are publicly insured. This gives us a total sample of over 7,500 individuals. We use their expenditures (all inflated with the CPI-U to 2007 dollars) to calculate average expenditures per outpatient visit (including office visits, outpatient surgery, and outpatient visits to any other facilities), average expenditures per ED visit, average expenditures per inpatient visit (for visits not related to childbirth). For medications, we calculate average spending per prescription drug by dividing total annual prescription drug costs by the total number of prescription drugs taken over the course of the year. All spending numbers are bases on total expenditures (i.e. not just expenditures among the insured or covered by insurance). The underlying costs are $150 per outpatient visit, $435 per ED visit, $7,523 per inpatient visit, and $312 per prescription drug. For each type of utilization we observe (office visit, outpatient surgery, ED visit, inpatient visit and prescription drug), we multiply the estimated change in number by the cost per visit estimated in the MEPS. For both office visits and outpatient surgery, we use the $150 outpatient visit estimate which is calculated across all outpatient visits (including doctor office visits, outpatient surgery and other outpatient visits).

Prevention

Our module on preventive care and screening used questions from the BRFSS (Centers for Disease Control and Prevention (CDC) 2000) for blood stool tests, colonoscopy, pap smear, mammogram and PSA tests. It also included questions from the NHANES (Centers for Disease Control and Prevention (CDC) National Center for Health Statistics (NCHS) 2000) on cholesterol-level screenings. We only asked individuals about their use of preventive care if they reported having used any medical care in the past 12 months. This means we do not know about use of preventive care prior to the past 12 months in individuals who did not report any medical care in the past 12 months. To avoid potential bias from these missing data, we have adopted a 12-month time frame for all types of preventive care, even when the recommended interval is longer.

We consider an individual as having had a **cholesterol-level screening in the last 12 months** (1) if the individual answered "yes" to the survey question "Has a doctor or other health professional ever told you that you had high cholesterol?" and "within the last year" to the survey question "When were you first told that you had high cholesterol?", or (2) for individuals who have not been diagnosed with high cholesterol, if the individual answered "within the last year" when asked "long has it been since you last had your cholesterol checked?" Cholesterol

testing is recommended every 5 years starting at age 20 (Expert Panel on Detection Evaluation And Treatment of High Blood Cholesterol In Adults 2001).

We consider an individual as having had a **fecal occult-blood test (blood stool test) in the last 12 months** if the individual answered "Yes" to the survey question "In the last 12 months, has a doctor asked you to do a blood stool test?" We consider an individual as having had a **colonoscopy in the last 12 months** if the individual answered "Yes" to the survey question "In the last 12 months, have you had a sigmoidoscopy or a colonoscopy?" We do not look at blood stool tests or colonoscopies for individuals younger than 50. The U.S. Preventive Services Task Force recommends screening using blood stool test, colonoscopy, or sigmoidoscopy for colorectal cancer in all adults beginning at age 50 years and continuing until age 75 years. The Task Force recommended annual screening with high-sensitivity blood stool test, or sigmoidoscopy every 5 years coupled with high-sensitivity blood stool test every 3 years, or screening colonoscopy every 10 years (U.S. Preventive Services Task Force 2008).

We consider an individual as having had a **flu shot in the last 12 months** if the individual answered "Yes" to the survey question "Have you had a flu shot in the last 12 months?" We do not look at flu shots for individuals younger than 50. Although the Advisory Committee on Immunization Practices (ACIP) in the Center for Disease Control (CDC) recommends annual flu shots for everyone older than 6-months, it also recommend priorities to be given to young children and those 50 or older in cases of limited supply (Fiore et al. 2010).

We consider a woman as having had a **papanicolaou smear in the last 12 months** if she answered "Yes" to the survey question "In the last 12 months, have you had a Pap test or Pap smear?" This variable is not applicable to men. The U.S. Preventive Services Task Force recommends initial screening for cervical cancer with Pap smear or liquid-based cytology starting by age 21 years or approximately 3 years after the first sexual intercourse. Future screenings should occur every year with a traditional Pap smear or every 2 years with liquid-based cytology. At or after age 30 years and with three normal test results, intervals can be decreased to every 2 to 3 years with traditional Pap smear or every 3 years with HPV assay testing plus cervical cytology (U.S. Preventive Services Task Force 2003).

We consider a woman as **having had a mammogram in the last 12 months** if she answered "Yes" to the survey question "In the last 12 months, have you had a Mammogram?" This variable is not applicable to men, and we limit to women 50 or older. According to updated guidelines in 2009 from the U.S. Preventive Services Task Force, biennial screening mammography is recommended for women aged 50 to 74 years (U.S. Preventive Services Task Force 2009).

We consider a man **having had a PSA test in the last 12 months** if he answered "Yes" to the survey question "In the last 12 months, have you had a blood test to check for prostate cancer?" This variable is not applicable to women, and we limit to men 50 or older. PSA screening may not be beneficial, and the US Preventive Task Force recently circulated draft recommendations against such screening (Chou et al. 2011). It is, however, quite common, with 54% of the U.S. men aged 50-64 reporting have received a test in the last year,[11] and as of 2009, American Urological Association and American Cancer Society have recommended that early

[11] Estimated from 2008 BRFSS data, N= 28380.

detection begin at age 50 years for men at average risk of prostate cancer (Greene et al. 2009). We include it because access to health insurance may increase use of commonly used tests, even if those tests are of limited value.

Access and quality

We considered a number of questions on individuals' access to health care. We asked if they **had a usual place of clinic care**. We defined this to exclude emergency rooms but include all doctors' offices in a hospital, a private clinic, or a community health center. We also asked if individuals needed medical care in the last 12 months, and if so, whether they received all needed care. These questions focused on care for a physical illness, injury, or condition and excluded dental care or routine vision services. We consider people having **gotten all needed medical care in the last 12 months** if they reported needing care and receiving all needed care or if they reported not needing care (23% of controls).

We asked individuals to rate the **quality of care they received in the past 12 months,** conditional on receiving care, and analyzed if it was **good, very good, or excellent**, vs. **fair or not so good**. This measure of quality of care is defined for the 78% of participants who reported receiving any medical care (include office visits, outpatient surgery, emergency room visits, hospital stays, and other care).

Smoking and obesity

We measured weight using a Seca 876 portable digital weight scale and height using a Seca 214 portable stadiometer. We calculate body mass index (BMI) as a function of height and weight. We define whether you are **obese** as BMI of at least 30. This is a standard clinical cut-points (Expert Panel on the Identification Evaluation and Treatment of Overweight and Obesity in Adults 1998, World Health Organization 2011) and is used by the NHANES in estimating the prevalence of obesity in the US population (Flegal et al. 2010)

Individuals answering "yes" to the survey question "Are you currently smoking?" are considered to be **currently smoking**.

Additional analysis

Pre-specified subgroups

Table S14 reports analysis for our health measures limited to pre-specified subgroups. In Table S14a the analysis is limited to those aged 50-64. In Table S14b the analysis is limited to those with a pre-randomization diagnosis making them "high-risk" for adverse outcomes. We use this composite measure, including diabetes, hypertension, high cholesterol, heart attack and congestive heart failure, rather than measure-specific diagnoses, in order to have a reasonable sample size in the limited group and because of the patterns of comorbidity. Individuals with any of these conditions are at increased risk of adverse cardiovascular outcomes and would particularly benefit from care and management. In Table S14c the analysis is limited to only those with the related pre-randomization diagnosis: the blood pressure measures limited to those with a pre-randomization diagnosis of hypertension, the cholesterol measures limited to those with a pre-randomization diagnosis of high cholesterol, the glycosylated hemoglobin measures limited to those with a pre-randomization diagnosis of diabetes, and the depression measures

limited to those with a pre-randomization diagnosis of depression or anxiety. These three sets of subgroup analysis were all selected as being of particular interest a prior and pre-specified. As we were interested in the effects in these groups per se, rather than the comparison between those in these groups and not in these groups, we did not test for the heterogeneity of effects. We did not estimate effects in any other subgroups.

Sensitivity of results

Tables S15a-d reports analysis of our binary outcome measures using a logistic regression rather than a linear probability regression model. We present the control means and linear probability results (also shown in Tables S1-S4) as well as average marginal effects from the logistic model estimation. The average marginal effects are calculated by predicting the outcome as a treatment and as a control for each individual, taking the difference in the two predictions, and averaging those differences across the whole sample. The results are robust to the choice of model specification.

Tables S16a-d reports the sensitivity of our results to the inclusion of covariates. We present the unadjusted results (also shown in Tables S1-S4, except for blood pressure outcomes), the age-and-sex adjusted results (also shown in Table S1 for blood pressure), and a set of "fully adjusted" results. Age-and-sex adjustment includes fixed effects for sex and age (in decile bins). The fully adjusted model adds to the age-and-sex adjustment additional fixed effects for black race, other non-white race, and Hispanic ethnicity (all as measured in the survey) and six "lottery list variables." These variables are constructed from the information that participants provided at the time of lottery sign-up and include: whether English is the preferred language for receiving materials; whether the individuals signed them- selves 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. Our results are robust to the choice of included covariates.

Table S17 reports analysis of alternate cut-points for our physical measures. We present results for less stringent definitions (e.g. pre-hypertension or hypertension as opposed to hypertension only in the main results). The definitions of these cut-points are given in the section on *Clinical measures of health* above. As with our primary clinical measures, we do not find any change in these clinical health measures.

Comparison to previous results

Table S18 compares our results from the in-person interview to results from earlier work. We present 4 sets of results: results that were published previously (Finkelstein et al. 2012) based on responses to a mail survey fielded approximately 12 months after the lottery the in-person results for all in-person interview respondents (mostly presented in Tables 2-5 of the main text), the mail survey results limited to those who were sampled and completed an in-person interview and the in-person results limited to those who were sampled and responded to the mail survey. The first two sets of results are replicated here for ease in comparison; the second two sets are to present the results limited to the overlapping sample. To further assist in the comparison, we present standardized treatment effects, which are discussed in more detail elsewhere (Finkelstein et al. 2012). In most cases, we limit to the questions asked in both surveys, which means that the

standardized treatment effects presented here do not summarize over the same set of variables as those previously published.

The results from the two studies may vary for a variety of reasons. The targeted samples for the surveys differ, with the in-person interview sample being limited to the Portland-metro area. The response rates differ, with a weighted response rate of 50% in the 12-month mail survey compared to 73% in the in-person interview. The sample sizes differ, with 23,741 respondents to the mail survey, 12,229 respondents to the in-person survey and 5,750 individuals in the overlap sample. The time-frame of analysis is different, with the 12-month mail survey data having an average response date of September 23, 2009 and the in-person interviews having an average response date of April 23, 2010. Finally, the survey mode differed, and mail surveys may elicit different responses than in-person interviews, even to similarly worded questions.

On self-reported health and happiness, we have broadly similar findings from the two studies in terms of Medicaid reducing depression and the proportion reporting health staying the same or getting better. However, we no longer find improvements in general self-reported health status or in happiness, as we did with the previous study. For happiness, limiting to the overlap samples, we do see improvements in both studies. Overall, our findings on financial hardship and on utilization and spending are quite similar for the two studies. The financial hardship standardized treatment effects are almost identical. For utilization and spending, they are also very similar. When limiting to the overlap sample, however, the utilization and spending results become very noisy.

As in the previous study, we continue to find that Medicaid increases the use of preventive care and perceived access and quality. As in the previous study, we find no change in smoking status.

Comparison to observational estimates in the same setting

Table S19 compares our experimental estimates on the clinical health measures to what we might have estimated using observational data. The first column replicates our LATE estimates that used the lottery as an instrument for insurance coverage. We then compare outcomes for the insured to those for the uninsured in our sample: the next columns present various "as treated" comparisons of people with and without insurance within our full study population (column 2), our control group (column 3), and our treatment group (column 4). Unlike our LATE estimates, these observational comparisons capture the effect of endogenous take-up of Medicaid, which may be driven by factors including selection in completing the application or coverage generated by visiting a provider.

These results highlight the importance of random assignment in identifying the impact of insurance coverage. In particular, observational comparisons of those with and without insurance could suggest that expanded insurance coverage is associated with increased rates of depression, while our estimates using random assignment indicate that insurance in fact decreases the rates of depression.

Figures and Tables

Figure S1: Enrollment, treatment assignment, sampling and survey response

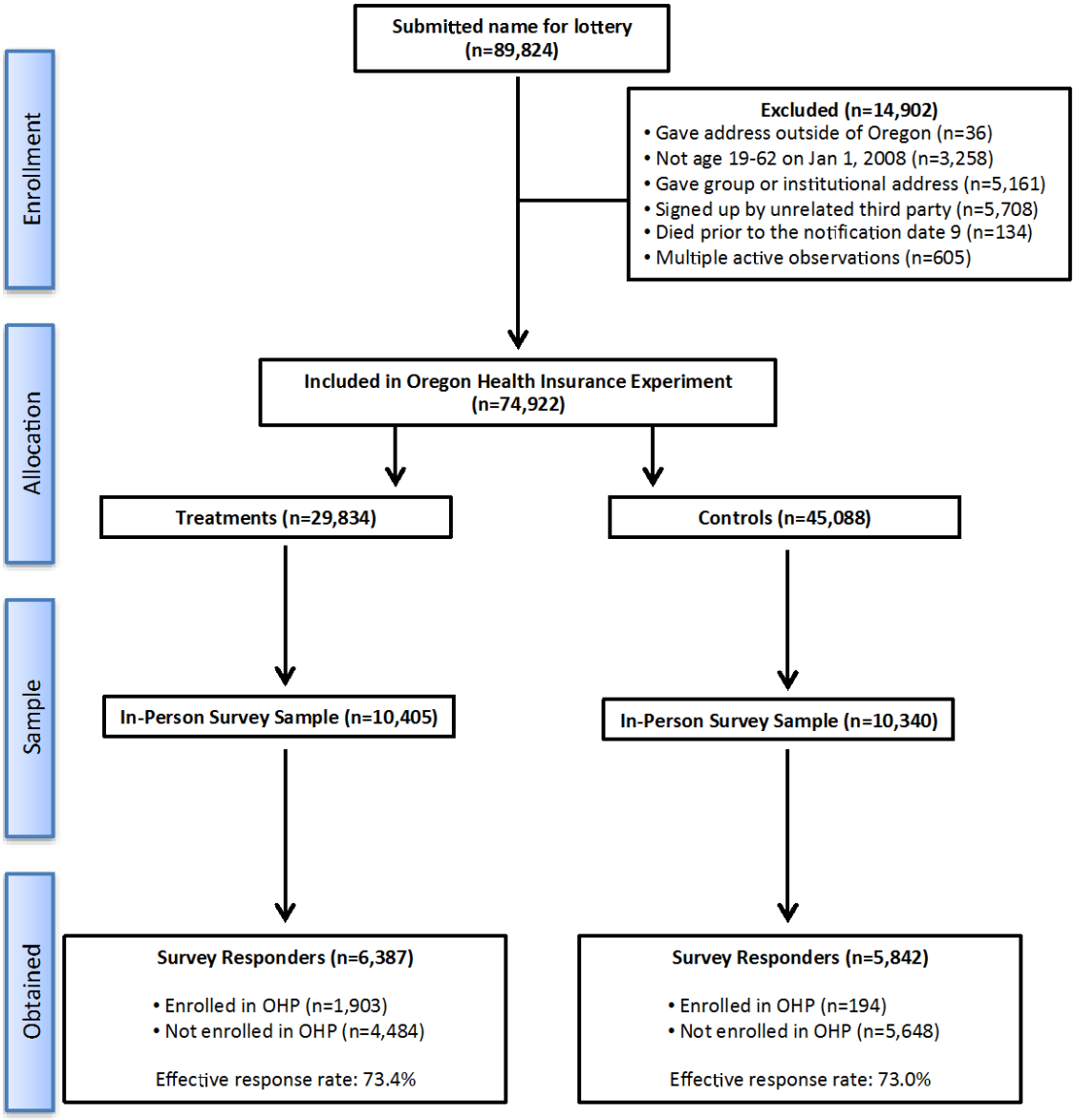
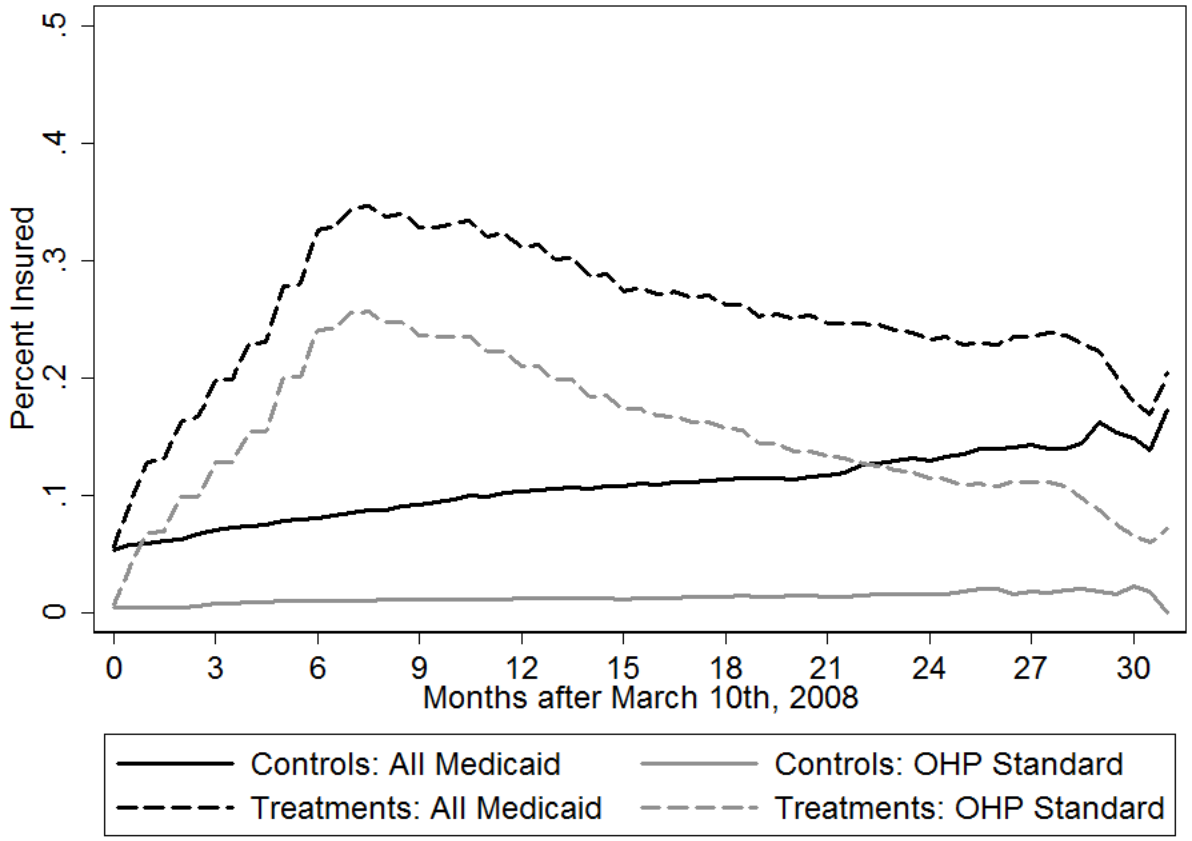


Figure S2: Enrollment in OHP Standard and all Medicaid



Notes: Figure shows the weighted percent with public insurance coverage over time. Weighted percent with insurance is shown separately for treatments and controls, and both all Medicaid coverage and OHP Standard coverage percentages are given. Time is measured in months from March 10, 2008; percent enrolled is observed twice a month. Individuals are censored following interview date and no longer contribute to the weighted percent. This mimics how we define the study period for each individual (from March 10, 2008 to interview date). The numbers closest to the end of the time period (28 months or more from March 10, 2008) are thus based on small numbers for respondents and not estimated precisely. Sample consists of survey respondents (N=12,229)

Table S1: Mean Values and Absolute Change in Clinical Measures and Health Outcomes

	Mean Value in Control Group	Effect of Lottery Selection (95% CI)	Effect of Medicaid Coverage (95% CI)	P Value
	(1)	(2)	(3)	(4)
Blood pressure				
Systolic blood pressure (mmHg)	119.3 (16.9)	-0.13 (-0.72 to 0.47)	-0.52 (-2.97 to 1.93)	0.68
Diastolic blood pressure (mmHg)	76.0 (12.1)	-0.19 (-0.64 to 0.25)	-0.81 (-2.65 to 1.04)	0.39
Elevated blood pressure (%)[1]	16.3	-0.32 (-1.73 to 1.09)	-1.33 (-7.16 to 4.49)	0.65
^Hypertension diagnosis post-lottery (%)[5]	5.6	0.42 (-0.46 to 1.30)	1.76 (-1.89 to 5.40)	0.34
^Current medication for hypertension (%)[6]	13.9	0.16 (-1.09 to 1.41)	0.66 (-4.48 to 5.80)	0.80
Cholesterol				
Total cholesterol (mg/dL)	204.1 (34.0)	0.53 (-0.83 to 1.89)	2.20 (-3.44 to 7.84)	0.45
High total cholesterol (%)[2]	14.1	-0.59 (-1.87 to 0.70)	-2.43 (-7.75 to 2.89)	0.37
HDL cholesterol (mg/dL)	47.6 (13.1)	0.20 (-0.32 to 0.72)	0.83 (-1.31 to 2.98)	0.45
Low HDL cholesterol (%)[2]	28.0	-0.68 (-2.48 to 1.12)	-2.82 (-10.28 to 4.64)	0.46
^High cholesterol diagnosis post-lottery (%)[5]	6.1	0.58 (-0.37 to 1.52)	2.39 (-1.52 to 6.29)	0.23
^Current medication for high cholesterol (%)[6]	8.5	0.92 (-0.18 to 2.02)	3.80 (-0.75 to 8.35)	0.10
Glycated hemoglobin				
Glycated hemoglobin level (% glycated)	5.3 (0.6)	0.0029 (-0.022 to 0.028)	0.01 (-0.09 to 0.11)	0.82
Glycated hemoglobin level ≥ 6.5% (%)[3]	5.1	-0.22 (-1.07 to 0.62)	-0.93 (-4.44 to 2.59)	0.61
^Diabetes diagnosis post-lottery (%)[5]	1.1	0.92 (0.47 to 1.38)	3.83 (1.93 to 5.73)	<.001
^Current medication for diabetes (%)[6]	6.4	1.31 (0.34 to 2.29)	5.43 (1.39 to 9.48)	0.008
Depression measures				
Positive depression screen (%)[4]	30.0	-2.21 (-4.02 to -0.40)	-9.15 (-16.70 to -1.60)	0.018
^Depression diagnosis post-lottery (%)[5]	4.8	0.91 (0.034 to 1.79)	3.81 (0.15 to 7.46)	0.041
^Current medication for depression (%)[6]	16.8	1.33 (-0.12 to 2.77)	5.49 (-0.46 to 11.45)	0.071
Predicted Cardiovascular Risk				
Framingham Risk Score (%)[7]	8.2 (7.5)	-0.05 (-0.39 to 0.28)	-0.21 (-1.56 to 1.15)	0.76
Limit to "high risk" diagnoses (%)[8]	11.6 (8.3)	0.40 (-0.27 to 1.07)	1.63 (-1.11 to 4.37)	0.24
Limit to aged 50-64 (%)	13.9 (8.2)	-0.10 (-0.71 to 0.51)	-0.37 (-2.64 to 1.90)	0.75

^This analysis was not pre-specified.

Notes: Column 1 reports the weighted mean of the dependent variable in the control sample of survey respondents and standard deviation (in parenthesis) for continuous outcomes. Column 2 reports intent-to-treat estimates (95% confidence interval in parenthesis), which compare the average outcome for all individuals selected in the lottery to the average outcome for all control individuals, as calculated by ordinary least squares regression. Column 3 reports the local-average-treatment-effect for insurance coverage as estimated by instrumental variable regression. Column 4 reports the per-comparison p-value, which is the same for the effect of lottery selection and the effect of insurance. All regressions include indicators for number of household members on the lottery list, and all standard errors are clustered on household. For the blood pressure measures, all regressions also include controls for age (in decile bins) and sex. All analyses are weighted using survey weights. Sample size is 12229, except for the Framingham risk scores which have sample sizes of 9525, 3099, and 3372 respectively.

Table S1 notes, continued.

[1]Elevated blood pressure is defined as having a systolic BP measure of 140 mmHg or above and a diastolic BP measure of 90 mmHg or above.

[2]High total cholesterol is defined as having a total cholesterol level of 240 mg/dL or higher. Low HDL cholesterol is defined as having a HDL level lower than 40 mg/dL. We do not have a separate measurement of LDL cholesterol.

[3]Glycated Hemoglobin A1c measure of 6.5% or higher is used as a criterion for diabetes.

[4]Positive depression screen is defined as reporting a PHQ-8 score of 10 or higher. The PHQ-8 ranges from 0-24.

[5]A participant is considered as being diagnosed for a certain condition post-lottery if they indicate a first diagnosis after March 2008 (the start of the lottery).

[6]A participant is considered as being medicated for a certain condition if one or more of the medications recorded during the interview are classified as relevant for that condition.

[7]The Framingham risk score was used to predict the 10-year cardiovascular risk. Risk scores were calculated separately for men and women on the basis of the following variables: age, total cholesterol and HDL cholesterol levels, measured blood pressure and use or nonuse of medication for high blood pressure, current smoking status, and status with respect to a glycated hemoglobin level ≥6.5%. Framingham risk scores, which are calculated for persons 30 years of age or older, range from 0.99 to 30%.

[8]A high-risk diagnosis was defined as a diagnosis of diabetes, hypertension, hypercholesterolemia, myocardial infarction, or congestive heart failure before the lottery (i.e., before March 2008).

Table S2: Mean Values and Absulute Change in Health-Related Quality of Life and Happiness

	Mean Value in Control Group	Effect of Lottery Selection (95% CI)	Effect of Medicaid Coverage (95% CI)	P Value
	(1)	(2)	(3)	(4)
Health-related quality of life				
Health same or better (%)	80.4	1.89 (0.36 to 3.43)	7.84 (1.45 to 14.23)	0.016
SF-8 mental component score[1]	44.4 (11.4)	0.47 (0.0097 to 0.93)	1.95 (0.028 to 3.88)	0.047
SF-8 physical component score[1]	45.5 (10.5)	0.29 (-0.13 to 0.71)	1.2 (-0.54 to 2.93)	0.18
No or very mild pain (%)	56.4	0.28 (-1.68 to 2.23)	1.16 (-6.94 to 9.26)	0.78
Happiness				
Very or pretty happy (%)	74.9	0.29 (-1.41 to 1.98)	1.18(-5.85 to 8.21)	0.74

Notes: Column 1 reports the weighted mean of the dependent variable in the control sample (standard deviations for continuous outcomes are in parentheses). Column 2 reports intent-to-treat estimates (with the 95% confidence interval in parentheses), which compare the average outcome for all individuals selected in the lottery to the average outcome for all control individuals, as calculated by ordinary least squares regression. Column 3 reports the local-average-treatment-effect for insurance coverage as estimated by instrumental variable regression. Column 4 reports the per-comparison p-value, which is the same for the effect of lottery selection and the effect of insurance. All regressions include indicators for number of household members on the lottery list, and all standard errors are clustered on household. All analyses are weighted using survey weights. Sample size is 12,229.

[1] Higher SF-8 mental or physical subscale score indicates higher self-reported quality of life. The scale is normalized to yield means of 50 and standard deviations of 10 in the general U.S. population; the range is 0 to 100.

Table S3: Mean Values and Absolute Change in Financial Hardship

	Mean Value in Control Group	Effect of Lottery Selection (95% CI)	Effect of Medicaid Coverage (95% CI)	P Value
	(1)	(2)	(3)	(4)
Any out-of-pocket spending (%)	58.8	-3.70 (-5.64 to -1.75)	-15.30 (-23.28 to -7.32)	<0.001
Amount of out-of-pocket spending	552.8 (1219.5)	-52.13 (-99.14 to -5.12)	-215.35 (-408.75 to -21.95)	0.029
Catastrophic expenditures (%)	5.5	-1.08 (-1.98 to -0.17)	-4.48 (-8.26 to -0.69)	0.020
Any medical debt (%)	56.8	-3.21 (-5.23 to -1.20)	-13.28 (-21.59 to -4.96)	0.002
Borrowed or skipped bills (%)	24.4	-3.44 (-5.09 to -1.79)	-14.22 (-21.02 to -7.43)	<0.001

Notes: Column 1 reports the weighted mean of the dependent variable in the control sample and standard deviation (in parentheses) for continuous outcomes. Column 2 reports intent-to-treat estimates (with the 95% confidence interval in parentheses), which compare the average outcome for all individuals selected in the lottery to the average outcome for all control individuals, as calculated by ordinary least squares regression. Column 3 reports the local-average-treatment-effect for insurance coverage as estimated by instrumental variable regression. Column 4 reports the per-comparison p-value, which is the same for the effect of lottery selection and the effect of insurance. All regressions include indicators for household size, and all standard errors are clustered on household. All analyses are weighted using survey weights. Sample is all survey respondents (N=12, 229).

Table S4: Mean Values and Absolute Change in Utilization, Prevention, Access and Quality, and Smoking and Obesity

	Mean Value in Control Group	Effect of Lottery Selection (95% CI)	Effect of Medicaid Coverage (95% CI)	P Value
	(1)	(2)	(3)	(4)
Utilization (number of visits or medications)				
Prescription drugs (currently taking)	1.8 (2.8)	0.16 (0.05 to 0.27)	0.66 (0.21 to 1.11)	0.004
Office Visits (last 12 months)	5.5 (11.6)	0.65 (0.21 to 1.08)	2.70 (0.91 to 4.49)	0.003
Outpatient Surgery (last 12 months)	0.1 (0.4)	0.01 (-0.01 to 0.02)	0.03 (-0.03 to 0.09)	0.28
ED visits (last 12 month)	1.0 (2.0)	0.02 (-0.06 to 0.10)	0.09 (-0.23 to 0.42)	0.57
Hospital admissions (last 12 months)	0.2 (0.6)	0.02 (-0.01 to 0.04)	0.07 (-0.03 to 0.17)	0.17
Spending Estimate				
Annual spending ($)[1]	3,257.28	282.75	1171.63	0.018
Standard error for spending		121.13	496.06	
Prevention (last 12 months)				
Cholesterol-level screening (%)	27.2	3.55 (1.71 to 5.38)	14.57 (7.09 to 22.04)	<.001
Fecal occult-blood test (age >=50) (%)	19.1	0.34 (-2.55 to 3.22)	1.26 (-9.44 to 11.96)	0.82
Colonoscopy (age >=50) (%)	10.4	1.12 (-1.16 to 3.40)	4.19 (-4.25 to 12.62)	0.33
Flu shot (age >=50) (%)	35.5	-1.54 (-5.14 to 2.07)	-5.74 (-19.31 to 7.83)	0.41
Papanicolaou smear (women) (%)	44.9	3.21 (0.56 to 5.86)	14.44 (2.64 to 26.24)	0.016
Mammogram (women >=50) (%)	28.9	7.75 (3.12 to 12.37)	29.67 (11.96 to 47.37)	0.001
PSA (men >=50) (%)	21.4	5.02 (0.25 to 9.79)	19.18 (1.14 to 37.21)	0.037
Access and Quality				
Have a usual place of care (%)	46.1	5.74 (3.73 to 7.76)	23.75 (15.44 to 32.06)	<.001
Received all needed care (%)	61.0	2.76 (0.88 to 4.65)	11.43 (3.62 to 19.24)	0.004
High quality, if received care (%)	78.4	2.46 (0.69 to 4.23)	9.85 (2.71 to 17.00)	0.007
Smoking and Obesity				
Currently Smoking (%)	42.8	1.35 (-0.61 to 3.31)	5.58 (-2.54 to 13.70)	0.18
Obese (%)	41.5	0.09 (-1.90 to 2.09)	0.39 (-7.89 to 8.67)	0.93

Notes: For the prevention measures, sample is all survey respondents (N=12,229), survey respondents at least 50 years of age (N=3374), female survey respondents (N=6915), female survey respondents at least 50 years of age (N=1864) or male survey respondents at least 50 years of age (N=1509), as indicated in the table. For all other measures, sample is all survey respondents (N=12, 229) except for quality of care which is only defined on those who received care in the last 12 months (N= 9,694)."Received all needed care" is defined over the last 12 months. Column 1 reports the weighted mean of the dependent variable in the control sample and standard deviation (in parentheses) for continuous outcomes. Column 2 reports intent-to-treat estimates (with the 95% confidence interval in parentheses), which compare the average outcome for all individuals selected in the lottery to the average outcome for all control individuals, as calculated by ordinary least squares regression. Column 3 reports the local-average-treatment-effect for insurance coverage as estimated by instrumental variable regression. Column 4 reports the per-comparison p-value, which is the same for the effect of lottery selection and the effect of insurance. All regressions include indicators for household size, and all standard errors are clustered on household. All analyses are weighted using survey weights.

[1] Annual spending is constructed based on the above measures of Rx drugs, office visits, ED visits, and hospital visits, combined with estimates of the cost of each. See appendix for further details.

^Table S5. Characteristics of the compliers

	Sample Mean (1)	Complier Mean (2)
Female (%)	56.4	52.5
Age[1]		
19-34 yrs (%)	35.6	29.8
35-49 yrs (%)	36.7	40.2
50-64 yrs (%)	27.7	30.8
Race or ethnic group[2]		
Non-Hispanic White (%)	68.6	76.2
Non-Hispanic Black (%)	10.3	11.0
Non-Hispanic other race (%)	14.9	11.2
Hispanic (%)	17.7	11.2
Interview conducted in English (%)	87.7	94.7

^This analysis was not pre-specified.

Notes: The first column reports the weighted mean of these variables for the sample; the second column reports the weighted mean for the compliers calculated following Angrist and Pischke (2009). Sample size is 12,229.

[1] The age listed was that calculated at the time of the inperson interview. Our study sample is restricted to those between 19 and 64 during the study period.

[2] The non-Hispanic racial categories (white, black and other) are not mutually exclusive, respondents could endorse as many races as desired.

Table S6: Summary of Weights

	Mean	Std Dev	Min	5th %tile	25th %tile	50th %tile	75th %tile	95th %tile	Max	N
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Sampling base	0.998	1.283	0.000	0.000	0.000	1.000	1.159	3.458	52.211	20745
Recruitment base	1.511	1.310	0.671	0.978	1.000	1.150	1.378	3.491	52.211	13707
Survey respondents	1.240	0.570	0.681	0.960	1.000	1.068	1.213	2.076	13.634	12229
Control group respondents	1.308	0.663	0.681	0.950	1.000	1.140	1.307	2.361	13.634	5842
Lottery winning respondents	1.178	0.461	0.862	1.000	1.000	1.003	1.152	1.815	10.872	6387

Notes: Zero weights are the result of being dropped from active follow-up. The recruitment base is the sampling base limited to those with non-zero weights. Respondents all have non-zero weights.

Table S7. Characteristics of the study population

	Mean Value in Control Group (1)	Difference between control group and lottery winners (95% CI) (2)	P Value (3)
Panel A: Response Rates			
Responded to survey	73.0 (44.4)	0.28 (-2.77 to 3.34)	0.86
Had Anthropometric Measures	97.7 (15.0)	0.037 (-0.54 to 0.61)	0.90
Had at least one DBS measure	99.5 (6.7)	0.030 (-0.24 to 0.30)	0.83
Had all DBS measures	99.1 (9.7)	0.21 (-0.25 to 0.68)	0.37
Had medication data or not needed	98.4 (12.6)	0.027 (-0.46 to 0.51)	0.91
Pooled F-Stats			0.24
p-value			0.91
N			12229
Panel B: Responder and interview characteristics, limited to responders			
Age	40.7 (11.7)	0.20 (-0.28 to 0.69)	0.41
Female	56.9 (49.6)	-0.46 (-2.16 to 1.24)	0.60
Black	10.5 (30.7)	0.14 (-1.06 to 1.34)	0.82
Other race	14.8 (35.5)	0.034 (-1.53 to 1.60)	0.97
Hispanic	17.2 (37.8)	-0.19 (-1.84 to 1.45)	0.82
English as preferred language	90.7 (29.0)	-0.17 (-1.52 to 1.18)	0.81
Signed self up	89.5 (30.6)	0.11 (-0.28 to 0.51)	0.57
Signed up first day of list	9.6 (29.5)	0.64 (-0.69 to 1.98)	0.34
Gave phone number	87.6 (32.9)	-0.029 (-1.51 to 1.45)	0.97
Address a PO Box	3.0 (17.0)	0.46 (-0.31 to 1.22)	0.24
Median hh income of zip code	44098.0 (9563.1)	-12.42 (-425.92 to 401.09)	0.95
Interview date (difference in days)	18375.8 (103.0)	-1.40 (-5.89 to 3.08)	0.54
Response time (days)	43.7 (52.3)	1.22 (-1.95 to 4.39)	0.45
Winter interview	20.2 (40.1)	0.11 (-1.29 to 1.52)	0.88
Spring interview	28.0 (44.9)	-1.14 (-2.96 to 0.67)	0.22
Summer interview	19.4 (39.5)	1.61 (-0.20 to 3.41)	0.08
Weekend interview	11.0 (31.3)	-0.37 (-1.62 to 0.88)	0.56
In-home interview	9.2 (28.9)	-0.55 (-1.85 to 0.75)	0.41
East side clinic interview	39.5 (48.9)	1.32 (-0.77 to 3.41)	0.21
South side clinic interview	20.9 (40.6)	0.31 (-1.51 to 2.12)	0.74
Spanish language instrument	9.2 (28.9)	0.097 (-1.17 to 1.37)	0.88
Interviewed with interpreter	2.8 (16.5)	-0.38 (-1.33 to 0.57)	0.43
Pooled F-Stats			0.7
p-value			0.84
N			12229

Panel C: Measurement variables, limited to responders	
Interviewer (pooled F-stat)	1.21
p-value	0.15
N	12229
Scale(pooled F-stat)	1.11
p-value	0.29
N	12229
Stadiometer (pooled F-stat)	1.15
p-value	0.23
N	12229
Sphygmomanometer (pooled F)	0.85
p-value	0.73
N	12229

Notes: Panel A shows the response rate to the survey.

Panel B variables are pre-randomization "demographics" taken from the lottery list (from January and February 2008) and characteristics of the interview itself. Age, sex, race and ethnicity are taken from information reported in the interview; respondents were allowed to report multiple races.
The next set of variables is from the lottery list. "English as preferred language" indicates whether you did not check a box requesting materials in a language other than English. "Signed up self?" is an indicator for whether you signed yourself up (as opposed to a household member including your name when they signed up). "Signed upon first day of list?" is an indicator variable for whether you signed up the first day the list was open. "Gave phone number" is an indicator variable for whether you provided a phone number when you signed up.

The remaining variables in Panel B are characteristics of the interview itself. "Interview date" is the time when the interview was conducted. The unit for the mean and the standard deviation are in days. "Response time" indicates the number of days between when the study participant was first released to an interviewer for recruitment and when the interview took place. "Response time" is missing for 885 survey responders because we could not accurately identify their release date. "Winter interview", "Spring interview", "Summer interview" are indicators for whether the interview was conducted in the corresponding season; the omitted category is "Fall interview." "Spring" is defined as March, April, and May, and all other seasons are defined accordingly. "Weekend interview" indicates if the interview took place on a weekend. "In-home interview", "East side clinic interview", "South side clinic interview" are indicators for whether the interview was conducted in the corresponding location; the omitted category is "West side clinic interview." "Spanish language instrument" is an indicator for whether the survey instrument is in Spanish. "Interviewed with interpreter" is an indicator for whether an interpreter was present during the interview.

All analysis is weighted using survey weights. The first column reports the mean of these variables for the control sample and standard deviation (in parenthesis) for continuous variables. Column (2) reports estimated differences between treatments and controls in the survey responders for the outcome shown in the left hand (except in Panel A where the whole survey sample is used). Specifically it reports the coefficient from a regression of the outcome on an indicator variable for having been selected by the lottery (treated). All regressions include indicators for household size, and all standard errors are clustered on household. We report the coefficient, 95% confidence interval, and per comparison p-value. The last row of panel B reports the pooled F-stat from estimating for all the variables in that panel jointly.

Panel C reports global tests for if there is any evidence of sorting across interviewers or equipment used. The scales are equipment used for measuring weight, the stadiometers are equipment used for measuring height and the sphygmomanometers are equipment used for measuring blood pressure. There are 49 interviewers, and we have interviewer information for all but 5 observations. We could identify the scale, stadiometers, sphygmomanometers used for 12202, 12211, 12189 observations, respectively. A few equipments are only used once or twice. To increase power, for each category, we grouped all equipments used for 10 or fewer observations into an "other" category. After this grouping, there are 44 different scale groups, 44 stadiometer groups, and 43 sphygmomanometer groups in our analysis. The global test for sorting across interviewers (scales, stadiometers, sphygmomanometers) calculated by estimating regression with each of the 49 interviewers (44 scales, 44 stadiometers, 43 sphygmomanometers) as the outcome, then testing whether the 49 coefficients on the indicator for having been selcted in the lottery (treated) are equal.

Table S8. Balance of pre-randomization diagnoses

	Mean Value in Control Group	Difference between control group and lottery winners (95% CI)	P Value
	(1)	(2)	(3)
Asthma	19.9 (39.9)	-0.68 (-2.26 to 0.89)	0.40
Diabetes	7.2 (25.9)	-0.16 (-1.14 to 0.82)	0.76
Hypertension	18.1 (38.5)	0.21 (-1.29 to 1.70)	0.79
High cholesterol	12.7 (33.3)	-0.15 (-1.46 to 1.17)	0.83
Heart attack	2.0 (13.9)	-0.12 (-0.64 to 0.40)	0.66
Congestive heart failure	1.0 (9.8)	0.16 (-0.21 to 0.53)	0.39
Emphysema/COPD	2.3 (15.0)	0.018 (-0.58 to 0.61)	0.95
Failing kidneys	1.8 (13.3)	-0.049 (-0.52 to 0.42)	0.84
Cancer	4.3 (20.2)	0.15 (-0.65 to 0.95)	0.72
Depression/anxiety	35.0 (47.7)	-0.80 (-2.67 to 1.07)	0.40
Pooled F-stat			0.30
p-value			0.98
N			12229
Standardized treatment effect			-0.0026
p-value			0.76
N			12229
Composite of diabetes, hypertension, high cholesterol, heart attack, or congestive heart failure	27.3 (44.6)	-0.26 (-2.02 to 1.49)	0.77

Notes: All analysis weighted using survey weights. The first column reports the mean of these variables for the control sample, with standard deviations in parenthesis. Column (2) reports estimated differences between treatments and controls in the survey responders for the outcome shown in the left hand. Specifically it reports the coefficient from a regression of the outcome on an indicator variable for having been selected by the lottery (treated). All regressions include household fixed effects. All standard errors are clustered on household. We report the coefficient, 95% confidence interval, and per comparison p-value. We report the pooled F-stat and the standardized treatment effect from estimating for all the variables (except the composite measure) jointly.

We asked individuals about whether they were ever diagnosed with the following conditions: asthma, diabetes, hypertension, high cholesterol, heart attack, congestive heart failure, emphysema/COPD, kidney failure, cancer, and depression, and when they were diagnosed. If an individual was interviewed in 2010 and answered "more than 3 years ago" to the question "when were you first diagnosed", or if the individual was interviewed in 2009 and answered "more than 2 years ago" to the question "when were you first diagnosed", we knew that the diagnosis was before the lottery. In other cases we asked explicitly about the month and year of diagnosis to determine whether the diagnosis was before or after the lottery. For each of these conditions, we consider the individual to have a pre-randomization diagnosis of asthma, diabetes, hypertension, high cholesterol, heart attack, congestive heart failure, emphysema/COPD, kidney failure, cancer, or depression if we could identify that the diagnosis of the specific condition happened before March 10, 2008 (the earliest possible selection date for the lottery).

We consider an individual high risk if there was a pre-randomization diagnosis of diabetes, high blood pressure, high cholesterol, heart attack, or congestive heart failure (although there are likely other high risk individuals with such conditions who were never diagnosed). The last row reports the results for this composite measure,which is the one we use to select the sub-sample for the middle panel of Table S9.

Table S9. First Stage

	Mean Value in Control Group (1)	Difference between control group and lottery winners (95% CI) (2)	P Value (3)
(1) Ever on Medicaid during study period (%)	18.5	24.14 (22.37 to 25.92)	<0.001
(2) Ever on OHP Standard during study period(%)	3.3	26.49 (25.11 to 27.87)	<0.001
(3) # of Months on Medicaid during study period	2.6	4.16 (3.84 to 4.49)	<0.001
(4) Currently on Medicaid (%)	13.3	11.35 (9.81 to 12.89)	<0.001
(5) Currently have any insurance (self-report, %)	35.8	11.13 (9.13 to 13.13)	<0.001
(6) Currently have Medicaid (self-report, %)	12.8	12.32 (10.82 to 13.83)	<0.001
(7) Currently have private insurance (self-report, %)	14.7	-0.40 (-1.82 to 1.03)	0.58

Notes: The first column reports the weighted control mean for the measure of "INSURANCE" defined in the left-hand column; The second column reports the effect on insurance coverage, which compares the average of the insurance measure for all individuals selected in the lottery to the average of the insurance measure for all control individuals, as calculated by ordinary least squares regression. All regressions include household size fixed effects and are weighted using survey weights. All standard errors are clustered on the household.
The insurance measures are taken from the Medicaid enrollment administrative data except for those labeled "self-report" (rows 5-7) 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.
The study period is defined as running from March 10, 2008 to the date of interview; variables defined as "ever" (rows 1-3) cover this entire period; variables defined as "currently" are current for the interview date.
Sample consists of survey responders (N = 12,229).

Table S10: Summary of Variables

	Time frame of question	Survey question name(s)*	Non-missing data (N)	Non-missing data (%)
Clinical measures of health				
Systolic blood pressure	Current	physical measures	12188	99.7
Diastolic blood pressure	Current	physical measures	12188	99.7
Total cholesterol	Current	physical measures	12174	99.6
HDL cholesterol	Current	physical measures	12172	99.5
Glycated hemoglobin	Current	physical measures	12140	99.3
Positive depression screen	Current	physical measures	12161	99.4
Diagnoses and medications				
Post-lottery dx of hypertension	Last 3 years*	hbp_dx/when/year/month_inp	11945	97.7
Post-lottery dx of high chol	Last 3 years*	chl_dx/when/year/month_inp	11894	97.3
Post-lottery dx of diabetes	Last 3 years*	dia_dx/when/year/month_inp	12186	99.6
Post-lottery dx depression	Last 3 years*	dep_dx/when/year/month_inp	12095	98.9
Current medication for hypertension	Current	medication survey	12229	100.0
Current medication for high chol	Current	medication survey	12229	100.0
Current medication for diabetes	Current	medication survey	12229	100.0
Current medication for depression	Current	medication survey	12229	100.0
Framingham risk score				
Framingham risk score (age≥30)	Current	from the above + physical measures	9421	98.9
Health-related quality of life				
Health change not worse	Last 12 months	health_change_inp	12226	100.0
SF-8 physical component score	Last 4 weeks	sf1_inp to sf8_inp	12204	99.8
SF-8 mental component score	Last 4 weeks	sf1_inp to sf8_inp	12204	99.8
No or mild pain	Last 4 weeks	sf4_inp	12225	100.0
Happiness				
Very or pretty happy	Current	happy_inp	12206	99.8
Finances				
Any out-of-pocket spending	Last 12 months	doc/surg/ed/hosp/other_cost_inp plus probes	12194	99.7
Amount of out-of-pocket spending	Last 12 months		12145	99.3
Had catastrophic expenditures	Last 12 months	from the above + hh_income_inp	11795	96.5
Any medical debt	Current	owe_inp	12108	99.0
Borrow money or skipped bills	Last 12 months	borrow_inp	12212	99.9
Health care utilization				
Number of prescription drugs taking	Current	rx_any_inp + medication survey responses	11912	97.4
Number of office visits	Last 12 months	doc_use_inp and doc_use_probe_inp	12158	99.4
Number of outpatient surg	Last 12 months	surg_use_inp and surg_use_probe_inp	12188	99.7
Number of ED visits	Last 12 months	ed_use_inp and ed_use_probe_inp	12175	99.6
Number of hospital visits	Last 12 months	hosp_use_inp and hosp_use_probe_inp	12175	99.6

Table S10, Continued

Preventive care in last 12 months (all use care_any_inp and probe)				
Has cholesterol-level screening	Last 12 months	chl_dx_inp, chl_test_when_inp	11382	93.1
Had a fecal occult-blood test (age≥50)	Last 12 months	fobt_ever_inp	3358	99.5
Had a colonoscopy (age≥50)	Last 12 months	col_ever_inp	3,361	99.6
Had a flu shot (age≥50)	Last 12 months	did_flu_inp	3364	99.7
Had a papanicolaou smear (women)	Last 12 months	pap_inp	6673	96.5
Had mammogram (women, age≥50)	Last 12 months	mam_inp	1858	99.7
Had a PSA test (men, age≥50)	Last 12 months	psa_inp	1381	91.5
Access and quality				
Have usual place of clinic-based care	Current	usual_place_inp, usual_place_where_inp	12219	99.9
Got all needed medical care	Last 12 months	got_care_phs_inp, needed_care_phys_inp	12216	99.9
Got all needed mental health care	Last 12 months	got_care_ment_inp, needed_care_med_inp	12192	99.7
Got all needed drugs	Last 12 months	rx_delay_inp, rx_inp	12215	99.9
Quality of care (cond. on any)	Last 12 months	satisfaction_inp	12137	99.1
Health behavior				
Obese	Current	physical measures	12175	99.6
Current smoking	Current	smk_now_inp	12225	100.0

* We used the date of interview and the responder's survey answer to the timing of diagnoses to determine whether a diagnosis happened before or after the lottery

Table S11: Distribution of variables (control sample only)

Panel A: Clinical Measures of Health

	Mean	SD	5th %tile	25 %tile	Median	75th %tile	95th %tile
Systolic Blood Pressure	119.28	16.85	97	107	117	128	149
Diastolic Blood Pressure	76.01	12.14	58	67	75	83	98
Total Cholesterol	204.06	34.00	154.38	180.46	201.00	224.86	263.71
HDL Cholesterol	47.56	13.15	28.08	38.58	46.83	55.08	71.58
Glycated Hemoglobin	5.34	0.63	4.79	5.04	5.20	5.45	6.51
PHQ total score	701.57	601.10	0	200	500	1100	1600

Note: All numbers are weighted

Panel B: Health Status

	N	%
General health, last 12 mo.		
1: Very poor	134.73	2.31
2: Poor	692.40	11.86
3: Fair	1533.75	26.27
4: Good	2100.89	35.99
5: Very good	1024.71	17.55
6: Excellent	351.52	6.02
Health status compared to 12 mo. ago		
0: Better	1461.07	25.02
1: Worse	1144.10	19.59
2: About the same	3233.83	55.38

Note: All numbers are weighted

Panel C: Self Reported Health

	Mean	SD	5th %tile	25 %tile	Median	75th %tile	95th %tile
SF-8 physical component score	45.49	10.50	25.65	37.83	47.71	54.16	57.32
SF-8 mental component score	44.39	11.38	22.62	36.48	46.76	53.31	57.67

Note: All numbers are weighted

Panel D: Health Care Use

	Percent reporting any	Mean	SD	Median	75th %tile	95th %tile	Cutpoint for truncation	% of data truncated
Doctor office visits	64.55	8.61	13.49	4	10	30	164	0.5
Outpatient surgery visits	7.81	1.25	0.58	1	1	2	4	0.2
ED visits	40.23	2.48	2.5	2	3	7	20	0.1
Inpatient hospital visit	12.66	1.5	1.04	1	2	4	6	0.3
Number of Rx drugs	53.89	3.48	3.04	2	5	10	n/a	n/a

Note: In Panels D and E, the mean, standard deviation, median, 75th and 95th percentile values reflect non-zero observations only, after truncating at 2*99% based on the unweighted distribution. "Number of Rx drugs" is not truncated. Percent reporting any use, cutpoint for censoring and percent of data censored reflect all valid non-missing data, including observations with zero values. The value for “percent reporting any use” for “Number of Rx drugs” includes 157 control respondents who reported taking Rx drugs but for whom we could not accurately count the number of Rx drugs. All numbers in table except “Cutpoint for truncation” are weighted.

Panel E: Financial Strain

	Percent reporting any	Mean	SD	Median	75th %tile	95th %tile	Cutpoint for truncation	% of data truncated
Total out of pocket expense	58.8	942.18	1472.37	440	1075	3661	15200	0.3

Note: In Panels D and E, the mean, standard deviation, median, 75th and 95th percentile values reflect non-zero observations only, after truncating at 2*99% based on the unweighted distribution. "Total medical expense" and "total other expense" are not truncated. Percent reporting any expenses, cutpoint for censoring and percent of data censored reflect all valid non-missing data, including observations with zero values. Missing values are largely due to individuals answering "don't know" or "prefer not to answer" to the survey question. All numbers except “Cutpoint for truncation” are weighted.

Table S12: Comparison of physical measures to a national sample

	Mean Value in Control Group	NHANES Adults <200% FPL	NHANES Uninsured Adults <200% FPL	NHANES Adults <100% FPL	NHANES Uninsured Adults <100% FPL
	(1)	(2)	(3)	(4)	(5)
Blood pressure					
Systolic blood pressure (mmHg)	119.3 (16.9)	119.1 (16.5)	119.1 (16.3)	118.2 (16.1)	118.8 (16.2)
Diastolic blood pressure (mmHg)	76.0 (12.1)	71.0 (12.5)	70.7 (12.7)	70.4 (12.7)	70.3 (13.0)
Elevated blood pressure (%)[1]	16.3	13.0	12.0	11.8	11.3
Cholesterol					
Total cholesterol (mg/dL)	204.1 (34.0)	196.3 (44.1)	195.9 (41.6)	194.8 (46.0)	194.0 (41.6)
High total cholesterol (%)[2]	14.1	14.6	14.1	14.2	12.5
HDL cholesterol (mg/dL)	47.6 (13.1)	50.5 (15.4)	50.1 (15.5)	49.8 (15.0)	49.9 (15.2)
Low HDL cholesterol (%)[2]	28.0	23.8	24.5	25.6	25.0
Glycated hemoglobin					
Glycated hemoglobin level (% glycated)	5.3 (0.6)	5.5 (1.0)	5.4 (1.0)	5.5 (1.1)	5.5 (1.1)
Glycated hemoglobin level ≥ 6.5% (%)[3]	5.1	6.5	5.0	6.8	5.2
Sample size	5933	8486	3761	3967	1813

Notes: Column 1 reports the weighted mean of the dependent variable in the control sample of survey respondents and standard deviation (in parenthesis) for continuous outcomes. Columns 2-4 report the weighted means of the variable in the 1999-2008 NHANES data, limited to subgroups by reported income and insurance status.

[1]Elevated blood pressure is defined as having a systolic BP measure of 140 mmHg or above and a diastolic BP measure of 90 mmHg or above

[2]High total cholesterol is defined as having a total CHL level of 240 mg/dL or higher. Low HDL cholesterol is defined as having a HDL level lower than 40 mg/dL.

[3]Glycated Hemoglobin A1c measure of 6.5% or higher is used as a criterion for diabetes.

Table S13: Classification of medications

Panel A: Distribution of Anti-hypertensives

Medication Name	Frequency	Percent	Cumulative
Lisinopril	950	38.28	38.28
Hydrochlorothiazide	531	21.39	59.67
Furosemide	151	6.08	65.75
Clonidine	137	5.52	71.27
Lisinopril-Hydrochlorothiazide	106	4.27	75.54
Cozaar	66	2.66	78.20
Spironolactone	64	2.58	80.78
Triamterene-Hydrochlorothiazid	47	1.89	82.68
Doxazosin	36	1.45	84.13
Enalapril Maleate	32	1.29	85.41
Losartan	24	0.97	86.38
Terazosin	23	0.93	87.31
Accupril	21	0.85	88.15
Prazosin	17	0.68	88.84
Benicar	16	0.64	89.48
Diovan	16	0.64	90.13
Lasix	16	0.64	90.77
Chlorthalidone	16	0.64	91.42
Hydralazine	15	0.60	92.02
Amlodipine-Benazepril	14	0.56	92.59
Quinapril	14	0.56	93.15
Accuretic	10	0.40	93.55
Aldactone	10	0.40	93.96
Cardura	10	0.40	94.36
Atenolol-Chlorthalidone	10	0.40	94.76
Other	130	5.2	100

Panel B: Distribution of Antihyperlipidemics

Medication Name	Frequency	Percent	Cumulative
Simvastatin	411	31.79	31.79
Lipitor	228	17.63	49.42
Lovastatin	214	16.55	65.97
Pravastatin	87	6.73	72.70
Gemfibrozil	69	5.34	78.04
Fish Oil	68	5.26	83.29
Crestor	32	2.47	85.77
Atorvastatin	23	1.78	87.55
Omega 3 Fish Oil	14	1.08	88.63
Niacin	13	1.01	89.64
Niaspan Extended-Release	12	0.93	90.56
Lovaza	10	0.77	91.34
Tricor	10	0.77	92.11
Zocor	10	0.77	92.88
Other	92	7.13	100

Panel C: Distribution of Diabetes Medicine

Medication Name	Frequency	Percent	Cumulative
Metformin	657	42.22	42.22
Glipizide	174	11.18	53.41
Lantus	140	9.00	62.40
Glyburide	96	6.17	68.57
Humalog	53	3.41	71.98
Novolog	39	2.51	74.49
Novolin R	35	2.25	76.74
Actos	34	2.19	78.92
Novolin 70/30	33	2.12	81.04
Novolin N	27	1.74	82.78
Glucotrol Xl	23	1.48	84.25
Glimepiride	20	1.29	85.54
Glucophage	15	0.96	86.50
Humulin N	15	0.96	87.47
Humulin R	15	0.96	88.43
Lantus Solostar	14	0.90	89.33
Glucotrol	11	0.71	90.04
Novolog Flexpen	11	0.71	90.75
Other	144	9.19	100

Panel D: Distribution of Antidepressants

Medication Name	Frequency	Percent	Cumulative
Trazodone	430	14.83	14.83
Citalopram	384	13.24	28.07
Fluoxetine	257	8.86	36.93
Amitriptyline	215	7.41	44.34
Sertraline	200	6.90	51.24
Cymbalta	196	6.76	58.00
Bupropion HCl	137	4.72	62.72
Paroxetine HCl	121	4.17	66.90
Lexapro	113	3.90	70.79
Effexor XR	96	3.31	74.10
Zoloft	95	3.28	77.38
Prozac	68	2.34	79.72
Celexa	63	2.17	81.90
Venlafaxine	52	1.79	83.69
Bupropion (bulk)	49	1.69	85.38
Nortriptyline	49	1.69	87.07
Wellbutrin	46	1.59	88.66
Wellbutrin SR	36	1.24	89.90
Mirtazapine	35	1.21	91.10
Doxepin	33	1.14	92.24
Paxil	32	1.10	93.34
Wellbutrin XL	25	0.86	94.21
Pristiq	23	0.79	95.00
Effexor	20	0.69	95.69
Budeprion XL	19	0.66	96.34
Budeprion SR	18	0.62	96.97
Fluoxetine HCL	17	0.59	97.55
Escitalopram	10	0.34	97.90
Other	61	2.10	100.00

Table S14a: Mean Values and Absolute Change in Clinical Measures and Health Outcomes: age 50-64

	Mean Value in Control Group (1)	Effect of Lottery Selection (95% CI) (2)	Effect of Medicaid Coverage (95% CI) (3)	P Value (4)
Blood pressure				
Systolic blood pressure (mmHg)	127.3 (19.4)	-0.67 (-2.15 to 0.81)	-2.50 (-7.99 to 3.00)	0.37
Diastolic blood pressure (mmHg)	79.6 (12.8)	-0.74 (-1.70 to 0.22)	-2.76 (-6.34 to 0.81)	0.13
Elevated blood pressure (%)	28.3	-2.05 (-5.45 to 1.34)	-7.66 (-20.32 to 4.99)	0.24
^Hypertension diagnosis post-lottery (%)	7.4	1.45 (-0.46 to 3.35)	5.47 (-1.73 to 12.67)	0.14
^Current medication for hypertension (%)	32.3	-1.20 (-4.61 to 2.21)	-4.50 (-17.33 to 8.33)	0.49
Cholesterol				
Total cholesterol (mg/dL)	207.1 (36.6)	0.50 (-2.13 to 3.13)	1.86 (-7.96 to 11.69)	0.71
High total cholesterol (%)	18.3	-1.32 (-3.98 to 1.34)	-4.93 (-14.87 to 5.00)	0.33
HDL cholesterol (mg/dL)	48.1 (14.1)	0.62 (-0.41 to 1.64)	2.31 (-1.55 to 6.16)	0.24
Low HDL cholesterol (%)	28.1	-1.61 (-4.88 to 1.66)	-6.01 (-18.28 to 6.25)	0.34
^High cholesterol diagnosis post-lottery (%)	10.3	1.42 (-0.87 to 3.71)	5.21 (-3.17 to 13.58)	0.22
^Current medication for high cholesterol (%)	20.0	1.50 (-1.48 to 4.48)	5.60 (-5.44 to 16.65)	0.32
Glycated hemoglobin				
Glycated hemoglobin level (% glycated)	5.5 (0.7)	-0.0052 (-0.059 to 0.049)	-0.019 (-0.22 to 0.18)	0.85
Glycated hemoglobin level ≥ 6.5% (%)	9.1	-0.55 (-2.66 to 1.57)	-2.04 (-9.97 to 5.88)	0.61
^Diabetes diagnosis post-lottery (%)	1.9	1.61 (0.45 to 2.76)	6.03 (1.64 to 10.41)	0.007
^Current medication for diabetes (%)	13	1.58 (-0.93 to 4.09)	5.90 (-3.44 to 15.25)	0.22
Depression measures				
Positive depression screen (%)	39.1	-5.54 (-9.08 to -2.00)	-20.81 (-34.50 to -7.13)	0.003
^Depression diagnosis post-lottery (%)	4.7	0.20 (-1.32 to 1.72)	0.77 (-4.99 to 6.53)	0.79
^Current medication for depression (%)	25.4	-0.59 (-3.75 to 2.56)	-2.22 (-14.04 to 9.59)	0.71

^This analysis was not pre-specified.

See Table S1 notes. Sample is limited to survey responders aged 50-64 (N=3372).

Table S14b: Mean Values and Absolute Change in Clinical Measures and Health Outcomes: pre-randomization "high risk" diagnosis

	Mean Value in Control Group (1)	Effect of Lottery Selection (95% CI) (2)	Effect of Medicaid Coverage (95% CI) (3)	P Value (4)
Blood pressure				
Systolic blood pressure (mmHg)	126.1 (19.6)	0.72 (-0.74 to 2.19)	2.90 (-3.00 to 8.80)	0.34
Diastolic blood pressure (mmHg)	80.6 (13.1)	0.22 (-0.77 to 1.20)	0.86 (-3.09 to 4.81)	0.67
Elevated blood pressure (%)	29.3	1.35 (-2.05 to 4.74)	5.41 (-8.22 to 19.04)	0.44
^Hypertension diagnosis post-lottery (%)	3.5	0.65 (-0.71 to 2.01)	2.56 (-2.80 to 7.92)	0.35
^Current medication for hypertension (%)	40.6	0.23 (-3.25 to 3.70)	0.91 (-12.92 to 14.73)	0.90
Cholesterol				
Total cholesterol (mg/dL)	205.9 (36.9)	0.14 (-2.57 to 2.86)	0.58 (-10.28 to 11.44)	0.92
High total cholesterol (%)	16.7	-0.87 (-3.45to 1.71)	-3.46 (-13.78 to 6.85)	0.51
HDL cholesterol (mg/dL)	46.6 (13.4)	0.35 (-0.66 to 1.35)	1.39 (-2.64 to 5.43)	0.50
Low HDL cholesterol (%)	31.3	-0.37 (-3.15 to 3.90)	1.49 (-12.60 to 15.58)	0.84
^High cholesterol diagnosis post-lottery (%)	8.5	-0.79 (-2.84 to 1.25)	-3.20 (-11.51 to 5.10)	0.45
^Current medication for high cholesterol (%)	25.3	0.31 (-2.95 to 3.57)	1.22 (-11.77 to 14.21)	0.85
Glycated hemoglobin				
Glycated hemoglobin level (% glycated)	5.7	0.025 (-0.045 to 0.095)	0.10 (-0.18 to 0.38)	0.48
Glycated hemoglobin level ≥ 6.5% (%)	16.0	-.97 (-3.66 to 1.73)	-3.88 (-14.72 to 6.96)	0.48
^Diabetes diagnosis post-lottery (%)	1.6	1.66 (0.53 to 2.79)	6.63 (2.04 to 11.22)	0.005
^Current medication for diabetes (%)	20.8	3.10 (0.041 to 6.15)	12.37 (0.12 to 24.61)	0.048
Depression measures				
Positive depression screen (%)	41.9	-2.91 (-6.63 to 0.81)	-11.68 (-26.80 to 3.44)	0.13
^Depression diagnosis post-lottery (%)	3.8	0.86 (-0.57 to 2.30)	3.45 (-2.28 to 9.19)	0.24
^Current medication for depression (%)	27.9	1.27 (-2.01 to 4.54)	5.06 (-7.94 to 18.06)	0.45

^This analysis was not pre-specified.

See Table S1 notes. Sample is limited to those with a pre-exising "high-risk" diagnosis, defined as a pre-randomization diagnosis of diabetes, high blood pressure, high cholesterol, heart attack, or congestive heart failure (N=3314).

Table S14c: Mean Values and Absolute Change in Clinical Measures and Health Outcomes: pre-randomization specific diagnoses

	Mean Value in Control Group (1)	Effect of Lottery Selection (95% CI) (2)	Effect of Medicaid Coverage (95% CI) (3)	P Value (4)
Blood pressure (N=2225)				
Systolic blood pressure (mmHg)	129.8 (20.7)	0.98 (-0.92 to 2.88)	3.80 (-3.63 to 11.24)	0.32
Diastolic blood pressure (mmHg)	82.9 (13.7)	0.42 (-0.82 to 1.67)	1.65 (-3.20 to 6.50)	0.51
Elevated blood pressure (%)	38.0	1.16 (-3.25 to 5.58)	4.52 (-12.61 to 21.64)	0.61
^Hypertension diagnosis post-lottery (%)	N/A	N/A	N/A	N/A
^Current medication for hypertension (%)	51	-0.0089 (-4.25 to 4.27)	0.034 (-16.28 to 16.35)	1.00
Cholesterol (N=1549)				
Total cholesterol (mg/dL)	211.5 (39.4)	-0.37 (-4.55 to 3.81)	-1.68 (-20.44 to 17.08)	0.86
High total cholesterol (%)	21.9	-0.51 (-4.79 to 3.78)	-2.28 (-21.53 to 16.98)	0.82
HDL cholesterol (mg/dL)	47.4 (13.5)	0.36 (-1.07 to 1.80)	1.64 (-4.84 to 8.13)	0.62
Low HDL cholesterol (%)	28.6	0.72 (-4.23 to 5.67)	3.24 (-18.96 to 25.45)	0.77
^High cholesterol diagnosis post-lottery (%)	N/A	N/A	N/A	N/A
^Current medication for high cholesterol (%)	36.5	3.21 (-2.17 to 8.58)	14.31 (-9.38 to 38.00)	0.24
Glycated hemoglobin (N=872)				
Glycated hemoglobin level (% glycated)	6.7 (1.2)	0.031 (-0.14 to 0.21)	0.20 (-0.88 to 1.27)	0.72
Glycated hemoglobin level ≥ 6.5% (%)	54.0	-4.39 (-11.59 to 2.81)	-27.00 (-71.91 to 17.92)	0.24
^Diabetes diagnosis post-lottery (%)	N/A	N/A	N/A	N/A
^Current medication for diabetes (%)	72.3	6.78 (0.50 to 13.07)	40.86 (-0.45 to 82.16)	0.053
Depression measures (N=4166)				
Positive depression screen (%)	52.1	-3.16 (-6.53 to 0.21)	-13.12 (-27.22 to 0.98)	0.07
^Depression diagnosis post-lottery (%)	N/A	N/A	N/A	N/A
^Current medication for depression (%)	37.5	2.38 (-0.89 to 5.64)	9.87 (-3.62 to 23.36)	0.15

^This analysis was not pre-specified.

See Table S1 notes. Sample is limited for each set of analyses to those with the relevent pre-randomization diagnosis (e.g. hypertension diagnosis for the blood pressure measures). Sample sizes are given above.

Table S15a: Mean Values and Absolute Change in Clinical Measures and Health Outcomes, Logistic Specification

	Mean Value in Control Group	Effect of Lottery Selection (95% CI) primary spec	P Value	Effect of Lottery Selection (95% CI) logistic spec	P Value
	(1)	(2)	(3)	(4)	(5)
Blood pressure					
Elevated blood pressure (%)	16.3	-0.32 (-1.73 to 1.09)	0.65	-0.26 (-1.55 to 1.04)	0.70
^Hypertension diagnosis post-lottery (%)	5.6	0.42 (-0.46 to 1.30)	0.34	0.39 (-0.40 to 1.17)	0.34
^Current medication for hypertension (%)	13.9	0.16 (-1.09 to 1.41)	0.80	0.13 (-0.85 to 1.11)	0.79
Cholesterol					
High total cholesterol (%)	14.1	-0.59 (-1.87 to 0.70)	0.37	-0.59 (-1.87 to 0.70)	0.37
Low HDL cholesterol (%)	28.0	-0.68 (-2.48 to 1.12)	0.46	-0.68 (-2.48 to 1.12)	0.46
^High cholesterol diagnosis post-lottery (%)	6.1	0.58 (-0.37 to 1.52)	0.23	0.58 (-0.37 to 1.52)	0.23
^Current medication for high cholesterol (%)	8.5	0.92 (-0.18 to 2.02)	0.10	0.91 (-0.18 to 2.01)	0.10
Glycated hemoglobin					
Glycated hemoglobin level ≥ 6.5% (%)	5.1	-0.22 (-1.07 to 0.62)	0.61	-0.22 (-1.07 to 0.62)	0.61
^Diabetes diagnosis post-lottery (%)	1.1	0.92 (0.47 to 1.38)	<.001	0.91 (0.47 to 1.34)	<0.001
^Current medication for diabetes (%)	6.4	1.31 (0.34 to 2.29)	0.008	1.30 (0.34 to 2.27)	0.008
Depression measures					
Positive depression screen (%)	30.0	-2.21 (-4.02 to -0.40)	0.017	-2.22 (-4.03 to -0.40)	0.017
^Depression diagnosis post-lottery (%)	4.8	0.91 (0.034 to 1.79)	0.042	0.91 (0.037 to 1.78)	0.041
^Current medication for depression (%)	16.8	1.33 (-0.12 to 2.77)	0.072	1.31 (-0.12 to 2.74)	0.072

^This analysis was not pre-specified.

Notes: See Table S1 notes for variable definitions. Columns 1 through 3 report the control mean, intent-to-treat estimate and 95% confidence interval (in parentheses), and associated p-value as reported in Table S1. Columns 4 and 5 report the intent-to-treat estimate, 95% confidence interval (in parentheses) and associated p-value estimated using logistic regression. The logistic regressions have the same weights, covariates, and clustering as the linear estimates and are presented as average marginal effects. Sample is survey respondents (N=12,229).

Table S15b: Mean Values and Absolute Change in Health-related quality of life and Happiness, Logistic Specification

	Mean Value in Control Group	Effect of Lottery Selection (95% CI) primary spec	P Value	Effect of Lottery Selection (95% CI) logistic spec	P Value
	(1)	(2)	(3)	(4)	(5)
Health-related quality of life					
Health same or better (%)	80.4	1.89 (0.36 to 3.43)	0.016	1.89 (0.36 to 3.42)	0.016
No or very mild pain (%)	56.4	0.28 (-1.68 to 2.23)	0.78	0.28 (-1.69 to 2.25)	0.78
Happiness					
Very or pretty happy (%)	74.9	0.29 (-1.41 to 1.98)	0.74	0.29 (-1.41 to 1.98)	0.74

Notes: See Table S15a notes. Sample is survey respondents (N=12,229).

Table S15c: Mean Values and Absolute Change in Financial Hardship, Logistic Specification

	Mean Value in Control Group	Effect of Lottery Selection (95% CI) primary spec	P Value	Effect of Lottery Selection (95% CI) logistic spec	P Value
	(1)	(2)	(3)	(4)	(5)
Any out-of-pocket spending (%)	58.8	-3.70 (-5.64 to -1.75)	<0.001	-3.70 (-5.64 to -1.76)	<0.001
Any medical debt (%)	56.8	-3.21 (-5.23 to -1.20)	0.002	-3.25 (-5.28 to -1.21)	0.002
Borrowed or skipped bills (%)	24.4	-3.44 (-5.09 to -1.79)	<0.001	-3.44 (-5.09 to -1.79)	<0.001
Catastrophic expenditures (%)	5.5	-1.08 (-1.98 to -0.17)	0.020	-1.04 (-1.92 to -0.16)	0.020

Notes: See Table S15a notes. Sample is survey respondents (N=12,229).

Table S15d: Mean Values and Absolute Change in Utilization, Prevention, Access, and Smoking and Obesity, Logistic Specification

	Mean Value in Control Group	Effect of Lottery Selection (95% CI) primary spec	P Value	Effect of Lottery Selection (95% CI) logistic spec	P Value
	(1)	(2)	(3)	(4)	(5)
Utilization (percentage using any)					
Prescription drugs (currently taking)	53.9	2.96 (0.99 to 4.93)	0.003	3 (1 to 5)	0.003
Office Visits (last 12 months)	64.6	4.00 (2.12 to 5.89)	<0.001	4.02 (2.12 to 5.91)	<0.001
Outpatient Surgery (last 12 months)	7.8	0.70 (-0.38 to 1.77)	0.20	0.69 (-0.38 to 1.76)	0.20
ED visits (last 12 month)	40.2	1.30 (-0.65 to 3.25)	0.19	1.32 (-0.66 to 3.29)	0.19
Hospital visits (last 12 months)	12.7	0.75 (-0.53 to 2.03)	0.25	0.75 (-0.53 to 2.02)	0.25
Prevention (last 12 months)					
Cholesterol-level screening (%)	27.2	3.55 (1.71 to 5.38)	<0.001	3.55 (1.72 to 5.39)	<0.001
Fecal occult-blood test (age >=50) (%	19.1	0.34 (-2.55 to 3.22)	0.82	0.34 (-2.54 to 3.22)	0.82
Colonoscopy (age >=50) (%)	10.4	1.12 (-1.16 to 3.40)	0.33	1.12 (-1.15 to 3.38)	0.33
Flu shot (age >=50) (%)	35.5	-1.54 (-5.14 to 2.07)	0.40	-1.54 (-5.14 to 2.07)	0.40
Papanicolaou smear (women) (%)	44.9	3.21 (0.56 to 5.86)	0.018	3.21 (0.56 to 5.86)	0.018
Mammogram (women >=50) (%)	28.9	7.75 (3.12 to 12.37)	0.001	7.76 (3.12 to 12.4)	0.001
PSA (men >=50) (%)	21.4	5.02 (0.25 to 9.79)	0.039	5.02 (0.26 to 9.78)	0.039
Access and Quality					
Have a usual place of care (%)	46.1	5.74 (3.73 to 7.76)	<0.001	5.75 (3.73 to 7.77)	<0.001
Received all needed care (%)	61.0	2.76 (0.88 to 4.65)	0.004	2.77 (0.88 to 4.67)	0.004
High quality, if received care (%)	78.4	2.46 (0.69 to 4.23)	0.006	2.46 (0.69 to 4.22)	0.006
Smoking and Obesity					
Currently Smoking (%)	42.8	1.35 (-0.61 to 3.31)	0.18	1.38 (-0.63 to 3.4)	0.18
Obese (%)	41.5	0.094 (-1.90 to 2.09)	0.93	0.094 (-1.90 to 2.09)	0.93

Notes: See Table S15a notes. For the prevention measures, sample is all survey respondents (N=12,229), survey respondents at least 50 years of age (N=3374), female survey respondents (N=6915), female survey respondents at least 50 years of age (N=1864) or male survey respondents at least 50 years of age (N=1509), as indicated in the table. For all other measures, sample is all survey respondents (N=12, 229) except for quality of care which is only defined on those who received care in the last 12 months (N= 9,694).

Table S16a: Mean Values and Absolute Change in Clinical Measures and Health Outcomes, Sensitivity to Covariates

	Mean Value in Control Group	Effect of Lottery Selection (95% CI) unadjusted	P Value	Effect of Lottery Selection (95% CI) age-sex adjusted	P Value	Effect of Lottery Selection (95% CI) fully adjusted	P Value
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Blood pressure							
Systolic blood pressure (mmHg)	119.3 (16.9)	-0.011 (-0.67 to 0.65)	0.97	-0.13 (-0.72 to 0.47)	0.68	-0.13 (-0.74 to 0.47)	0.67
Diastolic blood pressure (mmHg)	76.0 (12.1)	-0.17 (-0.64 to 0.31)	0.50	-0.19 (-0.64 to 0.25)	0.39	-0.21 (-0.66 to 0.24)	0.36
Elevated blood pressure (%)	16.3	-0.21 (-1.67 to 1.26)	0.78	-0.32 (-1.73 to 1.09)	0.65	-0.20 (-1.63 to 1.23)	0.78
^Hypertension diagnosis post-lottery (%)	5.6	0.43 (-0.46 to 1.32)	0.34	0.42 (-0.46 to 1.30)	0.34	0.43 (-0.46 to 1.32)	0.35
^Current medication for hypertension (%)	13.9	0.32 (-1.03 to 1.67)	0.64	0.16 (-1.09 to 1.41)	0.80	0.058 (-1.20 to 1.32)	0.93
Cholesterol							
Total cholesterol (mg/dL)	204.1 (34.0)	0.53 (-0.83 to 1.89)	0.44	0.53 (-0.82 to 1.89)	0.44	0.63 (-0.74 to 2.00)	0.37
High total cholesterol (%)	14.1	-0.59 (-1.87 to 0.70)	0.37	-0.62 (-1.89 to 0.66)	0.34	-0.67 (-1.96 to 0.62)	0.31
HDL cholesterol (mg/dL)	47.6 (13.1)	0.20 (-0.32 to 0.72)	0.45	0.22 (-0.29 to 0.74)	0.39	0.27 (-0.25 to 0.79)	0.31
Low HDL cholesterol (%)	28.0	-0.68 (-2.48 to 1.12)	0.46	-0.72 (-2.52 to 1.07)	0.43	-0.97 (-2.79 to 0.86)	0.30
^High cholesterol diagnosis post-lottery (%)	6.1	0.58 (-0.37 to 1.52)	0.23	0.5 (-0.44 to 1.43)	0.30	0.55 (-0.40 to 1.50)	0.26
^Current medication for high cholesterol (%)	8.5	0.92 (-0.18 to 2.02)	0.10	0.78 (-0.27 to 1.83)	0.15	0.78 (-0.29 to 1.85)	0.15
Glycated hemoglobin							
Glycated hemoglobin level (% glycated)	5.3 (0.6)	0.0029 (-0.022 to 0.028)	0.82	).00090 (-0.023 to 0.025	0.94	).00017 (-0.024 to 0.025	0.99
Glycated hemoglobin level ≥ 6.5% (%)	5.1	-0.22 (-1.07 to 0.62)	0.61	-0.25 (-1.10 to 0.59)	0.56	-0.30 (-1.17 to 0.56)	0.49
^Diabetes diagnosis post-lottery (%)	1.1	0.92 (0.47 to 1.38)	<0.001	0.89 (0.44 to 1.35)	<0.001	0.83 (0.37 to 1.28)	<0.001
^Current medication for diabetes (%)	6.4	1.31 (0.34 to 2.29)	0.008	1.24 (0.29 to 2.20)	0.011	1.18 (0.21 to 2.15)	.018
Mental health measures							
Positive depression screen (%)	30.0	-2.21 (-4.02 to -0.40)	0.017	-2.18 (-3.97 to -0.38)	0.02	-1.86 (-3.68 to -0.042)	0.045
^Depression diagnosis post-lottery (%)	4.8	0.91 (0.034 to 1.79)	0.042	0.93 (0.062 to 1.80)	0.04	0.93 (0.039 to 1.82)	0.041
^Current medication for depression (%)	16.8	5.49 (-0.46 to 11.45)	0.072	5.25 (-0.57 to 11.07)	0.078	5.29 (-0.61 to 11.19)	0.080
Predicted Cardiovascular Risk							
Framingham Risk Score	8.2 (7.5)	-0.051 (-0.39 to 0.28)	0.77	-0.073 (-0.29 to 0.15)	0.51	-0.054 (-0.27 to 0.17)	0.63
Limited to “high risk” Dx	11.6 (8.3)	0.40 (-0.27 to 1.07)	0.24	0.22 (-0.25 to 0.69)	0.36	0.30 (-0.17 to 0.77)	0.21
Limited to aged 50-64	13.9 (8.2)	-0.098 (-0.71 to 0.51)	0.75	-0.19 (-0.67 to 0.29)	0.45	-0.15 (-0.64 to 0.34)	0.55

^This analysis was not pre-specified.

Notes: See Table S1 notes for variables definition. Column 1 reports the weighted mean of the dependent variable in the control sample of survey respondents and standard deviation (in parenthesis) for continuous outcomes. Columns 2 and 3 reports intent-to-treat estimates and p-values for an unadjusted estimate. Columns 4 and 5 report the intent-to-treat estimates and p-values adding controls for age (in decile bins) and sex. Columns 6 and 7 report the intent-to-treat estimates and p-values adding controls for age (in decile bins), sex ,race and six additional pre-randomization variables as described in the appendix. All regressions include indicators for each household size, and all standard errors are clustered on the household. All analysis is weighted using survey weights. Sample size is 12229, except for the Framingham risk scores which have sample sizes of 9525, 3099, and 3372 respectively. Framingham risk scores are only defined for individuals of age 30 years or older.

Table S16b: Mean Values and Absolute Change in Health-related quality of life and Happiness, Sensitivity to covariates

	Control Mean (1)	Effect of Lottery unadjusted (2)	P Value (3)	Effect of Lottery age-sex adjusted (4)	P Value (5)	Effect of Lottery fully adjusted (6)	P Value (7)
Health-related quality of life							
Health same or better (%)	80.4	1.89 (0.36 to 3.43)	0.016	1.84 (0.31 to 3.36)	0.019	1.94 (0.39 to 3.49)	0.014
SF-8 mental component score	44.4 (11.4)	0.47 (0.0097 to 0.93)	0.047	0.45 (-0.0051 to 0.91)	0.054	0.44 (-0.018 to 0.90)	0.061
SF-8 physical component score	45.5 (10.5)	0.29 (-0.13 to 0.71)	0.18	0.29 (-0.11 to 0.69)	0.15	0.30 (-0.11 to 0.70)	0.15
No or very mild pain (%)	56.4	0.28 (-1.68 to 2.23)	0.78	0.30 (-1.59 to 2.20)	0.75	0.25 (-1.67 to 2.18)	0.80
Happiness							
Very or pretty happy (%)	74.9	0.29 (-1.41 to 1.98)	0.74	0.29 (-1.39 to 1.96)	0.74	0.52 (-1.18 to 2.23)	0.55

Notes: See Table S16a notes. Sample is all survey respondents (N=12, 229)

Table S16c: Mean Values and Absolute Change in Financial Hardship, Sensitivity to covariates

	Control Mean (1)	Effect of Lottery unadjusted (2)	P Value (3)	Effect of Lottery age-sex adjusted (4)	P Value (5)	Effect of Lottery fully adjusted (6)	P Value (7)
Any out-of-pocket spending (%)	58.8	-3.70 (-5.64 to -1.75)	<0.001	-3.73 (-5.64 to -1.82)	<0.001	-3.87 (-5.81 to -1.93)	<0.001
Amount of out-of-pocket spending ($)	552.8 (1219.5)	-52.13 (-99.14 to -5.12)	0.029	-52.97 (-99.54 to -6.39)	0.025	-53.42 (-100.75 to -6.09)	0.026
Catastrophic expenditures (% any)	5.5	-1.08 (-1.98 to -0.17)	0.020	-1.06 (-1.95 to -0.16)	0.021	-1.03 (-1.95 to -0.12)	0.027
Any medical debt (%)	56.8	-3.21 (-5.23 to -1.20)	0.002	-3.11 (-5.12 to -1.10)	0.002	-3.09 (-5.11 to -1.07)	0.003
Borrowed money or skipped bills (%)	24.4	-3.44 (-5.09 to -1.79)	<0.001	-3.43 (-5.06 to -1.81)	<0.001	-3.45 (-5.10 to -1.81)	<0.001

Notes: See Table S16a notes. Sample is all survey respondents (N=12, 229)

Table S16d: Utilization, Prevention, Access and Quality, and Smoking and Obesity, Sensitivity to Covariates

	Control Mean (1)	Effect of Lottery unadjusted (2)	P Value (3)	Effect of Lottery age-sex adjusted (4)	P Value (5)	Effect of Lottery fully adjusted (6)	P Value (7)
Utilization							
Prescription drugs (currently taking)	1.8 (2.8)	0.16 (0.05 to 0.27)	0.004	0.16 (0.051 to 0.26)	0.003	0.15 (0.047 to 0.26)	0.004
Office Visits (last 12 months)	5.5 (11.6)	0.65 (0.21 to 1.08)	0.003	0.68 (0.25 to 1.11)	0.002	0.69 (0.25 to 1.13)	0.002
Outpatient Surgery (last 12 months)	0.1 (0.4)	).0077 (-0.0064 to 0.022	0.28	).0078 (-0.0063 to 0.022	0.28	).0083 (-0.0061 to 0.023	0.26
ED visits (last 12 month)	1.0 (2.0)	0.023 (-0.056 to 0.10)	0.57	0.029 (-0.049 to 0.11)	0.46	0.031 (-0.048 to 0.11)	0.43
Hospital visits (last 12 months)	0.2 (0.6)	0.016 (-0.0074 to 0.040)	0.17	0.017 (-0.0066 to 0.041)	0.15	0.02 (-0.0042 to 0.044)	0.10
Prevention (last 12 months)							
Cholesterol-level screening (%)	27.2	3.55 (1.71 to 5.38)	<0.001	3.29 (1.52 to 5.06)	<0.001	3.08 (1.29 to 4.87)	0.001
Fecal occult-blood test (age >=50) (%)	19.1	0.34 (-2.55 to 3.22)	0.82	0.24 (-2.65 to 3.13)	0.87	0.32 (-2.62 to 3.27)	0.83
Colonoscopy (age >=50) (%)	10.4	1.12 (-1.16 to 3.40)	0.33	1.02 (-1.25 to 3.29)	0.37	1.22 (-1.10 to 3.53)	0.30
Flu shot (age >=50) (%)	35.5	-1.54 (-5.14 to 2.07)	0.41	-1.75 (-5.34 to 1.83)	0.34	-1.94 (-5.59 to 1.70)	0.30
Papancolaou smear (women) (%)	44.9	3.21 (0.56 to 5.86)	0.016	3.50 (0.90 to 6.10)	0.008	3.37 (0.74 to 6.00)	0.011
Mammogram (women >=50) (%)	28.9	7.75 (3.12 to 12.37)	0.001	7.59 (2.95 to 12.23)	0.001	7.30 (2.60 to 12.00)	0.002
PSA (men >=50) (%)	21.4	5.02 (0.25 to 9.79)	0.037	4.96 (0.23 to 9.69)	0.038	3.92 (-0.83 to 8.66)	0.10
Access and Quality							
Have a usual place of care (%)	46.1	5.74 (3.73 to 7.76)	<0.001	5.73 (3.75 to 7.70)	<0.001	5.62 (3.62 to 7.62)	<0.001
Received all needed care (%)	61.0	2.76 (0.88 to 4.65)	0.004	2.68 (0.79 to 4.56)	0.005	2.87 (0.98 to 4.76)	0.003
High quality, if received care (%)	78.4	2.46 (0.69 to 4.23)	0.007	2.36 (0.60 to 4.13)	0.009	2.25 (0.45 to 4.05)	0.014
Smoking and Obesity							
Currently Smoking (%)	42.8	1.35 (-0.61 to 3.31)	0.18	1.46 (-0.49 to 3.41)	0.14	1.64 (-0.31 to 3.58)	0.099
Obese (%)	41.5	0.094 (-1.90 to 2.09)	0.93	0.089 (-1.91 to 2.08)	0.93	-0.14 (-2.16 to 1.87)	0.89

Notes: See Table S16a notes. For the prevention measures, sample is all survey respondents (N=12,229), survey respondents at least 50 years of age (N=3374), female survey respondents (N=6915), female survey respondents at least 50 years of age (N=1864) or male survey respondents at least 50 years of age (N=1509), as indicated in the table. For all other measures, sample is all survey respondents (N=12, 229) except for quality of care which is only defined on those who received care in the last 12 months (N= 9,694).

Table S17: Mean Values and Absolute Change in Additional Clinical Health Measures

	Mean Value in Control Group	Effect of Lottery Selection (95% CI)	Effect of Medicaid Coverage (95% CI)	P Value
	(1)	(2)	(3)	(4)
Blood pressure				
Pre-hypertension or Hypertension(%)[1]	49.3	-0.38 (-2.21 to 1.45)	-1.57 (-9.12 to 5.98)	0.68
Cholesterol				
Elevated total cholesterol(%)[2]	50.9	1.74 (-0.28 to 3.77)	7.21 (-1.19 to 15.62)	0.092
Glycated hemoglobin				
Pre-diabetic or Diabetic(%)[3]	10.2	-0.11 (-1.31 to 1.09)	-0.45 (-5.43 to 4.52)	0.86

Notes: See Table S1 notes. Sample size is all survey responders (N=12229).

[1]Pre-hypertension is defined as having a systolic BP measure of 120 mmHg or above and a diastolic BP measure of 80 mmHg or above. Hypertension is defined as having a systolic BP measure of 140 mmHg or above and a diastolic BP measure of 90 mmHg or above.

[2]Elevated total cholesterol is defined as having a total cholesterol level of 200 mg/dL or higher.

[3]Glycated hemoglobin A1c measure of 6.5% or higher is used as a criterion for diabetes. Glycated Hemoglobin A1c measure of 5.7% or higher is used as a criterion for pre-diabetes.

Table S18a: Self-Reported Health and Happiness, comparing mail survey and inperson results

	Mail survey results for all mail survey respondents			In-person results for all in-person respondents		
	Mean Value in Control Group	**Effect of Medicaid Coverage (95% CI)**	**P Value**	**Mean Value in Control Group**	**Change with Medicaid Coverage (95% CI)**	**P Value**
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
Self-reported health						
Positive depression screen(%)	32.9	-7.83 (-12.67 to -2.99)	0.002	30.0	-9.15 (-16.70 to -1.60)	0.018
Health g/vg/e (%)	54.8	13.29 (8.11 to 18.47)	<0.001	59.6	7.08 (-1.10 to 15.27)	0.090
Health not poor or very poor(%)	86.0	9.90 (6.45 to 13.36)	<0.001	85.8	4.50 (-1.17 to 10.16)	0.12
Health same or gotten better (%)	71.4	11.28 (6.69 to 15.86)	<0.001	80.4	7.84 (1.45 to 14.23)	0.016
STE		0.27	<0.001		0.14	0.016
se		(0.038)			(0.058)	
Happiness						
Very or pretty happy (%)	59.4	19.11 (14.08 to 24.14)	<0.001	74.9	1.18 (-5.85 to 8.21)	0.74

Notes: The first set of results (columns 1-3) are for the mail survey respondents using outcomes from the mail survey and are also reported in Finkelstein et al 2012. The second set of results (columns 4-6) are for the inperson survey respondents using outcomes from the inperson survey and, with a few exceptions, are also reported in the mail paper. The mail survey results are weighed using the mail survey weights; the inperson results are weighted using the inperson weights. Columns 1 and 4 report the weighted mean of the dependent variable in the control sample of survey respondents and standard deviation (in parenthesis) for continuous outcomes. Columns 2 and 5 report local average treatment effect estimates, with corresponding p-values in columns 3 and 6. All regressions include indicators for each household size, and all standard errors are clustered on the household. All mail survey regressions also include indicators for survey wave and interactions between household size and survey wave. In the inperson survey, depression screening is based on answers to PHQ-8 questions. In the mail survey, depression screening is based on answers to two questions: "Over the past 2 weeks, how often have you been bothered by little interest or pleasure in doing things?" and "Over the past 2 weeks, how often have you been bothered by feeling down, depressed, or hopeless?" Sample size for the mail survey is 23,741 and sample size for the inperson is 12,229.

Table S18b: Financial hardship, comparing mail survey and inperson results

	Mail survey results for all mail survey respondents			In-person results for all in-person respondents		
	Mean Value in Control Group	**Effect of Medicaid Coverage (95% CI)**	**P Value**	**Mean Value in Control Group**	**Effect of Medicaid Coverage (95% CI)**	**P Value**
	(1)	**(2)**	**(3)**	**(1)**	**(2)**	**(3)**
Any out-of-pocket spending (%)	55.5	-19.95 (-25.09 to -14.81)	<0.001	58.8	-15.30 (-23.28 to -7.32)	<0.001
Any medical debt (%)	59.7	-17.98 (-23.17 to -12.80)	<0.001	56.8	-13.28 (-21.59 to -4.96)	0.002
Borrowed or skipped bills (%)	36.4	-15.43 (-20.36 to -10.49)	<0.001	24.4	-14.22 (-21.02 to -7.43)	<0.001
STE		-0.36	<0.001		-0.30	<0.001
se		(0.04)			(0.059)	

Notes: See Table S18a notes.

Table S18c: Utilization and spending, comparing mail survey and inperson results

	Mail survey results for all mail survey respondents			In-person results for all in-person respondents		
	Control Mean	**Effect of Medicaid Coverage (95% CI)**	**P Value**	**Control Mean**	**Effect of Medicaid Coverage (95% CI)**	**P Value**
	(1)	**(2)**	**(3)**	**(1)**	**(2)**	**(3)**
Prescription drugs (currently taking)	2.3 (2.9)	0.35 (0.0033 to 0.69)	0.048	1.8 (2.8)	0.66 (0.21 to 1.11)	0.004
Office Visits	3.8 (6.2)	2.17 (1.45 to 2.88)	<0.001	5.5 (11.6)	2.70 (0.91 to 4.49)	0.003
Outpatient Surgery	NA	NA	NA	0.1 (0.4)	0.032 (-0.027 to 0.091)	0.28
ED visits	0.9 (2.1)	0.051 (-0.17 to 0.27)	0.64	1.0 (2.0)	0.094 (-0.23 to 0.42)	0.57
Hospital visits	0.2 (0.8)	0.043 (-0.040 to 0.13)	0.31	0.2 (0.6)	0.068 (-0.030 to 0.17)	0.17
Annual Spending ($)	3172.97	778.15	0.036	3257.28	1171.63	0.018
se		(371.41)			(496.06)	

Notes: See Table S18a notes. In the 12-month survey all visits were in the past 6 months; in the inperson interviews all visits were in the past 12 months. For both, we calculate spending on an annual basis.

Table S18d: Prevention, Access and Quality, and Smoking, comparing mail survey and inperson results

	Mail survey results for all mail survey respondents			In-person results for all in-person respondents		
	Mean Value in Control Group	**Effect of Medicaid Coverage (95% CI)**	**P Value**	**Mean Value in Control Group**	**Effect of Medicaid Coverage (95% CI)**	**P Value**
	(1)	**(2)**	**(3)**	**(1)**	**(2)**	**(3)**
Prevention (last 12 months)						
Cholesterol-level screening (%)	62.5	11.43 (6.43 to 16.43)	<0.001	27.2	14.57 (7.09 to 22.04)	<0.001
Papanicolaou smear (women) (%)	40.6	18.26 (11.59 to 24.94)	<0.001	44.9	14.44 (2.64 to 26.24)	0.016
Mammogram (women >=50) (%)	29.8	18.69 (10.95 to 26.43)	<0.001	28.9	29.67 (11.96 to 47.37)	0.001
STE		0.34	<0.001		0.42	<0.001
se		(0.047)			(0.093)	
Access and Quality						
Have a usual place of care (%)	49.9	33.89 (28.5 to 39.28)	<0.001	46.1	23.75 (15.44 to 32.06)	<0.001
Received all needed care (%)	68.4	23.94 (19.58 to 28.30)	<0.001	61.0	11.43 (3.62 to 19.24)	0.004
High quality, if received care (%)	70.8	14.15 (8.93 to 19.38)	<0.001	78.4	9.85 (2.71 to 17.00)	0.007
STE		0.37	<0.001		0.27	<0.001
se		(0.028)			(0.046)	
Health Behavior						
Currently Smoking (%)	41.7	-0.36 (-5.56 to 4.83)	0.89	42.8	5.58 (-2.54 to 13.70)	0.18

Notes: See Table S18a notes. The mail survey results use "cholesterol checked ever" rather than "cholesterol checked in the last 12 months.

Table S18e: Self-Reported Health and Happiness, comparing mail survey and inperson results

	Mail survey results for overlap sample			In-person results for overlap sample		
	Mean Value in Control Group	**Effect of Medicaid Coverage (95% CI)**	**P Value**	**Mean Value in Control Group**	**Effect of Medicaid Coverage (95% CI)**	**P Value**
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
Self-reported health						
Positive depression screen(%)	34.5	-16.36 (-27.16 to -5.56)	0.003	31.0	-19.96 (-31.02 to -8.90)	<0.001
Health g/vg/e (%)	55.1	19.83 (8.03 to 31.63)	0.001	59.0	11.22 (-0.67 to 23.11)	0.064
Health not poor or very poor(%)	87.2	7.01 (-0.70 to 14.72)	0.075	85.9	2.95 (-5.33 to 11.23)	0.48
Health same or gotten better (%)	72.8	14.25 (4.20 to 24.30)	0.005	79.5	14.26 (4.77 to 23.76)	0.003
STE		0.34	<0.001		0.27	0.001
se		(0.088)			(0.086)	
Happiness						
Very or pretty happy (%)	57.5	21.8 (10.23 to 33.37)	<0.001	73.1	11.93 (1.46 to 22.41)	0.026

Notes: The first set of results (columns 1-3) are for the overlap sample and use outcomes from the mail survey. The second set of results (columns 4-6) are the overlap sample and use outcomes from the inperson survey. All analysis is weighted using the interaction between the 12-month mail survey weights and the inperson survey weights. Columns 1 and 4 report the weighted mean of the dependent variable in the control sample of survey respondents and standard deviation (in parenthesis) for continuous outcomes. Columns 2 and 5 report local average treatment effect estimates, with corresponding p-values in columns 3 and 6. All regressions include indicators for each household size, and all standard errors are clustered on the household. All mail survey regressions also include indicators for survey wave and interactions between household size and survey wave. In the inperson survey, depression screening is based on answers to PHQ-8 questions. In the mail survey, depression screening is based on answers to two questions: "Over the past 2 weeks, how often have you been bothered by little interest or pleasure in doing things?" and "Over the past 2 weeks, how often have you been bothered by feeling down, depressed, or hopeless?" Sample is the overlap in the mail survey respondents and the inperson respondents (N=5,750).

Table S18f: Financial hardship, comparing mail survey and inperson results

	Mail survey results for overlap sample			In-person results for overlap sample		
	Mean Value in Control Group	**Effect of Medicaid Coverage (95% CI)**	**P Value**	**Mean Value in Control Group**	**Effect of Medicaid Coverage (95% CI)**	**P Value**
	(1)	**(2)**	**(3)**	**(1)**	**(2)**	**(3)**
Any out-of-pocket spending (%)	51.9	-13.60 (-25.22 to -1.97)	0.022	60.4	-11.09 (-23.21 to 1.02)	0.073
Any medical debt (%)	60.1	-23.14 (-35.20 to -11.08)	<0.001	57.2	-21.26 (-33.61 to -8.90)	<0.001
Borrowed or skipped bills (%)	32.7	-8.73 (-19.41 to 1.94)	0.11	25.7	-16.12 (-26.31 to -5.93)	0.002
STE		-.031	<0.001		-0.34	<0.001
se		(0.092)			(0.087)	

Notes: See Table S18e notes.

Table S18g: Utilization and spending, comparing mail survey and inperson results

	Mail survey results for overlap sample			In-person results for overlap sample		
	Control Mean	**Effect of Medicaid Coverage (95% CI)**	**P Value**	**Control Mean**	**Effect of Medicaid Coverage (95% CI)**	**P Value**
	(1)	**(2)**	**(3)**	**(1)**	**(2)**	**(3)**
Prescription drugs (currently taking)	2.2 (2.8)	0.24 (-0.48 to 0.96)	0.51	2.1 (3.0)	0.34 (-0.33 to 1.01)	0.32
Office Visits	4.0 (6.7)	1.65 (0.056 to 3.25)	0.042	5.9 (11.7)	2.17 (-0.54 to 4.87)	0.12
Outpatient Surgery	NA	NA	NA	0.09 (0.4)	0.067 (-0.016 to 0.15)	0.11
ED visits	1.0 (2.2)	-0.018 (-0.54 to 0.50)	0.95	0.9 (1.9)	-0.13 (-0.56 to 0.30)	0.55
Hospital visits	0.2 (0.9)	-0.098 (-0.28 to 0.085)	0.29	0.2 (0.6)	0.016 (-0.13 to 0.16)	0.82
Annual Spending ($)	3467.61	-419.33	0.61	3299.97	507.14	0.48
se		(816.81)			(718.18)	

Notes: See Table S18e notes. In the 12-month survey all visits were in the past 6 months; in the inperson interviews all visits were in the past 12 months. For both, we calculate spending on an annual basis.

Table S18h: Prevention, Access and Quality, and Smoking, comparing mail survey and inperson results

	Mail survey results for overlap sample			In-person results for overlap sample		
	Mean Value in Control Group	**Effect of Medicaid Coverage (95% CI)**	**P Value**	**Mean Value in Control Group**	**Effect of Medicaid Coverage (95% CI)**	**P Value**
	(1)	**(2)**	**(3)**	**(1)**	**(2)**	**(3)**
Prevention (last 12 months)						
Cholesterol-level screening (%)	59.6	21.41 (9.88 to 32.94)	<0.001	28.2	17.84 (6.78 to 28.89)	0.002
Papanicolaou smear (women) (%)	41.4	26.88 (11.86 to 41.90)	0.001	45.9	7.33 (-8.73 to 23.40)	0.37
Mammogram (women >=50) (%)	28.9	23.96 (8.26 to 39.67)	0.003	31.9	18.45 (-2.65 to 39.54)	0.087
STE		0.50	<0.001		0.31	0.009
se		(0.10)			(0.12)	
Access and Quality						
Have a usual place of care (%)	47.0	33.67 (21.62 to 45.72)	<0.001	48.0	33.86 (21.48 to 46.24)	<0.001
Received all needed care (%)	69.7	19.32 (9.45 to 29.20)	<0.001	61.0	15.23 (3.67 to 26.79)	0.010
High quality, if received care (%)	74.6	3.16 (-8.23 to 14.56)	0.59	80.2	12.95 (3.02 to 22.88)	0.010
STE		0.30	<0.001		0.34	<0.001
se		(0.063)			(0.07)	
Health Behavior						
Currently Smoking (%)	39.0	0.93 (-10.35 to 12.21)	0.87	39.2	1.98 (-9.69 to 13.65)	1

Notes: See Table S18e notes. The mail survey results use "cholesterol checked ever" rather than "cholesterol checked in the last 12 months.

Table S19: Observational Estimates of Effect of Insurance in Study Population

		Random assignment	Any Medicaid vs. No Medicaid	Any medicaid vs. No Medicaid (controls only)	OHP Standard vs. No Medicaid (treatment only)
		(1)	(2)	(3)	(4)
Sample Size		12229	12229	5842	6387
% Insurance		31	31	19	30
Blood pressure					
Systolic blood pressure (mmHg)		-0.52	-1.74	-2.03	0.053
	se	(1.25)	(0.32)	(0.57)	(0.45)
	p	[0.68]	[<0.001]	[<0.001]	[0.91]
Diastolic blood pressure (mmHg)		-0.81	-0.60	-0.73	0.54
	se	(0.94)	(0.24)	(0.41)	(0.33)
	p	[0.39]	[0.012]	[0.076]	[0.11]
Elevated blood pressure (%)		-0.013	-0.012	-0.0088	0.0016
	se	(0.03)	(0.0074)	(0.013)	(0.0115)
	p	[0.65]	[0.11]	[0.49]	[0.88]
^Hypertension diagnosis post-lottery (%)		0.018	0.020	0.011	0.028
	se	(.019)	(0.0052)	(0.009)	(0.0078)
	p	[0.34]	[<0.001]	[0.20]	[<0.001]
^Current medication for hypertension (%)		0.0066	0.056	0.066	0.039
	se	(0.026)	(0.0072)	(0.013)	(0.01)
	p	[0.80]	[<0.001]	[<0.001]	[<0.001]
Cholesterol					
Total cholesterol (mg/dL)		2.20	-1.95	-2.87	0.093
	se	(2.88)	(0.72)	(1.25)	(1.00)
	p	[0.45]	[0.006]	[0.022]	[0.93]
High total cholesterol (%)		-0.024	-0.015	-0.022	-0.0034
	se	-0.027	(0.0069)	(0.012)	(0.0098)
	p	[0.37]	[0.034]	[0.055]	[0.73]
HDL cholesterol (mg/dL)		0.83	-0.74	-1.30	0.032
	se	(1.09)	(0.28)	(0.46)	(0.40)
	p	[0.45]	[0.008]	[0.005]	[0.94]
Low HDL cholesterol (%)		-0.028	0.018	0.022	0.0082
	se	(0.038)	(0.0097)	(0.017)	(0.013)
	p	[0.46]	[0.068]	[0.2]	[0.54]
^High cholesterol diagnosis post-lottery (%)		0.024	0.025	0.016	0.03
	se	(0.020)	(0.0055)	(0.01)	(0.0076)
	p	[0.23]	[<0.001]	[0.11]	[<0.001]
^Current medication for high cholesterol (%)		0.038	0.040	0.050	0.031
	se	(0.023)	(0.0065)	(0.012)	(0.0092)
	p	[0.10]	[<0.001]	[0.00004164]	[0.001]

Table S19, continued

Glycated hemoglobin					
Glycated hemoglobin (% glycated)		0.012	0.0062	0.048	-0.02
	se	(0.052)	(0.014)	(0.03)	(0.017)
	p	[0.82]	[0.656]	[0.11]	[0.26]
Glycated hemoglobin ≥ 6.5% (%)		-0.0093	0.00084	0.012	-0.0027
	se	(0.018)	(0.0048)	(0.0098)	(0.0061)
	p	[0.61]	[0.86]	[0.23]	[0.66]
^Diabetes diagonsis post-lottery (%)		0.038	0.012	0.012	0.007
	se	(0.0097)	(0.0029)	(0.0044)	(0.0044)
	p	[<0.001]	[<0.001]	[0.007]	[0.11]
^Current medication for diabetes (%)		0.054	0.016	0.024	-0.0019
	se	(0.021)	(0.0056)	(0.01)	(0.0077)
	p	[0.008]	[0.004]	[0.021]	[0.80]
Depression measures					
Positive depression screen (%)		-0.091	0.054	0.081	0.029
	se	(0.039)	(0.0098)	(0.018)	(0.013)
	p	[0.018]	[<0.001]	[<0.001]	[0.027]
^Depression diagnosis post-lottery (%)		0.038	0.027	0.022	0.015
	se	(0.019)	(0.0052)	(0.0084)	(0.0073)
	p	[0.041]	[<0.001]	[0.008]	[0.036]
^Current medication for depression (%)		0.055	0.089	0.10	0.067
	se	(0.030)	(0.0084)	(0.015)	(0.012)
	p	[0.071]	[<0.001]	[<0.001]	[<0.001]

Notes: See Table S1 for variable definitions. Column 1 reports our experiments estimates of the effect of insurance coverage (LATE) as reported in Table S1. All other columns are based on observational comparisons of insurance coverage in our population. Column 2 compares all those with any Medicaid coverage during our study period to those without Medicaid (regardless of lottery status); this represents the "as treated" analysis sometimes done in clinical trials. To avoid having much of the variation in insurance coming from the lottery, the third column performs the same analysis within the control group; here, most of the insurance coverage is OHP Plus which covers a somewhat different population than OHP Standard. The fourth column therefore performs the analysis within the treatment group, comparing those on OHP Standard to those with no Medicaid (and dropping the small percentage of treatment individuals on Plus). All regressions include indicators for household size, and all standard errors are clustered on the household in the calculation of confidence intervals. All analyses are weighted using survey weights. The regression with BP-related measures also control for gender and age-decile groups. For each estimate we report the coefficient, standard error (in parentheses), and p-value (in brackets). Sample is survey respondents (N=12,229).

References

Allen, H., K. Baicker, A. Finkelstein, S. Taubman, B. J. Wright & The Oregon Health Study Group (2010) What the Oregon health study can tell us about expanding medicaid. *Health Aff (Millwood)*, 29**,** 1498-506.

American Diabetes Association (2010) Standards of medical care in diabetes--2010. *Diabetes Care*, 33 Suppl 1**,** S11-61.

Angrist, J. D. (2001) Estimation of Limited Dependent Variable Models with Dummy Endogenous Regressors: Simple Strategies for Empirical Practice. *Journal of Business & Economic Statistics*, 19**,** 2-16.

Angrist, J. D., G. W. Imbens & D. B. Rubin (1996) Identification of Causal Effects Using Instrumental Variables. *Journal of the American Statistical Association*, 91**,** 444-455.

Angrist, J. D. & J.-S. Pischke. 2009. *Mostly Harmless Econometrics: An Empiricist's Companion*. Princeton, NJ: Princeton University Press.

Baicker, K., S. Taubman, H. Allen, M. Bernstein, J. Gruber, J. P. Newhouse, E. Schneider, B. Wright, A. Zaslavsky, A. Finkelstein & the Oregon Health Study Group. The Oregon Health Insurance Experiment: Evidence from the In-Person Interviews. Analysis Plan; Archived on April 16, 2012. Available at http://www.nber.org/oregon.

Centers for Disease Control and Prevention (CDC). 2000. Behavioral Risk Factor Surveillance System Survey Questionnaire. Atlanta, GA: U.S Department of Health and Human Services, Centers for Disease Control and Prevention.

Centers for Disease Control and Prevention (CDC) National Center for Health Statistics (NCHS). 2000. National Health and Nutrition Examination Survey Questionnaire. Hyattsville, MD: U.S. Department of Health and Human Services, Centers for Disease Control and Prevention.

Chobanian, A. V., G. L. Bakris, H. R. Black, W. C. Cushman, L. A. Green, J. L. Izzo, D. W. Jones, B. J. Materson, S. Oparil, J. T. Wright, E. J. Roccella, National Heart Lung and Blood Institute Joint National Committee on Prevention Detection Evaluation and Treatment of High Blood Pressure & National High Blood Pressure Education Program Coordinating Committee (2003) The Seventh Report of the Joint National Committee on Prevention, Detection, Evaluation, and Treatment of High Blood Pressure: the JNC 7 report. *JAMA*, 289**,** 2560-72.

Chou, R., J. M. Croswell, T. Dana, C. Bougatsos, I. Blazina, R. Fu, K. Gleitsmann, H. C. Koenig, C. Lam, A. Maltz, J. B. Rugge & K. Lin (2011) Screening for prostate cancer: a review of the evidence for the U.S. Preventive Services Task Force. *Ann Intern Med*, 155**,** 762-71.

Cole, S. R. & M. A. Hernán (2008) Constructing inverse probability weights for marginal structural models. *Am J Epidemiol*, 168**,** 656-64.

D'Agostino, R. B., R. S. Vasan, M. J. Pencina, P. A. Wolf, M. Cobain, J. M. Massaro & W. B. Kannel (2008) General cardiovascular risk profile for use in primary care: the Framingham Heart Study. *Circulation*, 117**,** 743-53.

Expert Panel on Detection Evaluation And Treatment of High Blood Cholesterol In Adults (2001) Executive Summary of The Third Report of The National Cholesterol Education Program (NCEP). *JAMA*, 285**,** 2486-97.

Expert Panel on the Identification Evaluation and Treatment of Overweight and Obesity in Adults (1998) Clinical guidelines on the identification, evaluation, and treatment of overweight and obesity in adults: executive summary. *Am J Clin Nutr*, 68**,** 899-917.

Finkelstein, A., S. Taubman, B. Wright, M. Bernstein, J. Gruber, J. P. Newhouse, H. Allen, K. Baicker & O. H. S. Group (2012) The Oregon Health Insurance Experiment: Evidence from the First Year. *The Quarterly Journal of Economics*, 127**,** 1057-1106.

Fiore, A. E., T. M. Uyeki, K. Broder, L. Finelli, G. L. Euler, J. A. Singleton, J. K. Iskander, P. M. Wortley, D. K. Shay, J. S. Bresee, N. J. Cox & Centers for Disease Control and Prevention (CDC) (2010) Prevention and control of influenza with vaccines: recommendations of the Advisory Committee on Immunization Practices (ACIP), 2010. *MMWR Recomm Rep*, 59**,** 1-62.

Flegal, K. M., M. D. Carroll, C. L. Ogden & L. R. Curtin (2010) Prevalence and trends in obesity among US adults, 1999-2008. *JAMA,* 303**,** 235-41.
Greene, K. L., P. C. Albertsen, R. J. Babaian, H. B. Carter, P. H. Gann, M. Han, D. A. Kuban, A. O. Sartor, J. L. Stanford, A. Zietman & P. Carroll (2009) Prostate specific antigen best practice statement: 2009 update. *J Urol,* 182**,** 2232-41.
Health and Retirement Study. 2000. Health and Retirement Study 2000 Questionnaire. Ann Arbor, MI: University of Michigan.
Huang, F. Y., H. Chung, K. Kroenke, K. L. Delucchi & R. L. Spitzer (2006) Using the Patient Health Questionnaire-9 to measure depression among racially and ethnically diverse primary care patients. *J Gen Intern Med,* 21**,** 547-52.
Imbens, G. W. & J. D. Angrist (1994) Identification and Estimation of Local Average Treatment Effects. *Econometrica,* 62**,** 467-475.
International Expert Committee (2009) International Expert Committee report on the role of the A1C assay in the diagnosis of diabetes. *Diabetes Care,* 32**,** 1327-34.
Kalton, G. (1986) Handling wave nonresponse in panel surveys. *Journal of Official Statistics,* 2**,** 303-314.
King, G., E. Gakidou, K. Imai, J. Lakin, R. T. Moore, C. Nall, N. Ravishankar, M. Vargas, M. M. Tellez-Rojo, J. E. Hernandez Avila, M. Hernandez Avila & H. Hernandez Llamas (2009) Public policy for the poor? A randomised assessment of the Mexican universal health insurance programme. *The Lancet,* 373**,** 1447-1454.
Kroenke, K., R. L. Spitzer & J. B. Williams (2001) The PHQ-9: validity of a brief depression severity measure. *J Gen Intern Med,* 16**,** 606-13.
Kroenke, K., T. W. Strine, R. L. Spitzer, J. B. Williams, J. T. Berry & A. H. Mokdad (2009) The PHQ-8 as a measure of current depression in the general population. *J Affect Disord,* 114**,** 163-73.
McDade, T. W., S. Williams & J. J. Snodgrass (2007) What a drop can do: dried blood spots as a minimally invasive method for integrating biomarkers into population-based research. *Demography,* 44**,** 899-925.
McHorney, C. A., J. E. Ware & A. E. Raczek (1993) The MOS 36-Item Short-Form Health Survey (SF-36): II. Psychometric and clinical tests of validity in measuring physical and mental health constructs. *Med Care,* 31**,** 247-63.
Office for Oregon Health Policy and Research. 2009. Trends in Oregon's HealthCare Market and the Oregon Health Plan: A Report to the 75th Legislative Assembly.
Shim, R. S., P. Baltrus, J. Ye & G. Rust (2011) Prevalence, treatment, and control of depressive symptoms in the United States: results from the National Health and Nutrition Examination Survey (NHANES), 2005-2008. *J Am Board Fam Med,* 24**,** 33-8.
U.S. Preventive Services Task Force (2003) Screening for cervical cancer: recommendations and rationale. *Am J Nurs,* 103**,** 101-2, 105-6, 108-9.
--- (2008) Screening for colorectal cancer: U.S. Preventive Services Task Force recommendation statement. *Ann Intern Med,* 149**,** 627-37.
--- (2009) Screening for breast cancer: U.S. Preventive Services Task Force recommendation statement. *Ann Intern Med,* 151**,** 716-26, W-236.
Wallace, N., J. McConnell, C. Gallia & J. Smith (2008) 'How Effective are Co-Payments for Reducing Expenditures for Low-Income Adult Medicaid Beneficiaries? Experience from the Oregon Health Plan. *Health Services Research,* 43**,** 515–530.
Ware, J. E., M. Kosinki, J. Dewey & B. Gandek. 2001. How to Score and Interpret Single-Item Health Status Measures: A Manual for Users of the SF-8 Health Survey. Boston: QualityMetric Inc.
World Health Organization. 2011. World Health Organization Global Database on Body Mass Index.
Xu, K., D. B. Evans, K. Kawabata, R. Zeramdini, J. Klavus & C. J. Murray (2003) Household catastrophic health expenditure: a multicountry analysis. *Lancet,* 362**,** 111-7.