The Oregon Health Insurance Experiment:
The Moderating Role of Neighborhood Characteristics

Katherine Baicker
Julia Dennett

+

Analysis Plan
April 21, 2020

INTRODUCTION

In 2008, Oregon held a lottery to allocate a limited number of Medicaid slots to low-income uninsured adults on a waiting list. This lottery provides researchers with the opportunity to evaluate the causal effect of Medicaid coverage on a range of outcomes. The lottery, data collection, and fundamental empirical strategy are described in detail elsewhere (see nber.org/Oregon for previous analysis plans and publications).

Although policymakers often focus on the effect of expanding access to health care through health insurance on health outcomes, other factors beyond the health care system – social determinants of health – may play a role in shaping outcomes. The goal of the analysis described here is to evaluate how the effects of Medicaid coverage vary across people living in neighborhoods with different characteristics.

This document pre-specifies our planned analysis of the interaction between neighborhood characteristics and health insurance. It has been created before comparing how the effect of treatment on health outcomes varies based on neighborhood features in order to minimize issues of data mining and specification searching. It was, however, constructed after completion of analyses using the lottery to estimate the effects of treatment itself on those health outcomes.[1,2,3,4] In addition, it was constructed after exploring the non-causal associations between neighborhood characteristics and health outcomes through descriptive analyses[5], as well as treatment-control balance tests of neighborhood characteristics to assess the validity of our empirical strategy. The methods proposed here follow those of our prior quantitative analyses very closely; however, we now include new data on neighborhood characteristics and the interaction of those characteristics with randomly-assigned access to health insurance.

BACKGROUND

Medicaid coverage has been hypothesized to improve health or general well-being by reducing the price and providing additional income in kind, and thereby increasing the use of health-improving health care, or by alleviating stress or stigma.[1] The Oregon Health Insurance Experiment used a randomized Medicaid lottery in Oregon to assess the causal effects of Medicaid on a wide range of outcomes, including physical health, mental health, and health-related quality of life. Findings from the OHIE indicate that roughly two years after obtaining randomized access to Medicaid, coverage significantly increased use of health care across settings; reduced financial strain; and lowered the prevalence of depression for low-income adults.[3] Impacts on physical health outcomes were statistically insignificant, although Medicaid coverage improved diabetes management.

Beyond access to health care, other factors like neighborhood characteristics and environmental conditions may also affect health. An extensive literature documents an association between neighborhood characteristics and health outcomes[6,7,8]. Neighborhood attributes may affect health through numerous potential mechanisms. For example, certain characteristics may encourage good behaviors, such as eating healthy food and exercising, or reduce stress[9]. However, identifying causal effects poses an analytic challenge due to the endogeneity of residing in a specific neighborhood. As an example, individuals with lower socioeconomic status may have

poorer health outcomes and simultaneously reside in areas with fewer amenities. One notable exception is the randomized U.S. Department of Housing and Urban Development's Moving to Opportunity (MTO) intervention, which found that providing families living in high-poverty housing with the ability to move to neighborhoods with less poverty reduced long-run obesity and diabetes, and improved "subjective well-being".[10,11,12]

Furthermore, the health effects of health insurance coverage and neighborhood characteristics may interact; the environment in which the newly insured live may moderate the effects of acquiring Medicaid. For example, a physician's instructions to a newly-diagnosed diabetic patient to eat healthier food or walk more may be more likely to be followed in neighborhoods with grocery stores stocking affordable produce, walkable sidewalks, and parks. However, there are very few data sources that combine data on detailed neighborhood characteristics, health insurance status, and direct clinical assessment of health outcomes – and even fewer with the variation needed to assess causal effects.

This analysis will investigate the extent to which living in neighborhoods with different attributes moderates the effects of health insurance on health, taking advantage of the health insurance lottery. To undertake this study, we constructed a novel dataset by combining outcomes and randomized access to Medicaid coverage from the OHIE with new data on neighborhood characteristics in the Portland metropolitan area. A separate, descriptive analysis uses this dataset to explore the relationship between neighborhood characteristics and OHIE outcomes, and finds evidence that specific attributes were associated with health.[13] Specifically, individuals residing in areas with lower socioeconomic deprivation, more grocery stores, and more active living features demonstrated improved health outcomes compared to residents of other areas.

We introduce causality into this work by exploiting the random assignment of access to health insurance through the lottery to identify how the effects of Medicaid coverage on health vary based on neighborhood features. As we describe in greater detail below, we have exogenous variation in insurance assignment, but *not* in neighborhood characteristics. Thus, a causal interpretation of the interaction between neighborhood characteristics and insurance status requires the assumption that neighborhood characteristics are not affected by insurance status (about which our balance tests provide some information). We cannot, however, separate out the effect of those neighborhood characteristics themselves from any endogenous pre-lottery sorting into neighborhoods based on unobserved individual characteristics. As a result, any interaction effects we observe should be interpreted as the way that the effect of insurance varies across *people living in neighborhoods with different characteristics*, rather than necessarily being attributable to the neighborhood characteristics themselves.

METHODS

Data

In-Person Survey Responses

As part of the Oregon Health Insurance Experiment, an in-person survey was administered in the greater Portland metro area approximately two years after the lottery (between September 2009 and December 2010) to investigate individuals' health, finances, and health care utilization. The in-person surveys have been described in detail elsewhere.[3] We select a subset of outcomes from the in-person survey for use in the analysis, described below.

Neighborhood features may moderate the effects of Medicaid coverage on multiple dimensions of health. We organize our outcomes into the following *three categories*: physical health, mental health, and health-related quality of life. Within each of these categories, there are multiple potential outcome measures (e.g. within physical health, we measure blood pressure, cholesterol, etc.). We select one or two measures within each category as the primary outcomes. We also analyze additional secondary outcomes to test additional hypotheses.

Primary Outcomes

Physical health: Framingham Risk Score
A continuous measure of 10-year cardiovascular risk calculated for individuals aged 30 or older.[14] Risk scores were calculated separately for men and women on the basis of 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%. Individuals with any of these conditions are at increased risk of adverse cardiovascular outcomes.

Mental health: Depression Screen Result
Continuously measured total score on the PHQ-8 questionnaire. The Patient Health Questionnaire is a standard scale for measuring depression. The PHQ-8 asks about the frequency of eight depression symptoms: depressed mood, anhedonia, trouble sleeping, fatigue, problems eating, feeling like a failure, trouble concentrating, and moving or speaking more slowly or rapidly than usual.[15] 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 (ranging from 0 to 24).

Health-related quality of life:
Physical Quality of Life: Continuously measured score on the SF-8 physical health composite.
Mental Quality of Life: Continuously measured score on the SF-8 mental health composite.
SF-8 physical component (PCS) and mental component (MCS) index scores were computed using standard SF survey methodologies, creating scores ranging from 0-100, with higher scores representing better subjective health ratings.[16]

Secondary Outcomes

Physical health:
Systolic Blood Pressure: Continuously measured systolic blood pressure reading (mmHg).
Diastolic Blood Pressure: Continuously measured diastolic blood pressure reading (mmHg).
Total Cholesterol: Continuously measured total cholesterol level (mg/dL)
HDL Cholesterol: Continuously measured HDL cholesterol level (mg/dL)
Glycated Hemoglobin Level: Continuously measured HbA1c level.
Body Mass Index**:** Continuously measured body mass index (BMI) score computed from measured height and weight.

Mental health:
Depression: Binary variable indicating a score of 10 or above on the 8-question version of the Patient Health Questionnaire (PHQ-8).

Health-related quality of life:
Self-reported general health in the last 12 months is not poor or very poor
Binary variable reporting health status as "excellent", "very good", "good", or "fair" health as compared to "poor" or "very poor" health. This outcome is based on self-reported general health in the last twelve months measured using a categorical variable.
Happiness: Binary variable indicating being "very happy" or "pretty happy" as compared to "not too happy".

Respondent Characteristics

We also use sociodemographic characteristics obtained from the in-person surveys in the analysis:
Age: Continuous variable indicating respondent's age.
Gender: Binary variable indicating respondent's gender.
Household income: Continuous variable indicating approximate average household income (using mid-points and top coded at $50,000).
Education: Binary variables indicating the highest level of education completed by the respondent – less than high school, high school diploma or GED, post-high school but less than a 4-year degree, and 4-year degree or more.
Employment status: Binary variables indicating employment status at the time the biomarker measures were taken – not currently employed, employed less than 20 hours per week, employed 20 to 30 hours per week, and employed 30 hours or more per week.
Race/Ethnicity: Binary variables indicating self-reported racial/ethnic identification (White, Black, Hispanic, and "Non-White Other"). Respondents can report more than one race/ethnicity identity.
Number of household members on the lottery list: Binary variables indicating the number of household members on the OHIE lottery list.

Neighborhood Characteristics

We collected address- and tract-level data on a wide range of environmental and neighborhood characteristics to represent the neighborhoods in which Oregon Health Insurance Experiment interview respondents lived. We organized these neighborhood variables into *five domains* to capture different aspects of the neighborhood environment: socioeconomic deprivation, access to different kinds of food, park access and green space, features conducive to active living, and land use. For each domain, we describe potential channels through which attributes may moderate the effects of insurance to provide context, but we are not able to test whether those particular mechanisms are at work.

Sampling, data sources and collection, variable construction and selection, and the association between neighborhood variables and health outcomes are described in detail elsewhere.[17] For this analysis, we identify one neighborhood characteristic to act as a representative attribute for each domain, based on the degree to which it captures neighborhood features that are associated with health outcomes in prior analyses.[18] These neighborhood features are briefly described below.

Socioeconomic Deprivation

Socioeconomic deprivation may influence the degree to which Medicaid coverage affects health through various channels. For example, residents of deprived areas may experience increased stress levels or contend with fewer social or health care resources,[19] which may moderate how Medicaid coverage affects health. In addition, social norms regarding healthy and unhealthy behaviors may differ[20] and these peer effects may influence the impact of Medicaid coverage on health. Residents of socioeconomically deprived areas may also be less connected to the health care system and beneficial social policies. Previous results from the OHIE reported that Medicaid coverage generated a small increase in receipt of SNAP (food stamp) benefits, but not benefits from TANF, SSI, or SSDI, suggesting limited scope for this feedback loop.[21]

To generate a socioeconomic deprivation index, following the methodology of Messer et al (2006)[22], we used PCA to combine tract-level information on ethnicity (percent Hispanic), education (percent with no high school diploma or GED), employment (percent of employed males in management, professional, or related occupations; percent of employed females in management, professional, or related occupations), poverty (percent of households in poverty; percent of households earning less than $30,000 per year), housing (percent with renter or owner costs in excess of 50 percent of income), and crowding (percent of households with more than one occupant per room) to generate a deprivation score. All data came from the 2005-2009 5-year American Community Survey. Higher scores on the index indicate greater deprivation.

Food Access

Access to food may influence the degree to which Medicaid coverage affects health, since food access may impact diet and the prevalence of chronic conditions.[23] For example, if medical care providers suggest dietary changes to improve health (e.g. more vegetables, fewer calories for weight loss, etc.) or provide specific dietary recommendations (e.g. less sodium for hypertension,

less sugar for diabetes), access to healthy food in grocery stores may influence the effect of Medicaid coverage on health.

We collected address-level information on access to different kinds of food outlets, including grocery, fast food, and convenience store counts within various radii, by mapping each sample member's address using ArcGIS to such stores in 2009 INFOUSA business licensing data. We select count of grocery stores in a one-mile radius as the representative characteristic for this analysis.

Park Access and Green Space

Exposure to green space itself may directly relieve stress or reduce cortisol levels,[24] and may alter the effects of Medicaid coverage on health. For instance, parks may moderate the effects of Medicaid coverage by providing space for exercise and recreation,[25] making physicians' recommendations to increase physical activity more effective.

Address-level data on park access and green space features was obtained from Metro regional planning data mapped to each sample member's address using ArcGIS. We select acres of tree coverage within a quarter-mile radius as the representative characteristic for this analysis.

Active Living

Neighborhood features conducive to active living may moderate the effects of Medicaid coverage on health by providing opportunities for physical and social activity (e.g. walking or taking public transportation, visiting business destinations, etc.), which may make it easier to be comply with recommendations for more exercise.[26,27]

Measures of the active living profile, or walkability, of a neighborhood were organized into two categories: the capacity for active living and the quality of active living. Address-level data collected on capacity for active living, including street intersection, business, and population density, as well as number of transit stops and frequency of transit service within various radii, were obtained from various sources and mapped to each sample member's address using ArcGIS. Data on the quality of active living was gathered from 2011 street audits using the Active Neighborhood Checklist,[28] and included measures of environmental features and mobility characteristics. We used this street audit data to construct variables representing the share of street segments within a tract with a given characteristic. We select frequency of transit service (average stops per day in a half-mile radius), a capacity for active living variable obtained from 2009 Tri-Met transit systems data, as the representative characteristic for this analysis.

Land Use

Neighborhoods with particular kinds of land uses may moderate the effects of Medicaid coverage on health. For example, retail land use may provide access to business or social destinations, which may support recommended physical activity and improve mental health.[29]

Data was collected on land use from street audits using the Active Neighborhood Checklist and used to construct tract-level variables representing the share of street segments within a tract with different types of land use. We select the percent of street segments in a tract with <u>retail or small business land use</u> as the representative characteristic for this analysis. Retail or small business land use is defined as a small grocery, convenience store, pharmacy, home-based business (e.g. daycare, tax prep, salon), food establishment, entertainment, small commercial building, indoor fitness facility, big box store, mall, strip mall, or supermarket present.

Statistical Analysis

For a full description of the Analytic Specifications used, please see analysis plans at www.nber.org/oregon (In-Person Analysis Plan, NEJM Appendix). We describe new specifications used here.

Intent-to-Treat Effect of the Lottery (ITT)

Our analytic approach begins with an intent-to-treat (ITT) model comparing outcomes for all those who were selected in the lottery (the study treatment group) to all those who were on the list but not selected (the study control group), or the effect of winning the lottery. We evaluate outcomes across neighborhood characteristics by introducing an interaction between neighborhood features and winning the lottery. Specifically, we estimate the ITT by fitting the following OLS equation:

(1)
$$Y_{ihj} = \beta_0 + \beta_1 LOTTERY_h + \beta_2 NC_j + \beta_3 LOTTERY_h * NC_j + X_{ih}\beta_4 + V_{ih}\beta_5 + \varepsilon_{ihj}$$

Here i denotes an individual and that individual's address (some individuals share the same address); h denotes a household; and j denotes the geographic level at which the neighborhood attribute is measured – the census tract or an individual address (in which case, j is equivalent to i). LOTTERY is an indicator variable for whether or not household h was selected by the lottery and NC represents a neighborhood characteristic. The coefficient on the interaction term (β_3) is the main coefficient of interest, and identifies the way in which the effect between the treatment group (the lottery winners) and the control group (those not selected by the lottery) varies for those people living in a neighborhood with characteristic NC compared to those who don't, or how a neighborhood attribute moderates the effect 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_3 give an unbiased estimate of the relationship between winning the lottery and the outcome across neighborhood characteristics. 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.

We denote by $\boldsymbol{V}_{ih}$ 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_3 to give an unbiased estimate of the relationship between winning the lottery and the outcome across neighborhood features, however, as they are not related to treatment status. $\boldsymbol{V}_{ih}$ includes sociodemographic covariates that do not vary based on lottery selection; specifically, indicators for age, sex, race/ethnicity (white, Black/African American, Hispanic, other), income, education (less than high school, high school diploma or GED, some college), and employment status (unemployed, employed less than 20 hours per week, employed 20-30 hours per week).

Local Average Treatment Effect of Medicaid (LATE)

The intent-to-treat estimates from equation (1) provide an estimate of the differential effect of winning the lottery (i.e. winning the opportunity to apply for OHP Standard) across people living in neighborhoods with different characteristics. We are also interested in the influence of neighborhood characteristics on the impact of insurance coverage itself. We model this as follows:

(2)
$$Y_{ihj} = \pi_0 + \pi_1 INSURANCE_{ihj} + \pi_2 NC_j + \pi_3 INSURANCE_{ihj} * NC_j$$
$$+ \boldsymbol{X}_{ih}\pi_4 + \boldsymbol{V}_{ih}\pi_5 + \nu_{ihj}$$

Here INSURANCE is a measure of insurance coverage and all other variables are as defined in equation (1). Insurance coverage may be endogenously determined. We estimate equation (2) by two stage least squares (2SLS), using the following first stage equations:

(3)
$$INSURANCE_{ihj} = \delta_0 + \delta_1 LOTTERY_h + \delta_2 NC_j + \delta_3 LOTTERY_h * NC_j$$
$$+ \boldsymbol{X}_{ih}\delta_4 + \boldsymbol{V}_{ih}\delta_5 + \mu_{ihj}$$

(4)
$$INSURANCE_{ihj} * NC_j = \gamma_0 + \gamma_1 LOTTERY_h + \gamma_2 NC_j + \gamma_3 LOTTERY_h * NC_j$$
$$+ \boldsymbol{X}_{ih}\gamma_4 + \boldsymbol{V}_{ih}\gamma_5 + \eta_{ihj}$$

Here the excluded instruments are the variables LOTTERY and LOTTERY*NC.

We interpret the coefficients on insurance from instrumental variable estimation of equation (2) as the local average treatment effect of insurance, or LATE.[30] In other words, our estimates of π_1 and π_3 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. Furthermore, the coefficient of interest on the interaction term (π_3) identifies the way in which the effect of insurance varies for those people living in a neighborhood with characteristic NC compared with those who don't (assuming that the neighborhood characteristic does not change in response to lottery selection).

In all analyses we cluster the standard errors on the census tract since it is the greatest unit of

geography on which Neighborhood Characteristics are measured. All regressions are weighted to account for survey design, as described elsewhere.[31,32]

ANALYSIS

Study Population

This analysis links data collected from the Oregon Health Insurance Experiment in-person survey sample, which includes 12,229 respondents in the greater Portland Metropolitan Area, with primary and secondary data describing characteristics of the neighborhoods where they reside. Our analytical sample includes 8,413 individuals residing in 196 Census tracts for whom we have complete data on health outcomes, respondent demographic information, and neighborhood characteristics. **Table 1** summarizes demographic and neighborhood characteristics for this sample, including the treatment group selected in the lottery and the control group not selected. There are no statistically significant differences in the demographic or neighborhood characteristics between the treatment and control groups, consistent with prior analyses.

In **Table 2**, we assess the sensitivity of the effects of Medicaid coverage itself to alternative samples and specifications. Specifically, we compare estimates from the original in-person sample[3] to estimates from alternative samples and specifications, including using this paper's analytic sample, clustering standard errors by tract identifiers, including respondent characteristics as covariates, and all of these alterations together.   These results are consistent with prior analyses.

Insurance Coverage

Table 3A reports the effect of being selected in the lottery on Medicaid coverage for both the original in-person sample and our new analytic sample. In our analysis, we define Medicaid coverage as being covered at any point between March 10, 2008 (the date of the first lottery notifications) and the sampling date. This definition of Medicaid includes both the lotteried Medicaid program (OHP Standard) and other non-lotteried Medicaid programs. The results indicate that winning the lottery increased the probability of being covered by Medicaid at any point during the study period by 22.7 percentage points in our analytic sample, compared to 24.1 percentage points among all in-person survey responders. The lottery affected coverage through increasing enrollment in OHP Standard.

Table 3B reports the estimates from our first stages, modeled by equations (3) [shown in the first column] and equation (4) [shown in subsequent columns], estimated separately for each neighborhood characteristic. There does not appear to be a statistically significant differential impact from winning the lottery on Medicaid coverage across neighborhood characteristics.

Causal Interpretation of the Influence of Neighborhoods

Using our novel data that links the health outcomes of low-income adults and their randomized health insurance status from the OHIE[33] with characteristics of their residential neighborhoods in Portland, Oregon, we can test numerous hypotheses regarding the extent to which a specific

neighborhood characteristic influences the effect of Medicaid coverage on health. The strength of this approach is that Medicaid coverage is randomly assigned.

However, neighborhood characteristics are not randomly assigned. Even though the selected neighborhood characteristics are similar for treatment and control groups, the neighborhood characteristics we observe may proxy for a number of other unobserved characteristics of the neighborhood or its residents. As a result, we are unable to test if the neighborhood characteristic itself moderates the effects of Medicaid coverage, or if factors correlated with the area's traits drive the results, such as the characteristics of individuals who chose to live in those neighborhoods or other unobserved neighborhood features.

We will therefore learn whether the effects of Medicaid are different for people living in areas with specific characteristics, rather than whether a specific feature of the neighborhood is driving the differential effects. For example, Medicaid may have a greater effect on health for people living in neighborhoods with parks, but this will not necessarily tell us that adding a park to an existing neighborhood would increase the effect of Medicaid. These estimates will nevertheless potentially be quite informative for policymakers in targeting resources and approaches to those neighborhoods where they can be most effective.

Planned Analysis of Health Outcomes

There are many pathways by which neighborhood characteristics may moderate the effects of health insurance coverage, and we have incredibly rich data sources. We have selected a parsimonious set of analyses based on the existing literatures exploring the influence of neighborhoods on health and the impact of Medicaid coverage on health. [34,35,3]

Table 4 presents the main results of our analyses: the extent to which a neighborhood characteristic moderates the effects of lottery selection (ITT) and Medicaid coverage (LATE) on the primary health outcomes.

We also plan to conduct secondary analyses informed by specific hypotheses. First, we test the influence of selected neighborhood characteristics on the effects of lottery selection and Medicaid coverage on the other physical health outcomes (systolic blood pressure, diastolic blood pressure, total cholesterol, HDL cholesterol, glycated hemoglobin level, and body mass index). **Table 5** reports the extent to which the count of grocery stores in a one-mile radius and the frequency of transit service (average stops per day in a half-mile radius) moderate the effects of lottery selection (ITT) and Medicaid coverage (LATE). These neighborhood characteristics were selected because of the previously observed[36] relationship between diet, active living, and physical health.

Furthermore, we test the influence of selected neighborhood characteristics on the effects of lottery selection and Medicaid coverage on the other mental health outcomes (positive depression screening) and the other health-related quality of life outcomes (self-reported general health and happiness). **Table 6** reports the extent to which the socioeconomic deprivation index and acres of tree coverage within a quarter-mile radius moderate the effects of lottery selection (ITT) and Medicaid coverage (LATE). These neighborhood characteristics were chosen because

of the potential relationship between environmental stress, green spaces, mental health, and subjective well-being.

Table 7 explores potential <u>heterogeneity</u> in the moderating effects of neighborhood characteristics. We focus on two subgroups of individuals and report the extent to which each of the selected neighborhood characteristics influences the effects of lottery selection (ITT) and Medicaid coverage (LATE).

First, we consider the summary physical health outcome (the Framingham Risk Score) limited to those with a "high risk" diagnosis, defined as a pre-randomization diagnosis of diabetes, hypertension, hypercholesterolemia, myocardial infarction, or congestive heart failure before the lottery (i.e., March 10th, 2008). In addition, we consider the summary mental health outcome (depression screening result) limited to those with a pre-randomization diagnosis of depression or anxiety.

REFERENCES

[1] Finkelstein A, Taubman S, Wright B, et al. The Oregon Health Insurance Experiment: Evidence from the First Year. Quarterly Journal of Economics 2012;127:1057-106.

[2] Allen H, Wright BJ, Baicker K. New medicaid enrollees in Oregon report health care successes and challenges. Health affairs (Project Hope) 2014;33:292-9.

[3] Baicker K, Taubman SL, Allen HL, et al. The Oregon Experiment -- Effects of Medicaid on Clinical Outcomes. New England Journal of Medicine 2013;368:1713-22.

[4] Taubman SL, Allen HL, Wright BJ, Baicker K, Finkelstein AN. Medicaid increases emergencydepartment use: evidence from Oregon's Health Insurance Experiment. Science 2014;343:263-8.

[5] Dennett J, Wright BJ, Cohen-Cline H, Baicker K. Connecting neighborhood characteristics and clinical health outcomes: novel data from the Oregon Health Insurance Experiment. 2020. Available at SSRN: https://ssrn.com/abstract=3582108

[6] Pickett KE, Pearl M. Multilevel analyses of neighbourhood socioeconomic context and health outcomes: a critical review. *Journal of Epidemiology and Community Health*. 2001; 55:111-122.

[7] Duncan DT, Kawachi I. "Neighborhoods and Health" Second edition. Oxford University Press. 2018.

[8] Kawachi I, Berkman LF. "Neighborhoods and Health" Oxford University Press. 2003.

[9] Duncan DT, Kawachi I. "Neighborhoods and Health" Second edition. Oxford University Press. 2018. p. 222 and p. 247

[10] Ludwig J, Sanbonmatsu L, Gennetian L, et al. Neighborhoods, obesity, and diabetes--a randomized social experiment. *The New England Journal of Medicine*. 2011; 365:1509-1519. doi: 10.1056/NEJMsa1103216 [doi].

[11] Ludwig J, Duncan GJ, Gennetian LA, et al. Neighborhood effects on the long-term well-being of low-income adults. *Science*. 2012; 337(6101):1505-1510.

[12] Ludwig J, Duncan GJ, Gennetian LA, et al. Long-term neighborhood effects on low-income families: evidence from Moving to Opportunity. *American Economic Review: Papers & Proceedings*. 2013; 103(3):226-231.

[13] Dennett J, Wright BJ, Cohen-Cline H, Baicker K. Connecting neighborhood characteristics and clinical health outcomes: novel data from the Oregon Health Insurance Experiment. 2020. Available at SSRN: https://ssrn.com/abstract=3582108

[14] D'Agostino R, Vasan R, Pencina M, Wolf P, Cobain M, Massaro J, Kannel W. General Cardiovascular Risk Profile for Use in Primary Care: The Framingham Heart Study. Circulation, 2008 Feb, 117:6, pp 743-753.

[15] Kroenke K. The PHQ-9: Validity of a Brief Depression Severity Measure. Journal of General Internal Medicine, 2001 Sep; 16(9), 606-13.

[16] Ware J, Kosinsky M, Dewey J, Gandek B. How to Score and Interpret Single Item Health Status Measures: A Manual for Users of the SF-8 Health Survey. Quality Metric, Feb 2001.

[17] Dennett J, Wright BJ, Cohen-Cline H, Baicker K. Connecting neighborhood characteristics and clinical health outcomes: novel data from the Oregon Health Insurance Experiment. 2020. Available at SSRN: https://ssrn.com/abstract=3582108

[18] Dennett J, Wright BJ, Cohen-Cline H, Baicker K. Connecting neighborhood characteristics and clinical health outcomes: novel data from the Oregon Health Insurance Experiment. 2020. Available at SSRN: https://ssrn.com/abstract=3582108

[19] Pickett KE, Pearl M. Multilevel analyses of neighbourhood socioeconomic context and health outcomes: a critical review. *Journal of Epidemiology and Community Health*. 2001; 55:111-122.

[20] Pickett KE, Pearl M. Multilevel analyses of neighbourhood socioeconomic context and health outcomes: a critical review. *Journal of Epidemiology and Community Health*. 2001; 55:111-122.

[21] Baicker K, Finkelstein A, Song J, Taubman S. The Impact of Medicaid on Labor Market Activity and Program Participation: Evidence from the Oregon Health Insurance Experiment. *American Economic Review: Papers & Proceedings* 2014, 104(5): 322–328.

[22] Messer LC, Laraia BA, Kaufman JS, et al. The development of a standardized neighborhood deprivation index. *Journal of Urban Health: Bulletin of the New York Academy of Medicine.* 2006; (83)6.

[23] Duncan DT, Kawachi I. "Neighborhoods and Health" Second edition. Oxford University Press. 2018. p. 247

[24] Duncan DT, Kawachi I. "Neighborhoods and Health" Second edition. Oxford University Press. 2018. p. 222

[25] Duncan DT, Kawachi I. "Neighborhoods and Health" Second edition. Oxford University Press. 2018. p. 222

[26] Duncan DT, Kawachi I. "Neighborhoods and Health" Second edition. Oxford University Press. 2018. p. 222

[27] Duncan DT, Kawachi I. "Neighborhoods and Health" Second edition. Oxford University Press. 2018. p. 222

[28] Hoehner C, Ivy A, Ramirez L, Handy S, Brownson R. Active Neighborhood Checklist; A User Friendly and Reliable Tool for Assessing Activity Friendliness. American Journal of Health Promotion. 2007. Jul-Aug, 21:6, 534-7.

[29] Duncan DT, Kawachi I. "Neighborhoods and Health" Second edition. Oxford University Press. 2018. p. 222

[30] Finkelstein A, Taubman S, Wright B, et al. The Oregon Health Insurance Experiment: Evidence from the First Year. Quarterly Journal of Economics 2012;127:1057-106.

[31] Baicker K, Taubman S, Allen H, et al. 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 2012.

[32] Finkelstein A, Taubman S, Allen H, et al. The short-run impact of extending public health insurance to low income adults: evidence from the first year of The Oregon Medicaid Experiment. Analysis Plan; Archived on December 1, 2010; Available at http://www.nber.org/oregon.

[33] Baicker K, Taubman SL, Allen HL, et al. The Oregon Experiment -- Effects of Medicaid on Clinical Outcomes. New England Journal of Medicine 2013;368:1713-22.

[34] Pickett KE, Pearl M. Multilevel analyses of neighbourhood socioeconomic context and health outcomes: a critical review. *Journal of Epidemiology and Community Health*. 2001; 55:111-122.

[35] Duncan DT, Kawachi I. "Neighborhoods and Health" Second edition. Oxford University Press. 2018.

[36] Dennett J, Wright BJ, Cohen-Cline H, Baicker K. Connecting neighborhood characteristics and clinical health outcomes: novel data from the Oregon Health Insurance Experiment. 2020. Available at SSRN: https://ssrn.com/abstract=3582108

<div align="center">

Table 1: Sample Characteristics

</div>

	Controls (1)	Lottery Winners (2)	P-Value (3)
N	4,029	4,384	
Demographic Characteristics			
Age (Years)	40.4	40.6	0.44
Female (%)	56.7	56.2	0.62
Race			
White (%)	66.2	68.0	0.10
Black (%)	12.3	12.3	0.98
Hispanic (%)	18.0	16.7	0.21
Non-White Other (%)	14.7	14.9	0.82
Household Income	18128.5	18090.5	0.91
Education			
Less than high school (%)	21.4	20.5	0.39
High school diploma or GED (%)	47.0	46.1	0.49
Post-high school (%)	21.6	23.0	0.18
Four year degree or more (%)	10.0	10.4	0.62
Employment Status			
Not currently employed (%)	50.5	51.3	0.50
Employed less than 20 hours a week (%)	10.0	9.6	0.64
Employed 20 to 30 hours a week (%)	11.1	11.6	0.56
Employed more than 30 hours per week (%)	28.4	27.6	0.43
F-Statistic for Above Variables		0.611	
P-Value		0.843	
Neighborhood Characteristics			
Socioeconomic Deprivation			
Deprivation Score	0.7	0.7	0.34
Food Access			
Count of grocery stores within a one mile radius of the dwelling	8.2	8.1	0.53
Park Access and Green Space			
Acres of mature tree cover within a quarter mile neighborhood of the dwelling	33.5	33.7	0.51
Active Living			
Frequency of service (avg. stops per day) within a half mile neighborhood	293.5	288.5	0.58
Land Use			
Percent of street segments within a tract with retail land use present	18.4	17.5	0.08
F-Statistic for Above Variables		1.009	
P-Value		0.413	

Notes: The first column shows the weighted mean value for control individuals who entered the lottery but were not selected, the second column reports the regression-adjusted weighted mean value for those randomly selected by the lottery to be able to apply for Medicaid coverage, and the third column shows the p-value of difference between the means. The sample is weighted using the in-person survey weights and standard errors are clustered by census tract. The final rows report the pooled F statistics and p-values from testing treatment-control balance on the above variables jointly.

Table 2. Means and the Effect of Medicaid Coverage on Health

	Sample with Health Outcomes Data (n=12,229)		(1) Analytic Sample with Neighborhood Data (n=8,413)		(2) Analytic Sample with Standard Errors Clustered by Tract (n=8,413)		(3) Respondent Characteristics Included as Covariates (n=12,229)		(1), (2), and (3) (n=8,413)	
	Mean Value in Control Group	Effect of Medicaid Coverage	Mean Value in Control Group	Effect of Medicaid Coverage	Mean Value in Control Group	Effect of Medicaid Coverage	Mean Value in Control Group	Effect of Medicaid Coverage	Mean Value in Control Group	Effect of Medicaid Coverage
Physical Health										
Framingham risk score (%)^	8.2	-0.21	8.1	0.43	8.1	0.43	8.2	-0.02	8.1	0.33
Systolic bp, avg	119.3	-0.51	119.0	0.32	119.0	0.32	119.3	-0.12	119.0	0.27
Diastolic bp, avg	76.0	-0.80	76.0	-0.57	76.0	-0.57	76.0	-0.86	76.0	-0.96
Total cholesterol	204.1	2.20	204.9	-1.36	204.9	-1.36	204.1	1.63	204.9	-2.07
HDL cholesterol	47.6	0.83	47.7	-0.53	47.7	-0.53	47.6	0.66	47.7	-0.69
Hemoglobin A1C	5.3	0.01	5.3	0.04	5.3	0.04	5.3	0.00	5.3	0.03
BMI	29.8	0.21	29.8	0.75	29.8	0.75	29.8	0.27	29.8	0.74
Mental Health										
PHQ total severity score	7.0	-1.18**	7.1	-1.50**	7.1	-1.50**	7.0	-1.11**	7.1	-1.67***
Screened positive for depression (phq8>=10) (%)	30.0	-9.15**	31.0	-12.07**	31.0	-12.07**	30.0	-8.54**	31.0	-13.06***
Health-Related Quality of Life										
PCS-8, physical subscale	45.5	1.20	45.5	1.06	45.5	1.06	45.5	1.23	45.5	1.42
MCS-8, mental subscale	44.4	1.95**	44.3	3.02**	44.3	3.02**	44.4	1.80*	44.3	3.30***
Health in the last 12 mos is not poor/very poor (%)	85.8	4.50	85.7	4.77	85.7	4.77	85.8	4.46	85.7	5.40
Happiness (%)	74.9	1.18	74.6	4.45	74.6	4.45	74.9	0.92	74.6	4.91

Notes: The first column reports the weighted mean of the dependent variable in the control sample and the second column reports the local-average-treatment-effect for insurance coverage as estimated by instrumental variable regression. This table compares the results from the in-person survey analyses (see Baicker et al., 2013 for details and specification) with different samples and specifications relevant to analyzing the analytic sample, which includes neighborhood characteristics. (1) Applies the in-person survey analyses specification to the analytic sample; (2) Applies the in-person survey analyses specification to the analytic sample, and clusters standard errors by census tract instead of household; (3) Applies the analytic sample covariates to the in-person survey analyses; and (4) combines previous specification changes to report results for the analytic sample with standard errors clustered by census tract and controlling for analytic sample covariates. All regressions include indicators for number of household members on the lottery list. All analyses are weighted using survey weights. * p<0.10, ** p<0.05, *** p<0.01. ^The sample sizes for the Framingham risk score, which is calculated for persons 30 years of age or older, are 9,525 in the in-person survey analyses and 6,545 in the analytic sample.

Table 3A: Insurance Coverage (First Stage Estimates)

	Sample with Health Outcomes Data (n=12,229)		Analytic Sample with Neighborhood Data (n=8,413)	
	Control Mean (1)	Estimated FS (2)	Control Mean (1)	Estimated FS (2)
Ever on Medicaid	18.5	24.14 (0.90)	19.3	22.72 (1.08)
Ever on OHP Standard	3.3	26.49 (0.70)	3.3	25.52 (0.94)
Number of Months on Medicaid	2.6	4.16 (0.16)	2.7	3.96 (0.21)
On Medicaid at the End of the Period	13.3	11.35 (0.79)	14.2	10.67 (0.99)

Notes: This table compares the first stage estimates from the in-person survey analyses (see Baicker et al., 2013 for details and specification) with the first stage estimates from the analytic sample, which includes neighborhood characteristics. The first column reports the control mean for alternate definitions of "MEDICAID," and the second column reports the coefficient (with standard error in parentheses) on LOTTERY from estimating the first-stage using the specified definition of "MEDICAID." All regressions include indicators for number of household members on the lottery list, and are weighted using survey weights. Analyses using the analytic sample also controls for respondent characteristics (age, gender, race/ethnicity, household income, education, and employment status). Analyses using the in-person survey data clusters standard errors by household, and analyses using the analytic sample clusters standard errors by census tract. In all our analyses of the local-average-treatment effect of Medicaid, we use the definition in the first row: "On Medicaid at any point in the study period."

Table 3B: Insurance Coverage and Neighborhood Characteristics (First Stage Estimates)

	Medicaid Coverage	Medicaid Coverage *SDI	Medicaid Coverage *Count of grocery stores	Medicaid Coverage *Acres of tree coverage	Medicaid Coverage *Frequency of transit service	Medicaid Coverage *Retail or small business land use
Lottery Selection	23.32*** (1.157)	-3.625 (2.354)				
Socioeconomic Deprivation Index (SDI)	0.831** (0.332)	21.21*** (1.563)				
Lottery Selection*SDI	-0.873* (0.520)	22.82*** (1.828)				
Lottery Selection	21.40*** (1.858)		-12.98 (23.09)			
Count of grocery stores in a one-mile radius	0.0517 (0.133)		21.95*** (2.947)			
Lottery Selection*Count of grocery stores	0.165 (0.180)		24.79*** (3.497)			
Lottery Selection	24.66*** (2.890)			17.33 (92.08)		
Acres of tree coverage within a quarter-mile radius	-0.111** (0.0563)			13.22*** (2.425)		
Lottery Selection*Acres of tree coverage	-0.0569 (0.0782)			22.09*** (3.072)		
Lottery Selection	21.76*** (1.248)				-1287.4** (629.8)	
Frequency of transit service (average stops per day in a half-mile radius)	0.000439				22.59***	

Lottery Selection*Frequency of transit service	(0.00140)	(2.720)
	0.00338*	29.22***
	(0.00188)	(2.979)
Lottery Selection	22.31***	-0.608
	(1.622)	(0.397)
Retail or small business land use (percent of street segments within a census tract)	1.623	21.11***
	(4.298)	(2.988)
Lottery Selection*Retail or small business land use	2.494	26.30***
	(6.726)	(3.128)

Notes: This table reports the first stage estimates from analyses using the analytic sample. All regressions include indicators for number of household members on the lottery list, are weighted using survey weights, and control for respondent characteristics (age, gender, race/ethnicity, household income, education, and employment status). Standard errors are in parentheses and are clustered by census tract. Sample sizes are N=8,413. * $p<0.10$, ** $p<0.05$, *** $p<0.01$.

Table 4: Intent-to-Treat Effect of Lottery and Local Average Treatment Effect of Medicaid on Primary Health Outcomes

	Physical Health		Mental Health		Health-Related Quality of Life			
	Framingham Risk Score		Depression Screen Result		Physical		Mental	
	ITT	LATE	ITT	LATE	ITT	LATE	ITT	LATE
Socioeconomic Deprivation								
Lottery Selection								
Socioeconomic Deprivation Index (SDI)								
Lottery Selection*SDI								
Medicaid Coverage								
Socioeconomic Deprivation Index (SDI)								
Medicaid Coverage*SDI								
Food Access								
Lottery Selection								
Count of grocery stores in a one-mile radius								
Lottery Selection*Count of grocery stores								
Medicaid Coverage								

Count of grocery stores in a one-mile radius

Medicaid Coverage*Count of grocery stores

Park Access and Green Space

Lottery Selection

Acres of tree coverage within a quarter-mile radius

Lottery Selection*Acres of tree coverage

Medicaid Coverage

Acres of tree coverage within a quarter-mile radius

Medicaid Coverage*Acres of tree coverage

Active Living

Lottery Selection

Frequency of transit service
(average stops per day in a half-mile radius)

Lottery Selection*Frequency of transit service

Medicaid Coverage

Frequency of transit service
(average stops per day in a half-mile radius)

Medicaid Coverage*Frequency of transit service

Land Use

Lottery Selection

Retail or small business land use
(percent of street segments within a census tract)

Lottery Selection*Retail or small business land use

Medicaid Coverage

Retail or small business land use
(percent of street segments within a census tract)

Medicaid Coverage*Retail or small business land use

Notes: All regressions include indicators for the number of household members on the lottery list and controls for respondent characteristics (age, gender, race/ethnicity, household income, education, and employment status). Standard errors are in parentheses and are clustered by census tract. All analyses are weighted using survey weights. Sample sizes are N=8,413, except for the Framingham risk score (N=6,542; three observations aged 30 or older in the analytic sample are missing this outcome). * p<0.10, ** p<0.05, *** p<0.01.

Table 5: Intent-to-Treat Effect of Lottery and Local Average Treatment Effect of Medicaid on Secondary Physical Health Outcomes

	Systolic bp, avg		Diastolic bp, avg		Total cholesterol		HDL cholesterol		Hemoglobin A1C		BMI	
	ITT	LATE	ITT	LATE	ITT	LATE	ITT	LATE	ITT	LATE	ITT	LATE
Food Access												
Lottery Selection												
Count of grocery stores in a one-mile radius												
Lottery Selection*Count of grocery stores												
Medicaid Coverage												
Count of grocery stores in a one-mile radius												
Medicaid Coverage*Count of grocery stores												
Active Living												
Lottery Selection												
Frequency of transit service (average stops per day in a half-mile radius)												
Lottery Selection*Frequency of transit service												
Medicaid Coverage												
Frequency of transit service (average stops per day in a half-mile radius)												
Medicaid Coverage*Frequency of transit service												

Notes: All regressions include indicators for the number of household members on the lottery list and controls for respondent characteristics (age, gender, race/ethnicity, household income, education, and employment status). Standard errors are in parentheses and are clustered by census tract. All analyses are weighted using survey weights. Sample sizes are N=8,413. * p<0.10, ** p<0.05, *** p<0.01.

Table 6: Intent-to-Treat Effect of Lottery and Local Average Treatment Effect of Medicaid on Secondary Mental and Health-Related Quality of Life Outcomes

| | Mental Health | | Health-Related Quality of Life | | Self-Reported Happiness | |
	Screened positive for depression (phq8>=10)		Health in the last 12 mos is not poor/very poor		(Very happy or pretty happy)	
	ITT	LATE	ITT	LATE	ITT	LATE
Socioeconomic Deprivation						
Lottery Selection						
Socioeconomic Deprivation Index (SDI)						
Lottery Selection*SDI						
Medicaid Coverage						
Socioeconomic Deprivation Index (SDI)						
Medicaid Coverage*SDI						
Park Access and Green Space						
Lottery Selection						
Acres of tree coverage within a quarter-mile radius						
Lottery Selection*Acres of tree coverage						
Medicaid Coverage						
Acres of tree coverage within a quarter-mile radius						
Medicaid Coverage*Acres of tree coverage						

Notes: All regressions include indicators for the number of household members on the lottery list and controls for respondent characteristics (age, gender, race/ethnicity, household income, education, and employment status). Standard errors are in parentheses and are clustered by census tract. All analyses are weighted using survey weights. Sample sizes are N=8,413. * p<0.10, ** p<0.05, *** p<0.01.

Table 7: Intent-to-Treat Effect of Lottery and Local Average Treatment Effect of Medicaid for Subgroups

	Physical Health in "High Risk" Diagnosis Subgroup Framingham Risk Score		Mental Health in Depression or Anxiety Diagnosis Subgroup Depression Screen Result	
	ITT	LATE	ITT	LATE
Socioeconomic Deprivation				
Lottery Selection				
Socioeconomic Deprivation Index (SDI)				
Lottery Selection*SDI				
Medicaid Coverage				
Socioeconomic Deprivation Index (SDI)				
Medicaid Coverage*SDI				
Food Access				
Lottery Selection				
Count of grocery stores in a one-mile radius				
Lottery Selection*Count of grocery stores				
Medicaid Coverage				
Count of grocery stores in a one-mile radius				
Medicaid Coverage*Count of grocery stores				
Park Access and Green Space				
Lottery Selection				

Acres of tree coverage within a quarter-mile radius

Lottery Selection*Acres of tree coverage

Medicaid Coverage

Acres of tree coverage within a quarter-mile radius

Medicaid Coverage*Acres of tree coverage

Active Living

Lottery Selection

Frequency of transit service
(average stops per day in a half-mile radius)

Lottery Selection*Frequency of transit service

Medicaid Coverage

Frequency of transit service
(average stops per day in a half-mile radius)

Medicaid Coverage*Frequency of transit service

Land Use

Lottery Selection

Retail or small business land use
(percent of street segments within a census tract)

Lottery Selection*Retail or small business land use

Medicaid Coverage

Retail or small business land use
(percent of street segments within a census tract)

Medicaid Coverage*Retail or small business land use

See Table 4 notes. "High risk" 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=2,095; 138 "high risk" observations are aged less than 30 and one observation aged 30 or older are missing a Framingham risk score in the analytic sample). Depression subgroup is limited to those with a pre-existing diagnosis of depression or anxiety (N=2,797). * p<0.10, ** p<0.05, *** p<0.01.