CESifo WORKING PAPERS

9664
2022

March 2022

Who Increases Emergency Department Use? New Insights from the Oregon Health Insurance Experiment

Augustine Denteh, Helge Liebert

CESifo

Impressum:

CESifo Working Papers
ISSN 2364-1428 (electronic version)
Publisher and distributor: Munich Society for the Promotion of Economic Research - CESifo GmbH
The international platform of Ludwigs-Maximilians University's Center for Economic Studies and the ifo Institute
Poschingerstr. 5, 81679 Munich, Germany
Telephone +49 (0)89 2180-2740, Telefax +49 (0)89 2180-17845, email office@cesifo.de
Editor: Clemens Fuest
https://www.cesifo.org/en/wp
An electronic version of the paper may be downloaded
· from the SSRN website: www.SSRN.com
· from the RePEc website: www.RePEc.org
· from the CESifo website: https://www.cesifo.org/en/wp

CESifo *Working Paper No. 9664*

Who Increases Emergency Department Use? New Insights from the Oregon Health Insurance Experiment

Abstract

We provide new insights regarding the finding that Medicaid increased emergency department (ED) use from the Oregon experiment. We find meaningful heterogeneous impacts of Medicaid on ED use using causal machine learning methods. The treatment effect distribution is widely dispersed, and the average effect is not representative of most individualized treatment effects. A small group—about 14% of participants—in the right tail of the distribution drives the overall effect. We identify priority groups with economically significant increases in ED usage based on demographics and prior utilization. Intensive margin effects are an important driver of increases in ED utilization.

JEL-Codes: H750, I130, I380.

Keywords: Medicaid, ED use, effect heterogeneity, causal machine learning, optimal policy.

*Augustine Denteh**
Department of Economics
Tulane University
6823 St. Charles Avenue
USA – New Orleans, LA 70118
adenteh@tulane.edu

Helge Liebert
Department of Economics
University of Zurich
Schönberggasse 1
Switzerland – 8001 Zurich
helge.liebert@econ.uzh.ch

*corresponding author

March 28, 2022
We thank Anirban Basu, Paul Goldsmith-Pinkham, Timothy Layton, Jacob Wallace, Sherri Rose, Mary Olson, Makayla Lavender, Prithvijit Mukherjee, Samuel Asare, Nicholas A. Wright, Anthony Strittmatter and seminar participants at Tulane University for helpful comments.

1. Introduction

The finding that Medicaid increased emergency department (ED) utilization in the 2008 Oregon Health Experiment drew widespread national attention (Taubman et al. 2014). Conventional economic theory suggests that insurance coverage increases access to primary care services and reduces the likelihood of needing ED care. One likely explanation for why Medicaid increased ED use in Oregon is newly insured people face a lower out-of-pocket cost (Sommers & Simon 2017). However, the positive effect of Medicaid on ED use in Oregon contradicted previous evidence from the 2006 Massachusetts health insurance reform where coverage decreased ED utilization (Chen et al. 2011, Miller 2012). Subsequent quasi-experimental studies also find that state Medicaid expansions in Kentucky and Arkansas reduced ED visits (Sommers et al. 2016). Despite the mixed evidence, moral hazard in ED utilization remains a crucial consideration in health insurance expansions because of the continued rise in ED visits, the declining number of emergency departments, and the effects of ED crowding on health (Tang et al. 2010, Moore & Liang 2020, Woodworth 2020).

This paper provides new insights regarding how Medicaid affects ED utilization when we go beyond the average effect to account for the characteristics of the expansion population. We do so by estimating the heterogeneous impacts of Medicaid on ED use, allowing us to characterize those driving the effects in the Oregon experiment. Oregon's coverage expansion to a population not typically covered by Medicaid motivates our investigation of its heterogeneous effects. Oregon randomly assigned the opportunity to apply for spots in its Oregon Health Plan Standard (OHP Standard) in 2008. Unlike OHP Plus, which serves Oregon's typical Medicaid population, OHP Standard was newly offered to a group of uninsured adults who were categorically ineligible for Medicaid under federal guidelines (Allen et al. 2013). Oregon's Medicaid expansion increased utilization for most types of ED visits, including those not requiring emergent or immediate care, for up to two years (Finkelstein et al. 2016). This study examines how representative these *average* estimates are of the *individualized* effects of Medicaid on ED use. We isolate the risk factors that matter most for increases in ED use and discuss how we can use such knowledge to inform policy.

This paper makes three salient contributions. First, we nonparametrically estimate the heterogeneous treatment effects of Medicaid on ED utilization using causal machine learning methods. We estimate various types of heterogeneous effects, including *individualized* treatment effects that condition on individual covariate values and higher-level *group* average treatment effects using generalized random forests (Athey et al. 2019). Previous studies report a small number of subgroup effects for a subset of people in the Oregon experiment who completed a follow-up survey (Taubman et al. 2014). Subsample analysis can miss interesting patterns of heterogeneity due to non-linear interactions between multiple covariates. In addition to the need to pre-specify the covariates, subsample heterogeneity analysis is prone to multiple hypothesis testing problems (e.g. List et al. 2019, Young 2019). Our results are consistent with the fact that the average treatment effect does not represent the individualized treatment effect in the population when there is substantial effect heterogeneity (Heckman et al. 1997, Angrist 2004, Deaton 2010).

Second, we uncover the risk factors for increasing ED use upon Medicaid coverage, opening the doorway for identifying priority groups for targeted policy interventions. Kowalski (2018) studies a related problem of reconciling the contradictory results in the Oregon experiment and the Massachusetts reform. She attributes the discrepancy to different local average treatment effects (LATE) arising from the same marginal treatment effect (MTE) function. She extrapolates an MTE function for Oregon to Massachusetts and finds that such re-weighting reconciles the opposite findings. We do not attempt to reconcile Oregon's ED finding with Massachusetts's since those differences could stem from multiple factors. Instead, we directly link treatment effect heterogeneity in Oregon to observed covariates in a data-driven manner without making additional assumptions beyond those needed to recover the LATE. Doing so sheds light on the drivers of Medicaid's positive average impact on ED use.

Third, we discuss the potential to leverage the heterogeneous effects of Medicaid to propose simple rules to identify people for either targeted educational interventions to curb undesirable ED use or for future coverage expansion. While it is essential to document heterogeneous effects, policymakers might be unclear how to use such evidence. We use the heterogeneous treatment effects to estimate assignment rules based on recently developed policy learning algorithms (Athey & Wager 2021, Kitagawa & Tetenov 2018). When faced with the decision to expand coverage, we show how an administrator can assign the opportunity to apply for the program to achieve the desired objective, such as minimizing undesirable ED use. Specifically, we estimate shallow decision trees for the intent-to-treat effect of winning the lottery to minimize unnecessary ED use. One advantage of such assignment rules is that fairness or legal considerations can be accounted for by excluding variables prohibited by discriminatory laws from the decision rule estimation.

Several broad findings emerge from our analysis. First, we find substantial heterogeneity in the impact of Medicaid coverage on ED visits. Individualized treatment effects are widely dispersed, with most people experiencing null effects or reductions in ED use. For ED use on the extensive margin, a small group of people in the right tail of the distribution—about 14% of recipients—with statistically significant increases in ED visits drive the positive average impact of Medicaid. We also document meaningful heterogeneity on the intensive margin and find similar patterns of heterogeneity for various types of ED visits. Moreover, we find that null effects for certain types of ED use hide significant countervailing forces—reductions in ED use by some people are offset by increases in utilization by others.

Second, when we aggregate the individualized treatment effects, we find weaker extensive margin results in magnitude and statistical significance. The aggregate estimate indicates that Medicaid increased the probability of any ED use by 4.5 percentage points (p-value = 0.111), which is about 65 percent of the magnitude from the linear instrumental variable (IV) method. For comparison, Taubman et al. (2014) find that Oregon's Medicaid expansion increased the probability of an ED visit by 7 percentage points (a 20 percent increase). We also find a smaller extensive margin impact of a statistically significant 5.7 percentage points for outpatient ED visits (90 percent of total visits), representing 70 percent of the linear IV method's effect. Our results suggest that the average impact of Medicaid on ED use poorly approximates the individualized treatment effect.

Third, we characterize the heterogeneous effects of Medicaid by observed covariates, yielding insights into who drives the average impact on ED use. The main risk factors for increasing ED use are gender, age, participation in other welfare programs, and past ED use for conditions not requiring emergency care. We identify four subgroups estimated to have statistically significant increases in ED use of at least twice the magnitude of the average effect. These groups are men, prior SNAP participants, younger adults less than 50 years old, and those with pre-lottery ED use classified as primary care treatable. The importance of previous ED use for effect heterogeneity and the weaker extensive margin effects point to a central role of intensive margin effects. Much of the overall positive Medicaid effect stems from increases in ED use by those accustomed to using the emergency department for their health care needs rather than people initiating ED use after obtaining coverage. Notably, the increases in ED use occur primarily for newly insured people with a history of using the emergency department for conditions that do not require it.

Fourth, we find noteworthy differences in subgroup effects between the nonparametric and linear IV methods, presenting a cautionary tale on ad hoc subgroup exploration of effect heterogeneity. While the linear IV and nonparametric methods yield similar treatment effects for some subgroups, the linear IV method estimates significant effects for subgroups that are likely to be spurious. For instance, the linear IV method finds significant effects for groups we would not expect to drive differences in treatment response, such as subgroups defined by lottery list variables collected during the sign-up process. Our results suggest that the nonparametric approach is more conservative at identifying group-level effects since it flexibly targets effect heterogeneity.

Finally, we illustrate how to use the heterogeneous effects to estimate simple assignment decision rules to identify those predicted to decrease undesirable ED utilization defined as non-emergent use. Our assignment rule solicits applications from those with little prior outpatient ED use and prioritizes applications from those with severe health problems justifying ED use or requiring inpatient care. While our decision rule reasonably prioritizes those needing ED care for emergencies, we caution that policy-relevant assignment rules must be designed to meet the policy maker's objectives subject to constraints imposed by legislative environments.

The rest of the paper is organized as follows. Section 2 describes the context of Oregon's health insurance experiment and the data. Section 3 presents the causal machine learning framework.

Section 4 presents and discusses the results. Section 5 concludes. Additional results are collected in Appendices A and B.

2. The Oregon Health Insurance Experiment

2.1. Background

We provide a brief overview of the essential features of the Oregon Health Insurance Experiment. Previous studies contain detailed institutional information (Finkelstein et al. 2012, Baicker et al. 2013, Allen et al. 2013, Taubman et al. 2014, Finkelstein et al. 2016). Oregon's split its Medicaid program into two due to a budget shortfall in 2003. The first program, OHP Plus, served categorically eligible people under federal rules—low-income income children, pregnant women, blind or disabled people, and Temporary Assistance to Needy Families (TANF) recipients. The second program, OHP Standard, served low-income and non-disabled adults who were not categorically eligible for its main OHP Plus program. Oregon's philosophy was to provide more people with coverage for fewer services rather than limiting participation. Thus, OHP Standard covered fewer services and came with a higher premium and cost-sharing requirements, resulting in a decline in enrolment (Allen et al. 2013). The resulting attrition in OHP Standard participation led to an accumulated budgetary surplus by 2007. It was then that Oregon decided to expand its OHP Standard program by offering about 10,000 spots using a lottery to solicit or invite applications.

Lottery winners applied for Medicaid, and those determined to be eligible received coverage. The main eligibility criteria were as follows—being 19–64 years of age; being an Oregon resident who is not otherwise eligible for public insurance; being a U.S. citizen or legal immigrant; being uninsured for the previous six months, and having income below the federal poverty level with assets not exceeding $2,000. Of the 89,824 individuals who signed up for the lottery, 35,169 won, and about 30% of the winners enrolled in Medicaid.

The low-income, uninsured population served by OHP Standard should be kept in mind when interpreting the Oregon experiment results. Finkelstein et al. (2012) elaborate on the difficulties with extrapolating Oregon's finding to other Medicaid expansions. The population served by OHP Standard is neither representative of the typical Medicaid population served by OHP Plus nor the low-income uninsured population in the United States. Compared to the low-income U.S. population, OHP Standard's target population in Oregon has more whites (84%), fewer Blacks (2%), is older, and reports being in worse health (Allen et al. 2010). Also, people voluntarily put themselves on the waiting list to participate in the lottery. These issues do not threaten the internal validity of the Oregon experiment but the inherent sample selection on observed (and unobserved dimensions) suggests a non-trivial role for moral hazard heterogeneity in understanding its findings (Einav et al. 2013).

2.2. Data

We use the publicly available data files from Taubman et al. (2014) (Finkelstein 2013). The sample contains ED visit information for 24,646 individuals from 2007 to 2009. The pre-lottery period spans January 1, 2007, through March 9, 2008. The study period is 18 months from the earliest notification date on March 10, 2008, through September 30, 2009. Medicaid coverage is constructed from state administrative records and defined as any receipt during the study period.

We analyze fourteen ED visit measures. These variables measure overall ED use and three categories of ED visits. The first category groups ED use by hospital admission, including outpatient and inpatient ED visits. The second category groups ED visits by the time of occurrence and consists of those occurring during on-time (7 am to 8 pm on Monday to Friday) and off-time hours (nights and weekends).

The final category groups ED visits by whether they required immediate care or not. We adopt the classification of ED visits based on the primary ICD-9 diagnosis code (Taubman et al. 2014, Billings et al. 2000). While we refer the reader to Taubman et al. (2014) for a detailed description, the algorithm assigns to each ED visit a probability that it is one of four types. The first type—emergent, non-preventable—includes unpreventable illnesses requiring immediate ED care (e.g., heart attacks and nonspecific chest pain). The second type—emergent, preventable—includes

ED visits that require immediate ED care but are avoidable (e.g., asthma attacks and urinary tract infections). The third type—primary care treatable—includes ED visits requiring immediate care but not through the emergency department (e.g., sprains, strains, and abdominal pain). The final type—non-emergent—contains ED visits that do not require immediate care (e.g., headaches and back problems). Each visit is assigned a probability of being in all four categories. The number of visits for each type is then obtained by summing the assigned probabilities across all visits (Taubman et al. 2014). We do not analyze unclassified visits not assigned to any of the above four categories.

We analyze both the binary and continuous versions of these types of ED visits. While Taubman et al. (2014) only analyzes the continuous versions, we create their binary counterparts as follows. Since the number of ED visits for each type is computed as a sum of probabilities, we classify the individual as having only one type of visit for those with one observed ED visit, which is set equal to the visit type with the highest probability. For those with two ED visits, we classify the individual as having at most two visit types, comprising the ones with the top two highest probabilities. We classify the individual as having at most three visit types for those with three ED visits, with the types being those with the top three highest probabilities. Finally, we classify individuals with four or more ED visits as having all visit types with non-zero assigned probabilities.

All our analyses include variables needed to ensure unbiased estimation of treatment effects, such as household size and an indicator to which of the eight lottery draws between March and September 2008 the individual was assigned. Because of our interest in heterogeneous effects along observed dimensions, we include two additional groups of covariates—lottery list variables and baseline characteristics measured before the lottery. We focus on these covariates—lottery list and pre-randomization measures—to explore heterogeneity to ensure that they are not endogenous to Medicaid receipt.

The lottery list contains eight variables constructed from the lottery sign-up sheet—age; sex; indicators for whether English is the preferred language for receiving materials; whether the individuals signed themselves up for the lottery; whether an individual provided a phone number on the sign-up form; whether the individuals listed their address as a P.O. Box; whether the individual signed up on the first day of the lottery, and the median household income in the applicant's zip code from the 2000 decennial census. We are only missing the last variable compared to the original analysis because it is not publicly available, but this does not affect the internal validity of the treatment effects.

The baseline characteristics are pre-lottery ED use, participation in the Supplemental Nutrition and Assistance Program (SNAP), and receiving TANF in the pre-lottery period. We include twenty pre-lottery ED measures covering overall ED use, on-/off-time use, ED visits resulting in hospital admission, visits for each type described above, and visits for specific types of injuries and health conditions. We include this expanded list of pre-lottery ED visit information to capture historical emergency care demand. The pre-lottery SNAP/TANF participation variables come from administrative data on program receipt and measure the total household benefit amounts received between January 1, 2007, through the individual's notification date.

We restrict the sample to those with pairwise non-missing observations in the covariates and each outcome. Our baseline sample consists of 24,615 individuals with non-missing information on all covariates (i.e., 99.9 percent of Taubman et al.'s (2014) sample). We use the baseline sample to analyze our primary outcome—whether the participant had any ED visits. The analysis samples for other outcomes are slightly smaller due to missing data. However, the difference is negligible; the smallest sample comprises 24,588 observations—only 27 fewer people relative to the baseline sample. Our results are unaffected by conditioning on a uniform sample of non-missing observations across all outcomes.

Table 1 presents the summary statistics for our baseline sample. The sample is 55 percent female with an average age of 40 years. About 54 percent of the sample received SNAP with an average SNAP benefit amount of $1,332 in the pre-lottery period. TANF receipt is much lower at 2 percent of the sample with an average benefit of $96. The sample averaged 0.77 ED visits in the pre-lottery

Table 1: Descriptive Statistics

Variable	Mean	SD	Median	Min	Max
Lottery list and baseline characteristics					
Age (years)	39.60	12.05	39.0	20	63
Gave phone number	0.87	0.34	1.0	0	1
English as preferred language	0.86	0.34	1.0	0	1
Female	0.55	0.50	1.0	0	1
Week of lottery sign up	1.58	1.62	1.0	0	5
Provided P.O. box address	0.03	0.16	0.0	0	1
Signed up self for lottery	0.90	0.30	1.0	0	1
Pre-lottery SNAP recipient	0.54	0.50	1.0	0	1
Pre-lottery SNAP benefit amount ($)	1331.94	1863.76	522.5	0	20745
Pre-lottery TANF recipient	0.02	0.15	0.0	0	1
Pre-lottery TANF benefit amount ($)	96.16	694.70	0.0	0	16031
Pre-randomization ED use					
Number of overall visits	0.77	1.86	0.0	0	17
Number of inpatient visits	0.09	0.41	0.0	0	6
Number of outpatient visits	0.69	1.71	0.0	0	16
Number of on-hours visits	0.45	1.23	0.0	0	13
Number of off-hours visits	0.33	0.92	0.0	0	10
Number of emergent, non-preventable visits	0.16	0.50	0.0	0	9
Number of emergent, preventable visits	0.06	0.28	0.0	0	6
Number of primary care treatable visits	0.27	0.75	0.0	0	12
Number of non-emergent visits	0.16	0.58	0.0	0	12
Number ambulatory-care-sensitive visits	0.05	0.30	0.0	0	5
Number of visits (chronic conditions)	0.14	0.64	0.0	0	9
Number of visits (injury)	0.17	0.56	0.0	0	6
Number of visits (skin conditions)	0.05	0.31	0.0	0	5
Number of visits (abdominal pain)	0.04	0.27	0.0	0	5
Number of visits (back pain)	0.03	0.26	0.0	0	5
Number of visits (chest pain)	0.02	0.18	0.0	0	3
Number of visits (headache)	0.02	0.23	0.0	0	4
Number of visits (mood disorders)	0.02	0.23	0.0	0	5
Number of visits (psychiatric conditions)	0.06	0.38	0.0	0	6
Sum of total ED charges	894.85	2593.46	0.0	0	42315
Any overall visit	0.34	0.47	0.0	0	1
ED use outcomes					
Any inpatient visit	0.07	0.26	0.0	0	1
Any outpatient visit	0.32	0.47	0.0	0	1
Any emergent, non-preventable visit	0.16	0.37	0.0	0	1
Any emergent, preventable visit	0.10	0.30	0.0	0	1
Any primary care treatable visit	0.22	0.41	0.0	0	1
Any non-emergent visit	0.16	0.37	0.0	0	1
Number of overall visits	1.00	2.41	0.0	0	22
Number of inpatient visits	0.11	0.53	0.0	0	7
Number of outpatient visits	0.89	2.21	0.0	0	21
Number of emergent, non-preventable visits	0.21	0.69	0.0	0	14
Number of emergent, preventable visits	0.07	0.35	0.0	0	8
Number of primary care treatable visits	0.35	0.96	0.0	0	16
Number of non-emergent visits	0.20	0.70	0.0	0	13

Notes: This table presents descriptive statistics based on the Oregon Health Insurance Experiment data. The sample consists of 24,615 individuals in the Taubman et al. (2014) sample with non-missing information on pre-lottery emergency department utilization and SNAP/TANF receipt. The pre-randomization SNAP and TANF benefits are total household benefits received from January 2007 through the individual's lottery notification date. The pre-randomization ED use measures are with respect to utilization between January 1, 2007 and March 9, 2008.

period, with the total emergency department facility charges being $895 on average.[1]

3. Methods

Our analysis of the impacts of Medicaid on ED use is twofold. First, we estimate different types of heterogeneous effects of Medicaid on ED utilization using causal forests. Causal forests are a special type of generalized random forests (GRF). GRFs, in turn, are an extension of random forests that can estimate population quantities identified as solutions to local moment conditions (Wager & Athey 2018, Athey et al. 2019). In the second part of our analysis, we use the heterogeneous treatment effects to estimate assignment rules following advances in policy learning algorithms (Athey & Wager 2021, Kitagawa & Tetenov 2018). We then describe the population that would be prioritized for coverage under optimal assignment given an objective function and constraints on program size.

3.1. Treatment effect heterogeneity using generalized random forests

We represent the treatment effect parameters in the potential outcomes framework. Let $D_i \in \{0, 1\}$ denote the individual's Medicaid coverage. Also, let $Y(D)$ represent the potential outcomes associated with Medicaid receipt ($D = 1$) or otherwise ($D = 0$). It is well-known that the individual treatment effect, $\tau_i = Y_i(1) - Y_i(0)$ is not identified because of the fundamental problem of causal inference. Under additional assumptions, we can estimate aggregate estimands such as the Average Treatment Effect (ATE). Since Medicaid coverage is endogenous, we pursue instrumental variable estimation using Oregon's lottery assignment as an instrument, denoted by $Z_i \in \{0, 1\}$. IV estimation identifies the LATE, that is, the ATE of Medicaid coverage for a sub-population of compliers (Imbens et al. 1994).

[1]The continuous measures of ED utilization have been top-coded in the public-use Oregon experiment files such that each value has as least ten observations. Nonetheless, this does not matter practically given Taubman et al.'s (2014) online appendix states that "[t]his means that for many of the number-of-visit outcomes, the publicly available data will not directly replicate the results presented in the main text and supplementary materials, although our findings are robust to the censoring we imposed in the public use data."

Suppose that the outcome is given by a structural model of the form $Y = \mu(X) + \tau(X)D + \varepsilon$, where X is a vector of covariates, $\mu(X)$ is the mean outcome, $\tau(X)$ is the treatment effect, and ε is the error term such that $\mathbb{E}[\varepsilon|X, Z] = 0$. Abadie (2003) shows that $\tau(X)$ identifies the *conditional* LATE defined as

$$\tau(X) = \frac{\mathbb{E}[Y|X, Z = 1] - \mathbb{E}[Y|X, Z = 0]}{\mathbb{E}[D|X, Z = 1] - \mathbb{E}[D|X, Z = 0]}. \tag{1}$$

Note that $\tau(X)$ refers to various types of conditional average treatment effects (CATE) at different covariate levels.[2] At the most granular level, $\tau(X)$ is an estimate of the individualized treatment effect (ITE) in the smallest covariate partition defined by an individual's covariate values. At higher levels, $\tau(X)$ represents the ATE for well-defined subgroups. We refer to this aggregate estimand as the group average treatment effect (GATE).

We can estimate the treatment effect, $\tau(X)$, via the population moment conditions given by

$$\mathbb{E}[Y - \mu(X) - \tau(X)D|X]\ (1\ Z')' = 0. \tag{2}$$

As a form of nearest-neighbor estimation, Athey et al. (2019) first define a similarity weight, $\alpha_i(X)$, and then obtain estimates $\left(\widehat{\tau}(X),\ \widehat{\mu}(X)\right)$ as solutions to the local estimation equations

$$\sum_{1}^{n} \alpha_i(X)[Y_i - \widehat{\mu}(X) - \widehat{\tau}(X)D_i|X = x_i]\ \left(1\ Z_i'\right)' = 0. \tag{3}$$

The weights, $\alpha_i(X)$, come from the causal forest as an average of tree-based neighborhoods around each point x.[3] Athey et al. (2019) show that using the data-adaptive forest-based weights, $\alpha_i(X)$, in equation (3) yields consistent and asymptotically normal estimates of $\tau(X)$.

[2]While $\tau(X)$ in equation (1) can be more appropriately referred to as the conditional local average treatment effect (CLATE), we refer to it as the CATE for easier exposition.

[3]The weights are obtained by first growing B trees using the standard Classification and Regression Tree (CART) algorithm. Unlike traditional CART, splits at the parent nodes are chosen to maximize heterogeneity in $\tau(X)$ estimates. Each tree marks observations that fall in the same leaf as x and gives them a positive weight. The similarity weights $\alpha_i(X)$ measure how often an observation falls in the same neighborhood ("leaf") as x. The forest-based algorithm then averages over the B tree-based neighborhoods to obtain $\alpha_i(X)$ which are then used to fit the estimation equations in (3).

We note some key implementation details. We grow all forests using 100,000 trees and sample half the data to build each tree. Tree building relies on further subsample splitting based on the honesty principle, using half of the sampled data to determine the tree splits and the other half for estimation. Since this may result in empty leaves, we prune these empty leaves so that each tree can handle all test points. We determine the number of variables considered at each split point using the default rule-of-thumb. For inference, the variance of the forest-based $\widehat{\tau}(X)$ are constructed using the bootstrap of little bags (Athey et al. 2019, Sexton & Laake 2009). We cluster standard errors at the household level.

After obtaining the nonparametric estimates of $\tau(X)$, we aggregate them to get a doubly robust estimate of the average impact using augmented inverse probability weighting (AIPW). Similarly, we can obtain doubly robust GATE estimates for specific subgroups defined by one or more covariates. We compute a variable importance ranking of all covariates used in growing the forest based on the number of times the variable is split upon weighted by the tree depth of the split. This measure yields valuable insights into which covariates are most important for driving treatment effect heterogeneity.

3.2. Learning optimal treatment assignment

Academic researchers and policy leaders usually promote randomization schemes to allocate scarce resources. Using a lottery seems fair and politically feasible. However, lotteries are one of many options. The Oregon Medicaid Director, Jim Edge, summarizes the difficulty of the treatment assignment problem in the New York Times: "[w]e thought about other options, such as should we try to pick all of the sickest people or the kids or the people with cancer or heart disease. But the Feds won't allow that, and there's just no way to guarantee the fairness of that. Why would cancer be more deserving than heart disease?" (Yardley 2008). Based on our heterogeneous effects, we discuss alternative treatment assignment schemes. These alternative assignments schemes are necessarily normative. Suppose that administrators are interested in maximizing some notion of welfare conditional on observed characteristics. Then random assignment may be inferior to a policy based on the empirical distribution of treatment effects.

In the context of Medicaid, we might be interested in soliciting applications from individuals in a way that minimizes undesirable utilization, such as ED visits for non-emergent conditions. Moreover, devising a rule-based allocation scheme may also improve targeting and free up funds to extend benefits to a larger pool of people. Alternatively, administrators might be interested in identifying people who should participate in an educational intervention about primary care providers due to cost constraints. We use the Oregon experiment to illustrate one way to answer the following questions: (1) How should a social planner who wants to minimize unnecessary ED use optimally choose people to treat, and how would the assigned population differ from the lottery population? (2) How would a simple decision rule prioritizing individuals least likely to increase undesirable ED visits look?

The literature on statistical decision rules utilizes the distribution of treatment effects to design treatment assignment rules that maximize mean social welfare, typically defined as the mean post-treatment outcome (Manski 2004, Dehejia 2005, Hirano & Porter 2009, Stoye 2009, 2012, Kitagawa & Tetenov 2018, Athey & Wager 2021). We rely on the theoretical guarantees from this literature to conduct two exercises. First, we compute the optimal allocation scheme based on a specified objective function to solicit Medicaid applications—the intent-to-treat (ITT) scenario—using integer programming (Kitagawa & Tetenov 2018). The integer programming approach lets us incorporate capacity constraints, allowing us to restrict the selected sample to equal the number of lottery winners in the Oregon experiment. We do not impose additional constraints on the budgetary costs of the program, but incorporating them is straightforward. We then compare the population selected based on the optimal linear program to those who won the Oregon lottery.

In the second exercise, we utilize Athey & Wager's (2021) algorithm to estimate a treatment assignment rule in the form of a simple depth-2 decision tree. We interpret these assignment rules as providing guidance regarding whom to solicit Medicaid applications given a sign-up list of people (instead of randomly selecting households from it). Administrators can use such decision rules in other ways, such as selecting people for specific educational interventions. We focus on learning assignment rules for the ITT context because it more closely aligns with the practical problem of allocating Medicaid spots in Oregon's expansion of OHP Standard to a new population. In other words, the social planner can only offer the opportunity to apply for coverage but not forcibly assign people to the program. Assignment rules for inviting applications for coverage may improve post-treatment outcomes relative to the lottery system. Of course, due to concerns regarding external validity, the assignment rules learned from Oregon may not directly apply to other states but rather provide a blueprint for similar expansions.

Learning treatment assignment rules mapping observed covariates to a binary treatment decision requires estimating doubly robust scores targeting the estimand of interest—the ITT effect in this case. Then we choose a policy π in the finite-depth class of policies Π to minimize regret, i.e., the loss resulting from not choosing the ideal policy.[4] Previous work has shown that this optimization problem is equivalent to a weighted binary classification problem, effectively predicting the sign of the treatment effect, where the weights are given by the size of the conditional treatment effect estimates (e.g. Zhao et al. 2012).

We need to make three practical choices for estimating the alternative treatment assignment regime. The first is to define the post-treatment outcome or utility constituting the objective function. We use the number of non-emergent ED visits described in the data section. These visits do not require immediate care and can be deemed unnecessary or undesirably requiring service in the emergency department. Our learned assignment rule minimizes the number of non-emergent visits.

Second, we choose which covariates to include in the decision rule estimation. We exclude variables that interfere with fairness, ethical and legal considerations. While we use the complete set of covariates to estimate the heterogeneous treatment effects, we limit the variables used to estimate the assignment rules to nine pre-lottery ED utilization variables. Specifically, we include the overall number of pre-lottery ED visits, the number of inpatient ED visits, the number of outpatient ED visits, the number of on-hour ED visits, and the number of off-hours ED visits. The remaining variables are indicators for the type of pre-lottery ED utilization—indicators of

[4]We refer the reader to Athey & Wager (2021) for a detailed technical discussion.

any pre-lottery emergent ED use, any pre-lottery primary care treatable ED use, any pre-lottery non-emergent ED use, and the pre-lottery sum of total ED facility charges. Finally, we estimate the assignment rules as decision trees of depth 2. Since the algorithm obtains the decision rule by exact (exhaustive) tree search, the required time for estimating more complex decision rules is exponential in tree depth.

4. Results

4.1. Distribution of individualized treatment effects of Medicaid on ED use

We first present the heterogeneous impacts of Medicaid coverage using causal forests as described in Section 3. We mainly focus on the extensive margin results of receiving Medicaid on binary indicators of ED use in the paper. However, we briefly refer to additional intensive margin results of Medicaid on continuous ED visit measures in Appendix A. We provide similar intent-to-treat heterogeneous effects of winning the lottery on ED utilization in Appendix B.

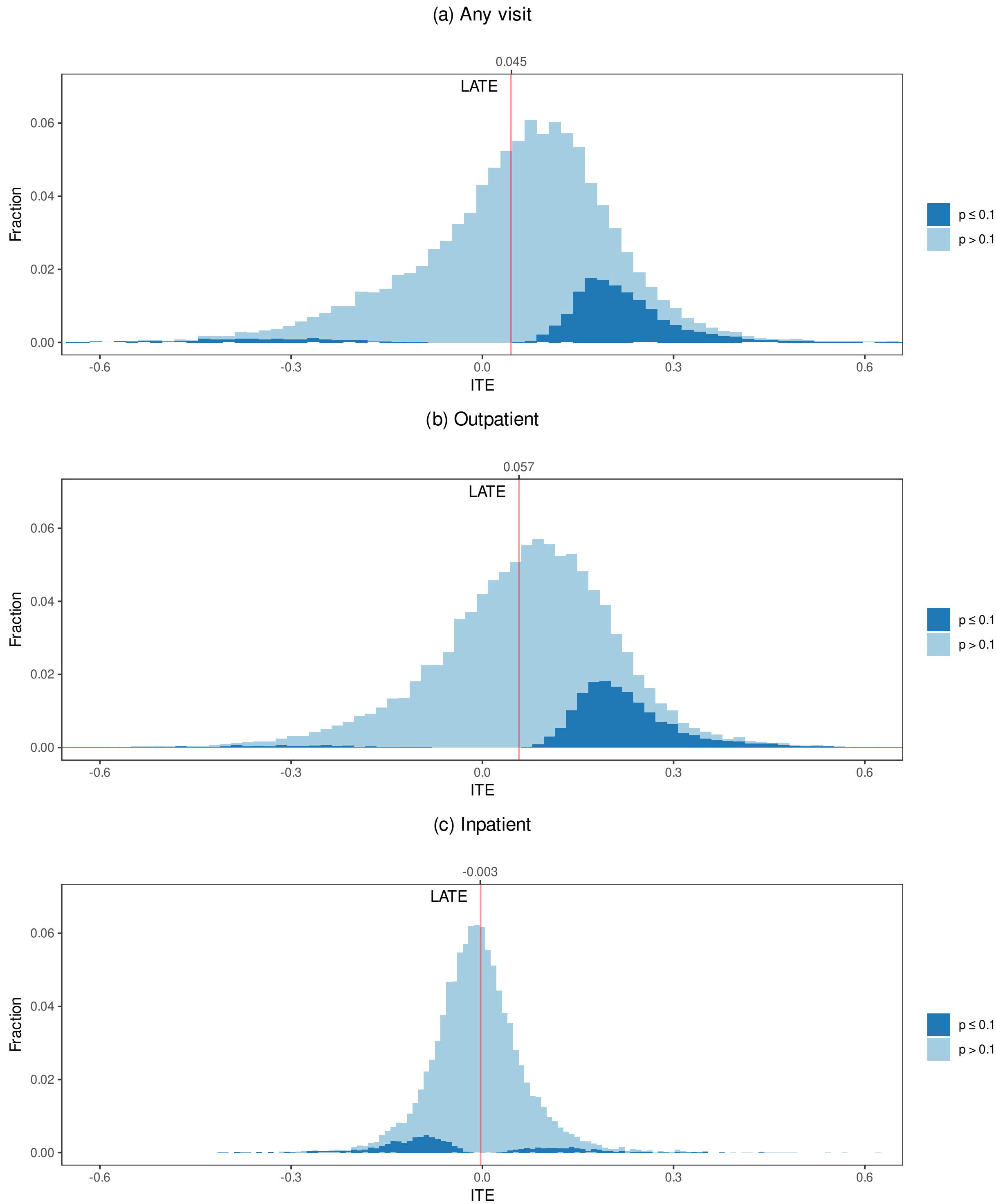
We find substantial and meaningful heterogeneous effects of Medicaid across all ED visit outcomes. Figure 1 displays the individualized treatment effects of Medicaid on the probability of any visit, with the darker blue shade indicating statistical significance at the 10% level. The subplots in Figure 1 display results for any overall visits (panel a), any outpatient visits (panel b), and any inpatient visits (panel c). These results come from $\tau(X)$ estimates at the smallest covariate partition, with each point X denoting each individual's covariate values. Table 2 reports selected quantiles of the empirical distribution of the ITE estimates for all outcomes.

The heterogeneous effects of Medicaid on any ED visits are widely dispersed and skewed to the left, showing positive and negative ITE estimates. About one-third of the ITE estimates are negative or practically zero. Table 2 reports the 25^{th} percentile to be a reduction of 3.1 percentage points, with the median of 7 percentage points increase, which is larger than the mean effect of 4.5 percentage points. We momentarily discuss the average (aggregate) effects in the "LATE" column of Table 2 in Section 4.2. The statistically significant increases in ED use at the 10% level are limited to about 14.2% of the sample (with 8% being significant at the 5% level). There is no overlap between the statistically significant ITE estimates and the overall mean effect. At the 5% level, the smallest ITE estimate among those who significantly increase ED usage is 7.6 percentage points, close to 1.7 times the mean effect. While most of the significant heterogeneous effects are positive, we find statistically significant reductions in ED use, making up only 2% of all individualized effects.

A comparison of panels (a) and (b) in Figure 1 reveals that Medicaid's impacts on ED utilization are concentrated among outpatient visits. We find a wide dispersion in the individualized treatment effects for outpatient ED visits, with a range similar to overall ED use. The 25^{th} percentile is a reduction of 1.3 percentage points, and the median ITE estimate of 7.8 percentage points continues to be larger than the mean effect of 5.7 percentage points. Notably, the distribution of individualized treatment effects shows that the null effects of Medicaid on inpatient ED use mask sizable proportions of counterbalancing statistically significant negative and positive effects (panel (c) of Figure 1).

Our results indicate that we obtain the same pattern of results when we analyze the intensive margin of the number of ED visits. The corresponding graphs for the number of visits are shown in Figure A.1 of Appendix A. Table 2 displays moments of the ITE distribution for different types of ED visits on the extensive and intensive margin. We find meaningful heterogeneity with individualized treatment effects ranging from a reduction of 1.50 to an increase of 5.94 ED visits. Unlike the extensive margin, the median of 0.27 is slightly smaller than the mean effect of 0.35 ED visits. The distribution has substantially more mass to the left, outweighing the effect of high-use individuals in the right tail. The results for outpatient ED visits closely mirror those of the overall number of ED visits. The outpatient ITE estimates span a decrease of 1.20 to an increase of 5.83 ED visits, with the median and mean effects being 0.29 and 0.36 visits, respectively. Similar to the extensive margin, the inpatient ITE estimates show sizable proportions of countervailing negative and positive effects, culminating in a statistically insignificant mean effect.

Figure 1: Distribution of individualized treatment effects of Medicaid on the propensity of any ED visit



Notes: This figure plots the individualized treatment effects of Medicaid on any overall ED visit (panel a), any outpatient ED visits (panel b), and any inpatient ED use (panel c) based on generalized random forests. The darker shade denotes statistical significance at the 10% level. The red vertical line indicates the local average treatment effect. The baseline sample consists of 24,615 individuals in the Taubman et al. (2014) sample with non-missing information on pre-lottery emergency department utilization and SNAP/TANF receipt. The estimates displayed exclude less than half a percentile at the top and bottom of the distribution, resulting in the axes corresponding approximately to the percentile range [0.5%, 99.5%]. Bin size is chosen according to the Freedman-Diaconis rule.

Table 2: Empirical quantiles of the distribution of individualized treatment effects of Medicaid on ED use

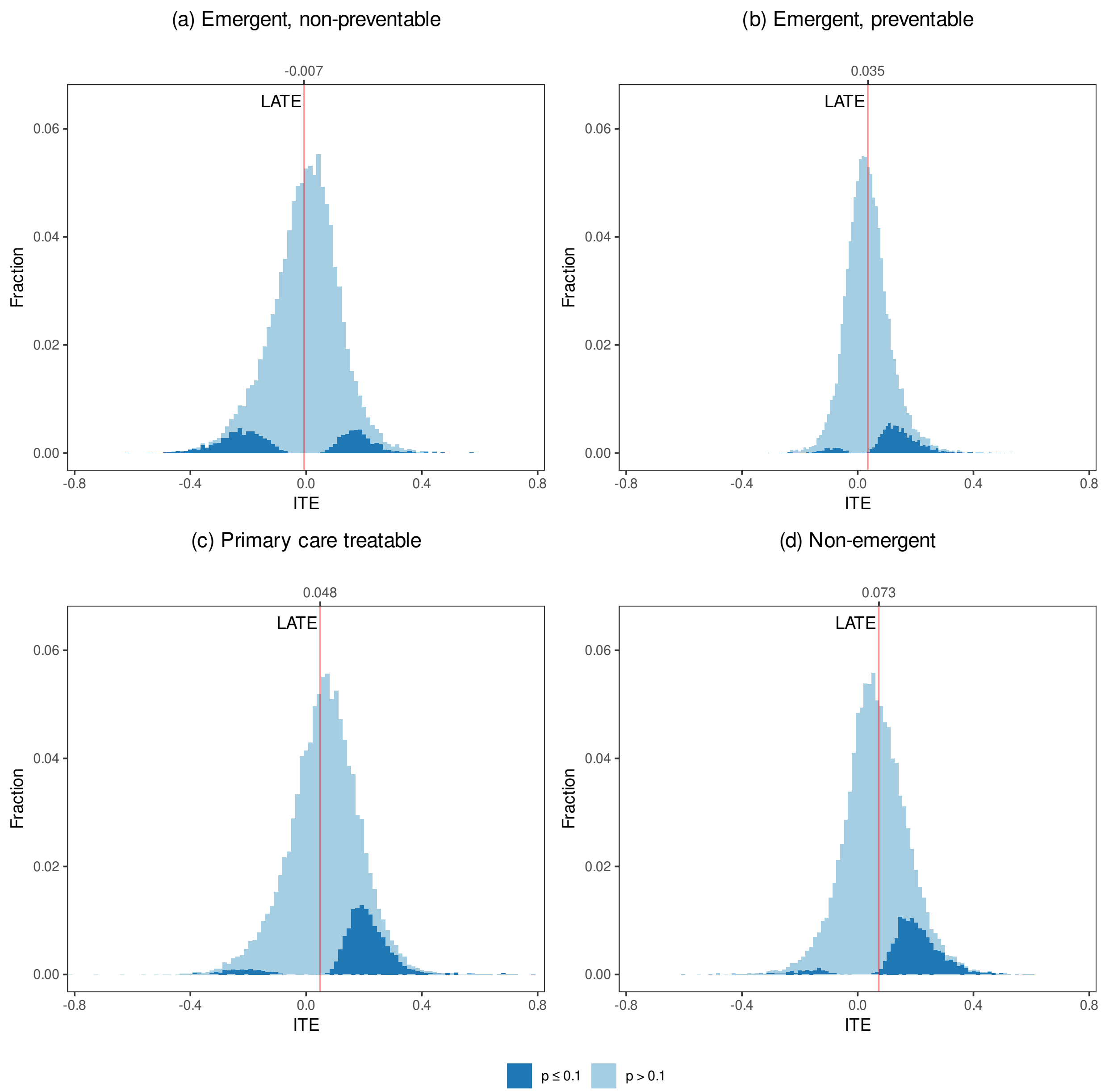
Variable	LATE	Min	25%	50%	75%	Max
Extensive margin						
Any overall visit	0.045	-1.122	-0.031	0.070	0.151	0.894
Any inpatient visit	-0.003	-0.415	-0.049	-0.010	0.031	0.619
Any outpatient visit	0.057	-1.121	-0.013	0.078	0.160	0.800
Any emergent, non-preventable visit	-0.007	-0.622	-0.075	0.003	0.072	0.583
Any emergent, preventable visit	0.035	-0.313	-0.014	0.030	0.077	0.532
Any primary care treatable visit	0.048	-0.813	-0.011	0.067	0.141	0.785
Any non-emergent visit	0.073	-0.607	-0.004	0.060	0.132	0.604
Intensive margin						
Number of overall visits	0.354	-1.503	0.006	0.273	0.567	5.939
Number of inpatient visits	-0.012	-0.941	-0.074	-0.013	0.046	1.072
Number of outpatient visits	0.360	-1.199	0.052	0.291	0.571	5.828
Number of emergent, non-preventable visits	0.028	-0.590	-0.066	0.020	0.109	3.025
Number of emergent, preventable visits	0.038	-1.533	-0.015	0.026	0.077	0.724
Number of primary care treatable visits	0.161	-0.550	0.018	0.131	0.255	3.689
Number of non-emergent visits	0.107	-0.663	-0.008	0.061	0.143	2.606

Notes: This table reports selected quantiles of the individualized treatment effects of Medicaid on ED use based on generalized random forests. The first column reports the average effect. The baseline sample consists of 24,615 individuals in the Taubman et al. (2014) sample with non-missing information on pre-lottery emergency department utilization and SNAP/TANF receipt.

We turn our attention to the heterogeneous results by type of ED visit presented in Figure 2 for the extensive margin. Consistent with the overall ED visits results, individualized treatment effects of Medicaid's impacts on the types of ED visits exhibit substantial heterogeneity. For any emergent, non-preventable visits, panel (a) in Figure 2 shows negative and positive statistically significant effects of sizable proportions. These two opposing forces are consistent with a null effect on average. A similar pattern holds for the number of emergent, non-preventable visits, with the corresponding graph in Figure A.2 of Appendix A.

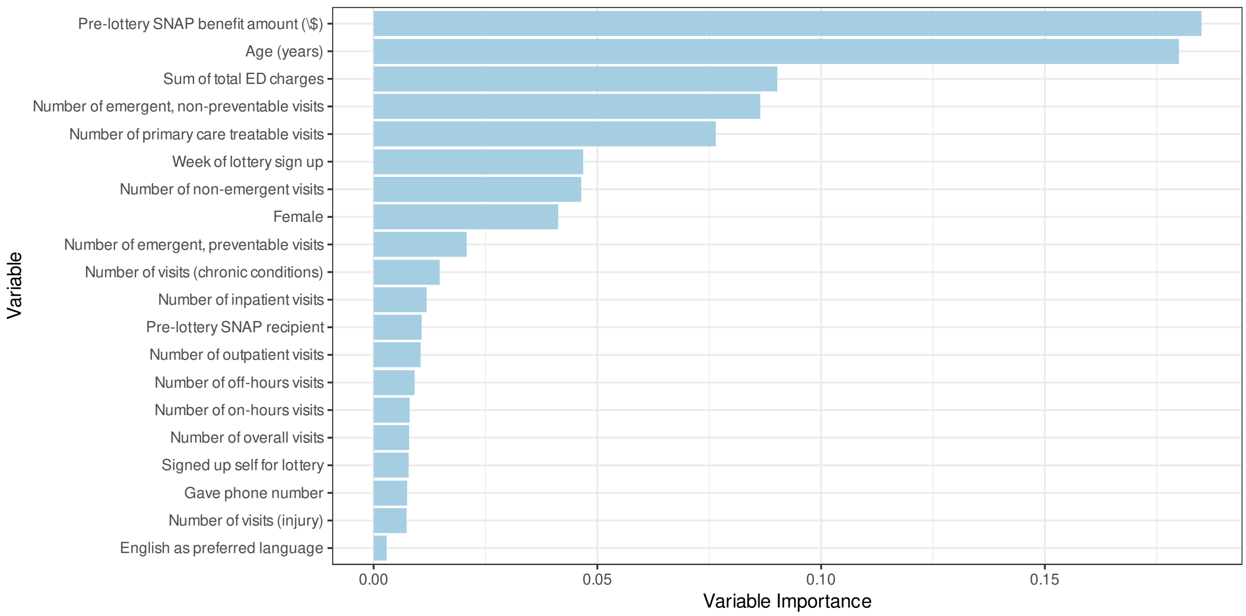
Figure 2 shows notable differences in the heterogeneous impacts for the remaining three types of ED visits. We find a larger share of positive and statistically significant individualized effects for primary care treatable and non-emergent visits than emergent visits—non-preventable and preventable—on both margins. Among all ED visit types, the two largest increases on the extensive and intensive margin always occur for primary-care treatable and non-emergent visits. Increases in primary care treatable visits are more pronounced on the intensive margin, whereas increases in non-emergent visits are larger on the extensive margin.

Figure 2: Distribution of individualized treatment effects of Medicaid on the propensity of any ED visit by type of condition



Notes: This figure plots the individualized treatment effects of Medicaid by type of ED visit based on generalized random forests for any emergent, non-preventable visit (panel a), any emergent, preventable visit (panel b), any primary care treatable visit (panel c), and any non-emergent visit (panel d). Measures of the type of ED visit are based on Billings et al.'s (2000) algorithm described in Taubman et al. (2014). We use these measures to construct binary indicators of ED visits by type of condition as described in the main text. The darker shade denotes statistical significance at the 10% level. The red vertical line indicates the local average treatment effect. The baseline sample consists of 24,615 individuals in the Taubman et al. (2014) sample with non-missing information on pre-lottery emergency department utilization and SNAP/TANF receipt. The estimates displayed exclude less than half a percentile at the top and bottom of the distribution, resulting in the axes corresponding approximately to the percentile range [0.5%, 99.5%]. Bin size is chosen according to the Freedman-Diaconis rule.

Figure 3: Variable importance scores in growing the causal forest (Any Visit)



Notes: This figure shows the variable importance scores of the top 20 characteristics in growing the generalized random forests used to estimate the individualized treatment effects of Medicaid for any overall ED visit. The variable importance measure is a simple weighted sum of the proportion of times a variable is used in a tree splitting step at each depth in growing the forest. The scores roughly capture how important a variable is for driving treatment effect heterogeneity. The baseline sample consists of 24,615 individuals in the Taubman et al. (2014) sample with non-missing information on pre-lottery emergency department utilization and SNAP/TANF receipt.

While the large increases for primary-care treatable and non-emergent visits are consistent with the original Oregon results, we expected to find some significant negative ITEs in these categories. This hypothesis is because those are the types of visits most likely affected by gaining health insurance coverage and are easier addressed by access to primary care physicians. In contrast, we only observe a sizable amount of negative ITEs for the emergent, non-preventable category. This finding suggests that some people reduce ED usage by not seeking treatment for conditions that potentially require treatment in the emergency department.

In summarizing, we find a consistent pattern of heterogeneous effects that offer a more nuanced interpretation of Medicaid's effect on ED utilization that is not discernible by focusing on the mean effect. Before examining the aggregate Medicaid effects in the next section, it is instructive to highlight the most important covariates driving these heterogeneous effects. Put differently, what are the most used individual characteristics for tree splitting in growing the causal forests? Table

1 lists the covariates utilized in growing the causal forests—pre-randomization ED use (overall use and by visit type), SNAP and TANF receipt measures, and lottery list variables. Figure 3 displays the variable importance scores for the top 20 characteristics. We present the corresponding variable importance plot for the number of ED visits and the scores for all variables in Appendix A.

For the extensive margin, Figure 3 shows that the causal forest splits the most on pre-lottery SNAP benefits followed by age, the sum of total ED facility charges, the number of pre-lottery emergent but non-preventable ED visits, and the number of pre-lottery primary care treatable ED visits. It is reassuring that the forest rarely splits on variables that we do not expect to drive treatment effect heterogeneity (e.g., lottery list variables such as whether the individual requested English language materials).

4.2. Aggregate effects of Medicaid on ED use

Based on the heterogeneous effects uncovered by the causal forests, we present doubly robust estimates of the *average* Medicaid effect (i.e., the LATE). Specifically, we estimate the mean effect by averaging the doubly robust scores (based on AIPW) targeting the LATE instead of computing a simple average of the individualized treatment effects, $\widehat{\tau}(X)$. By so doing, the forest-based LATE estimator is nonparametric and retains the desirable double robust property of being consistent if at least one component of the score is consistent (Chernozhukov et al. 2018, Athey et al. 2019). Table 3 presents the nonparametric and linear IV estimates of Medicaid's effect on all measures of ED utilization.

The nonparametric LATE estimates point to a weaker extensive margin effect of Medicaid on ED utilization. Regarding any overall ED visits, we find that Medicaid increased the probability of having any ED visit by 4.5 percentage points, which is not statistically significant at the 5% level (p-value=0.111). Moreover, this nonparametric estimate is only about 65 percent of the magnitude of the linear IV estimate of 6.9 percentage points (p-value=0.004). The forest-based mean results show a similar pattern for outpatient ED visits, with a smaller increase of 5.7 percentage points (p-value=0.037) relative to 8.2 percentage points (p-value=0.001) using the linear IV method.

Table 3: Treatment effect estimates of Medicaid coverage on ED use

	GRF estimates			Linear estimates		
Variable	LATE	SE	p-value	LATE	SE	p-value
Extensive margin						
Any overall visit	0.045	0.028	0.111	0.069	0.024	0.004
Any inpatient visit	-0.003	0.015	0.855	-0.011	0.013	0.424
Any outpatient visit	0.057	0.028	0.037	0.082	0.024	0.001
Any emergent, non-preventable visit	-0.007	0.021	0.743	0.007	0.019	0.731
Any emergent, preventable visit	0.035	0.016	0.029	0.039	0.015	0.009
Any primary care treatable visit	0.048	0.024	0.040	0.069	0.021	0.001
Any non-emergent visit	0.073	0.021	0.000	0.064	0.019	0.001
Intensive margin						
Number of overall visits	0.354	0.114	0.002	0.375	0.106	0.000
Number of inpatient visits	-0.012	0.028	0.676	-0.017	0.026	0.516
Number of outpatient visits	0.360	0.106	0.001	0.387	0.098	0.000
Number of emergent, non-preventable visits	0.028	0.034	0.414	0.040	0.032	0.223
Number of emergent, preventable visits	0.038	0.018	0.039	0.036	0.017	0.034
Number of primary care treatable visits	0.161	0.049	0.001	0.171	0.045	0.000
Number of non-emergent visits	0.107	0.036	0.003	0.105	0.033	0.002

Notes: This table reports the estimates of Medicaid coverage on ED use based on generalized random forests and a linear IV model. The sample consists of 24,615 individuals in the Taubman et al. (2014) sample with non-missing information on pre-lottery emergency department utilization and SP/TANF receipt.

Our nonparametric estimate for inpatient ED visits suggests a smaller reduction of 0.3 percentage points relative to 1.1 percentage points using the linear IV approach, although neither is statistically significant. The pattern of weaker extensive margin effects is less pronounced when we examine the results for ED use by type of condition. On the intensive margin, the nonparametric estimates more closely align with the linear method for the results by type of ED visits. However, the former is slightly smaller in magnitude in all but one instance.

To better understand the divergence of the nonparametric from the linear IV results, we first rule out that those differences arise from our sample composition relative to Taubman et al. (2014). The linear IV estimates in Table 3 confirm that we closely replicate Taubman et al.'s (2014) main findings, with minor differences in magnitude plausibly due to rounding and the top-coding in the public-use Oregon data files. The fact that we replicate their findings is not surprising given the negligible difference between our samples when we restrict the sample to observations with non-missing pre-lottery utilization.[5] Therefore, we are confident that our linear IV results are

[5] In their online appendix, Taubman et al. (2014) note using "a missing indicator to handle the small number of missing prerandomization observations."

practically the same as those in Taubman et al. (2014) and serve as a useful baseline for comparing our nonparametric results.

Eliminating sample composition issues, the difference between nonparametric and linear IV results on the extensive margin is likely due to at least two reasons. First, the general random forests directly target and discover substantial treatment effect heterogeneity, as shown above. In the presence of such meaningful heterogeneity, leveraging the individualized treatment effects (some of which suggest negative or zero effects) to estimate the mean effect is likely to yield different average impacts compared to the linear IV method. Second, we use a nonparametric estimator that can better handle the functional form of the dependent variable. The linear IV estimator might yield poor approximations of the average partial effect of an endogenous right-hand-side variable in binary choice models (Wooldridge 2010, Lewbel et al. 2012).

Overall, our nonparametric estimates suggest that Medicaid increased ED use on average, albeit the extensive margin effects are weaker than indicated by the linear method. The heterogeneous Medicaid effects and the divergence in the average results suggest that particular groups of people drive ED usage. The following section investigates the risk factors associated with significant changes in ED usage by examining treatment effects for selected subgroups.

4.3. Risk factors associated with changes in ED use

We summarize treatment effects for selected subgroups by aggregating the individualized treatment effects discussed above. Table 4 presents group average treatment effects for an extended list of characteristics from the causal forest (panel A) and classical subgroup analysis based on the linear IV method (panel B) for comparison.

Table 4 shows that the group treatment effects using both approaches can be materially different. There are two notable findings. First, using the causal forests, we find statistically significant GATE estimates for four subgroups—men, prior SNAP participants, young and middle-aged adults below age 50, and individuals with any previous ED use classified as primary care treatable. These groups exhibit treatment effects between 8 and 10 percentage points, at least twice the average effect. If we zero in on smaller subgroups at the intersection of these characteristics, we find even larger effects (not reported). For instance, the GATE we estimate for men below age 50 with any pre-lottery primary care treatable ED use implies a 21 percentage point increase in the likelihood of using the emergency department upon gaining coverage.

The forests do not reveal statistically significant group effects for other lottery list variables or pre-lottery ED utilization variables. The only subgroups with sizable negative effects are older adults (aged 50 and above) and those with any pre-lottery emergent, preventable ED visits. The effect for older adults is a 6-percentage-point reduction in the probability of using the ED. However, the estimate is insignificant (p-value=0.24), possibly due to a lack of power. Further analyses in this direction suggest that those who reduce inpatient visits are mostly older people—that is, individuals in the left tail of Figures 1c and 2a are older.

Second, we find important differences between the GATE based on the linear model and causal forests. While the linear method uncovers heterogeneity based on the four main groups identified by the nonparametric approach, the former yields a slightly bigger estimate for men (12 percentage points), a slightly smaller estimate for prior SNAP recipients (8 percentage points), a somewhat bigger estimate for people in the younger age group (10 percentage points), and an almost identical group effect for those with prior primary care treatable ED use (12 percentage points). Contrary to the causal forest, the linear method estimates statistically significant treatment effects for all binary indicators of prior ED use and those who did not receive TANF in the pre-lottery period. The same finding holds for some lottery list variables (e.g., households who provide a phone number on lottery sign-up, households who requested English language materials, and those who did not sign up in the first week). The linear IV method's subgroup effects are likely spurious, arising because subgroup analysis can miss meaningful non-linear relationships and are subject to the pitfalls of ex-post specification searching (e.g., Fink et al. 2014).

Figure 4 contrasts the distributions for the four main groups identified by the causal forest in more detail. The graphs highlight the differences in the empirical effect distribution that translate into the group average effects. Panel (a) shows a noticeable difference in the empirical distribution

Table 4: Group average treatment effects of Medicaid on the propensity of ED use

	Panel A: GRF estimates			Panel B: Linear estimates			
Group	GATE	SE	*p*-value	GATE	SE	*p*-value	% N
Female:	0.00	0.04	0.97	0.03	0.04	0.34	0.55
Male:	0.10	0.04	0.01	0.12	0.03	0.00	0.45
Gave phone number: No	0.00	0.08	0.99	0.03	0.08	0.66	0.13
Gave phone number: Yes	0.05	0.03	0.11	0.08	0.03	0.00	0.87
English as preferred language: No	0.01	0.07	0.88	0.02	0.06	0.73	0.14
English as preferred language: Yes	0.05	0.03	0.12	0.08	0.03	0.01	0.86
First week sign-up: No	0.04	0.04	0.31	0.09	0.03	0.01	0.62
First week sign-up: Yes	0.06	0.04	0.20	0.06	0.04	0.16	0.38
Pre-lottery SNAP recipient: No	-0.01	0.05	0.79	0.02	0.05	0.62	0.46
Pre-lottery SNAP recipient: Yes	0.09	0.03	0.01	0.08	0.03	0.01	0.54
Pre-lottery TANF recipient: No	0.04	0.03	0.17	0.07	0.03	0.00	0.98
Pre-lottery TANF recipient: Yes	0.05	0.18	0.79	0.08	0.37	0.82	0.02
Age ≥ 50: No	0.08	0.03	0.02	0.10	0.03	0.00	0.75
Age ≥ 50: Yes	-0.06	0.05	0.24	0.01	0.04	0.77	0.25
Two+ household members on lottery list: No	0.03	0.03	0.42	0.05	0.03	0.06	0.80
Two+ household members on lottery list: Yes	0.09	0.06	0.15	0.19	0.07	0.00	0.20
Any pre-lottery ED visit No	0.03	0.03	0.34	0.07	0.03	0.02	0.69
Any pre-lottery ED visit: Yes	0.06	0.05	0.24	0.07	0.04	0.08	0.31
Any pre-lottery on-hours ED visit: No	0.03	0.03	0.35	0.07	0.03	0.01	0.77
Any pre-lottery on-hours ED visit: Yes	0.07	0.05	0.18	0.08	0.05	0.12	0.23
Any pre-lottery off-hours ED visit: No	0.04	0.03	0.21	0.06	0.03	0.03	0.81
Any pre-lottery off-hours ED visit: Yes	0.06	0.06	0.29	0.07	0.05	0.16	0.19
Any pre-lottery emergent, non-preventable ED visit: No	0.04	0.03	0.19	0.07	0.03	0.01	0.87
Any pre-lottery emergent, non-preventable ED visit: Yes	0.06	0.07	0.40	0.07	0.06	0.28	0.13
Any pre-lottery emergent, preventable ED visit: No	0.05	0.03	0.10	0.09	0.03	0.00	0.92
Any pre-lottery emergent, preventable ED visit: Yes	-0.05	0.08	0.55	0.02	0.07	0.83	0.08
Any pre-lottery primary care treatable ED visit: No	0.02	0.03	0.59	0.06	0.03	0.02	0.81
Any pre-lottery primary care treatable ED visit: Yes	0.13	0.06	0.03	0.12	0.05	0.03	0.19
Any pre-lottery non-emergent ED visit: No	0.03	0.03	0.34	0.07	0.03	0.01	0.86
Any pre-lottery non-emergent ED visit: Yes	0.08	0.07	0.21	0.09	0.06	0.17	0.14

Notes: This table reports the group average treatment effects of Medicaid based on generalized random forests in Panel A and those based on a linear IV method's subsample analysis in Panel B. The first row reproduces the overall average effect. The baseline sample consists of 24,615 individuals in the Taubman et al. (2014) sample with non-missing information on pre-lottery emergency department utilization and SNAP/TANF receipt.

of the individualized treatment effects by pre-lottery SNAP receipt. The distribution of the ITEs for SNAP recipients is narrower and exhibits a denser mass in the positive effects region. This difference in the distribution of the heterogeneous effects reflects in the aggregated group average effect. Previous SNAP recipients are estimated to increase ED use by a statistically significant 9 percentage points, while there is no effect for SNAP non-recipients.

Interestingly, panel (b) of Figure 4 shows that the ITE distribution by pre-lottery primary care treatable ED use exhibits a similar pattern to prior SNAP receipt, but with a bigger group effect for those with any prior visit of 13 percentage points (p-value = 0.03). Finally, Panels (c) and (d) display the heterogeneous effects for age group (age $<$ 50 vs. age $\geq$ 50 years) and gender (male vs. female). In contrast to the previous groups, the distributions for partitioning these groups are similar in shape but shifted to the right for the younger age group and men.

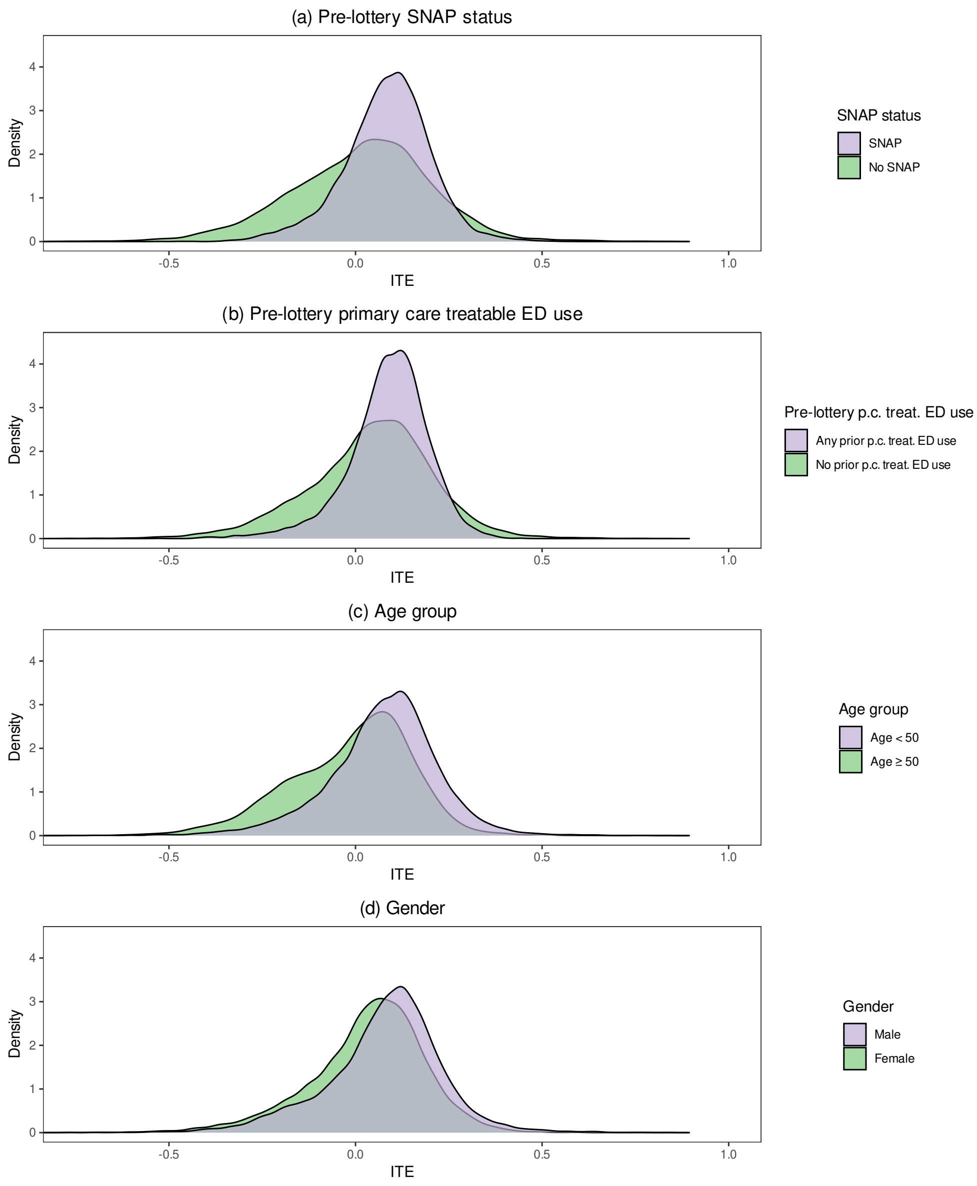
Finally, we divide the sample based on the individualized treatment effects to contrast the characteristics of those who increase and decrease usage ex-post. Table 5 presents the mean characteristics of individuals identified to increase and reduce emergency department utilization on the extensive margin.[6] In almost all cases, we find that those who increase and decrease are different on observed dimensions. To highlight a few key differences, those estimated to increase the likelihood of any ED use are almost three years younger (39.8 vs. 41.5 years), less likely to be female (0.51 vs. 0.63), more likely to have previously received SNAP (0.62 vs. 0.35) with higher total household benefits averaging ($1,607 vs. $717), and more likely to have been a previous TANF recipient (0.03 vs. 0.01) with higher benefits ($111 vs. $63). In terms of prior ED visits, those estimated to increase use have a higher number of visits across all visits. The largest mean differences are for the overall number of ED visits and outpatient ED visits.

4.4. Minimizing non-emergent use via assignment rules

This section presents results using the heterogeneous effects for the efficient policy learning exercises described in Section 3.2. Until now, we have focused on the effects of Medicaid coverage. However,

[6]The corresponding summary of characteristics for the intensive margin of ED use are available in Appendix A.

Figure 4: Distribution of individualized treatment effects of Medicaid by selected groups (Any visit)



Notes: This figure plots the individualized treatment effects of Medicaid on any overall ED visit for the four major groups identified with substantial group average effects—pre-lottery SNAP receipt (panel a), pre-lottery primary care treatable ED visit (panel b), age group (panel c), and gender (panel d). The baseline sample consists of 24,615 individuals in the Taubman et al. (2014) sample with non-missing information on pre-lottery emergency department utilization and SNAP/TANF receipt.

Table 5: Characteristics of individuals who increased and decreased ED use

Variable	Increased ED use (1)	Decreased ED use (2)	Difference (2)-(1)
Lottery list characteristics			
Age (years)	38.76	41.48	-2.72***
Gave phone number	0.87	0.87	0.00
English as preferred language	0.87	0.85	0.02***
Female	0.51	0.63	-0.12***
Week of lottery sign up	1.57	1.61	-0.04
Provided P.O. box address	0.02	0.03	-0.01***
Signed up self for lottery	0.89	0.92	-0.03***
Pre-lottery SNAP recipient	0.62	0.35	0.27***
Pre-lottery SNAP benefit amount ($)	1607.33	713.38	893.95***
Pre-lottery TANF recipient	0.03	0.01	0.02***
Pre-lottery TANF benefit amount ($)	111.08	63.09	47.99***
Pre-lottery ED usage			
Number of overall visits	0.88	0.49	0.39***
Number of inpatient visits	0.09	0.07	0.02***
Number of outpatient visits	0.78	0.42	0.36***
Number of on-hours visits	0.51	0.30	0.21***
Number of off-hours visits	0.37	0.20	0.17***
Number of emergent, non-preventable visits	0.18	0.09	0.09***
Number of emergent, preventable visits	0.07	0.04	0.03***
Number of primary care treatable visits	0.30	0.17	0.13***
Number of non-emergent visits	0.18	0.10	0.08***
Number ambulatory-care-sensitive visits	0.05	0.04	0.01***
Number of visits (chronic conditions)	0.14	0.11	0.03***
Number of visits (injury)	0.20	0.09	0.11***
Number of visits (skin conditions)	0.05	0.03	0.02***
Number of visits (abdominal pain)	0.04	0.02	0.02***
Number of visits (back pain)	0.04	0.02	0.02***
Number of visits (chest pain)	0.02	0.02	0.00***
Number of visits (headache)	0.03	0.02	0.01***
Number of visits (mood disorders)	0.02	0.02	0.00
Number of visits (psychiatric conditions)	0.06	0.05	0.01*
Sum of total ED charges	1013.28	584.41	428.87***
N	16816	7797	24613

Notes: This table reports the means of individual characteristics and pre-randomization ED use for those estimated to increase and decrease ED use upon receiving Medicaid coverage based on the generalized random forest ITE estimates. The sample consists of 24,613 individuals in the Taubman et al. (2014) sample with non-missing information on pre-lottery emergency department utilization and SNAP/TANF receipt.

we conduct the policy learning exercises using the intent-to-treat effects of winning the lottery. Our objective is to estimate an optimal allocation scheme and a simple assignment rule that prioritizes those least likely to increase undesirable ED use, defined as the number of non-emergent ED visits. We first characterize how the sample selected by such an optimal allocation scheme would compare to the lottery sample. We then estimate a simple decision rule to select such a sample as an alternative to a lottery. Of course, a straightforward decision rule might be to identify people based on their predicted individualized treatment effects, such as those with $\widehat{\tau}(X) \leq 0$. However, Athey & Imbens (2019) point out that such decision rules are not always optimal in the sense of minimizing the loss from not using the ideal assignment rule or policy.

Following Kitagawa & Tetenov (2018), we first compute the optimal treatment allocation scheme that minimizes the objective function of the number of non-emergent ED utilization using integer programming. We constrain the number of people selected by the allocation scheme to be the same number that won the Oregon lottery. We then compare the selected sample to the original lottery winners. As expected in Table 6, the selected sample based on the optimal allocation scheme has substantially lower ex-post ED utilization for non-emergent causes (0.18 fewer visits). However, other ED use measures following treatment are also substantially lower. The same finding holds for pre-lottery ED use, suggesting that ED use patterns are complex and treatment effects are driven by a population that uses the ED for most of their health care needs, regardless of the type or severity of the medical condition. Comparing the magnitude of pre-and post-randomization ED use outcomes within the same category suggests larger differences between the two samples, indicating that the estimated optimal assignment scheme successfully selects those who react to obtaining coverage by reducing ED usage. For instance, the difference in outpatient use between the two samples in the pre-lottery period is 0.24 visits, while the same difference in the post-lottery period is 0.51 visits.

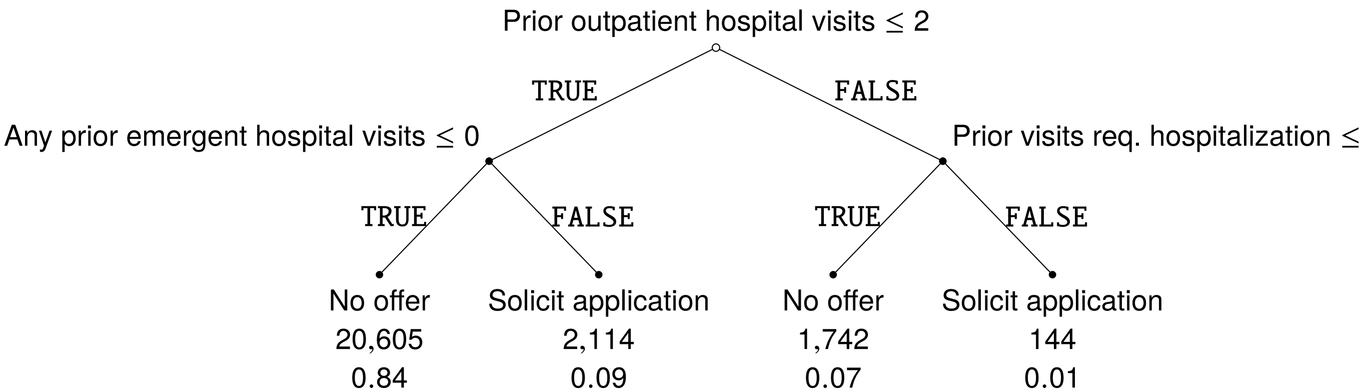
Next, we estimate the optimal assignment rule as a depth-2 decision tree (Athey & Wager 2021). Again, we set up the decision rule estimation to prioritize those least likely to increase non-emergent ED visits. We only use pre-lottery ED use variables to estimate the decision rule presented in Figure 5. The first node of the decision tree partitions the sample based on a threshold

Table 6: Individuals selected by lottery and the optimal linear program

Variable	Selected by lottery (1)	Selected by optimal linear program (2)	Difference (2)-(1)
Lottery list and baseline characteristics			
Age (years)	39.50	40.16	0.66***
Gave phone number	0.88	0.87	-0.01
English as preferred language	0.85	0.79	-0.06***
Female	0.54	0.53	-0.01
Week of lottery sign up	1.55	1.66	0.11***
Provided P.O. box address	0.03	0.03	0.00
Signed up self for lottery	0.85	0.85	0.00
Pre-lottery SNAP recipient	0.53	0.43	-0.10***
Pre-lottery SNAP benefit amount ($)	1376.53	1120.72	-255.81***
Pre-lottery TANF recipient	0.02	0.02	0.00***
Pre-lottery TANF benefit amount ($)	100.09	72.82	-27.27***
Number of overall visits	0.71	0.44	-0.27***
Pre-randomization ED use			
Number of inpatient visits	0.08	0.05	-0.03***
Number of outpatient visits	0.63	0.39	-0.24***
Number of on-hours visits	0.41	0.25	-0.16***
Number of off-hours visits	0.30	0.19	-0.11***
Number of emergent, non-preventable visits	0.15	0.09	-0.06***
Number of emergent, preventable visits	0.05	0.03	-0.02***
Number of primary care treatable visits	0.25	0.15	-0.10***
Number of non-emergent visits	0.14	0.08	-0.06***
Number ambulatory-care-sensitive visits	0.05	0.03	-0.02***
Number of visits (chronic conditions)	0.12	0.08	-0.04***
Number of visits (injury)	0.17	0.10	-0.07***
Number of visits (skin conditions)	0.04	0.03	-0.01***
Number of visits (abdominal pain)	0.03	0.02	-0.01***
Number of visits (back pain)	0.03	0.02	-0.01***
Number of visits (chest pain)	0.02	0.01	-0.01***
Number of visits (headache)	0.02	0.01	-0.01***
Number of visits (mood disorders)	0.02	0.01	-0.01**
Number of visits (psychiatric conditions)	0.05	0.03	-0.02***
Sum of total ED charges	808.36	504.40	-303.96***
Number of overall visits	0.98	0.43	-0.55***
ED use outcomes			
Number of inpatient visits	0.10	0.06	-0.04***
Number of outpatient visits	0.87	0.36	-0.51***
Number of emergent, non-preventable visits	0.20	0.10	-0.10***
Number of emergent, preventable visits	0.07	0.04	-0.03***
Number of primary care treatable visits	0.35	0.17	-0.18***
Number of non-emergent visits	0.20	0.02	-0.18***
N	9607	9607	

Notes: This table reports the means of individual characteristics, pre-randomization ED use and post-treatment outcomes for those selected by the lottery and selected by the optimal treatment assignment solution based on the optimal linear program which minimizes non-emergent ED use under the restriction that the same number of people are treated as within the lottery. The sample consists of 24,605 individuals in the Taubman et al. (2014) sample with non-missing information on pre-lottery emergency department utilization and SNAP/TANF receipt.

Figure 5: Optimal depth-2 policy tree for the number of non-emergent ED visits



Notes: This figure shows the optimal policy assignment rule obtained by optimizing the doubly robust scores of the conditional average treatment effect (CATE) of being randomly selected as a lottery winner in the Oregon Health Insurance Experiment using the efficient policy learning framework of Athey & Wager (2021). The sample consists of 24,615 individuals in the Taubman et al. (2014) sample with non-missing information on pre-lottery emergency department utilization and SNAP/TANF receipt.

of two prior outpatient ED visits. On the one hand, for people with two or fewer outpatient ED visits, the decision rule selects them if they had any previous ED use classified as emergent; otherwise, no applications are solicited from them. This assignment rule is interesting because, given the objective of minimizing non-emergent use, it prioritizes those with prior emergent use (usually the desirable use of emergency departments).

On the other hand, for those with more than two previous outpatient ED visits (potentially excessive ED use), only those with more than one prior ED visit resulting in hospitalization are selected. Again, this rule achieves the objective of minimizing non-emergent ED use and does so reasonably by prioritizing those who have presented with conditions severe enough to warrant inpatient care in the past.

We emphasize that our learned decision rule is not set in stone and may be impractical, especially if the previous ED utilization variables are not easily accessible. Policymakers may also want to specify a different objective function to align with their goals or accommodate additional constraints. Alternative welfare functions could place extra weight on targeting people with severe conditions who increase non-preventable ED use. Further capacity constraints can restrict the maximum allowable cost. Moreover, the set of allowable variables considered in the estimation of the decision rule can also be modified to respect fairness or other political/legal requirements.

Optimal policy rule estimation presents an alternative approach for policy administrators who are faced with constraints in allocating scarce resources, especially with public assistance programs relative to the random or automatic assignment status quo. Even when program spots are not scarce, optimal policy learning might help allocate recipients to different insurance plan options in programs such as Medicaid, where managed care is on the rise. About 25 million people will be auto-assigned into a Medicaid managed care plan as states continue contracting with risk-based managed care organizations to deliver benefits to Medicaid recipients (Ndumele & Wallace 2020). Policymakers have an opportunity to utilize advances in statistical treatment assignment rule estimation to assign people to health care plans to improve health and reduce spending.

5. Conclusion

This paper estimates the heterogeneous impacts of Medicaid coverage on emergency department visits using records from 12 hospitals in Portland, Oregon, matched to the Oregon Health Insurance Experiment. The headline result that Medicaid increased ED utilization in the 2008 Oregon Health Experiment was surprising, given the conventional insight that insurance coverage should make it easier for recipients to access primary care and reduce the need to use the emergency department. We provide new insights into the ED results of the Oregon experiment by estimating individualized treatment effects of Medicaid using nonparametric machine learning methods. We then leverage the heterogeneous effects to estimate simple optimal assignment rules to illustrate their usefulness for policymakers in similar settings.

We find substantial treatment effect heterogeneity in the impacts of Medicaid on emergency department utilization. The individualized treatment effects of Medicaid on ED use indicate positive and negative effects, suggesting a more nuanced interpretation of Medicaid's average impacts. We also find that coverage effects for different types of ED use exhibit meaningful heterogeneity. Due to the widely dispersed individualized effect distribution, the average treatment effect does not represent the individualized treatment effect. A small proportion of high-use individuals drive the positive (average) Medicaid effect. On the extensive margin, we estimate a weaker Medicaid effect on the probability of any overall ED use of a 4.5 percentage points increase, which is statistically insignificant and about 65 percent of the linear IV estimation method's magnitude. Despite the positive or null average effects, the predicted treatment effects for a considerable fraction of individuals are negative. We also find substantial treatment effect heterogeneity on the intensive margin. Overall, increased ED usage is driven by intensive margin effects rather than extensive margin effects—individuals who already utilize the emergency department further increase their usage.

Our findings suggest that the average effect sometimes masks countervailing forces. Although many people increase ED usage due to obtaining coverage in some instances, most people either decrease usage or do not respond. We find that reductions in ED use are pronounced for inpatient visits and emergent, non-preventable conditions. The overall null effects for these outcomes hide substantial opposing effects of obtaining coverage.

We also characterize the groups of people mainly driving the positive average Medicaid effects in the right tail of the distribution of the heterogeneous effects. Those with positive heterogeneous effects are predominantly younger, more likely to be men, more likely to receive SNAP and TANF in the pre-lottery period, and more likely to have higher baseline ED use. In particular, we identify four groups estimated to have statistically significant increases in ED use that are at least twice as large as the magnitude of the average effect. These groups are men, prior SNAP participants, younger adults less than 50 years old, and people with any pre-lottery ED use classified as primary care treatable. We also find that the nonparametric approach finds no effect for other groups that the linear model's subsample analysis finds statistically significant effects.

Finally, we illustrate how to use the estimated heterogeneous effects to estimate optimal decision rules that prioritize those likely to decrease undesirable emergency department utilization, defined as non-emergent use. Our estimated decision rule invites applications from those with minimal prior outpatient ED use as an alternative to the random assignment. The decision rule also prioritizes those with severe health problems justifying ED care or requiring inpatient treatment. While our decision rules are illustrative, they highlight the potential for using statistical decision rules to guide policymakers to achieve context-specific objectives.

References

Abadie, A. (2003), 'Semiparametric instrumental variable estimation of treatment response models', *Journal of econometrics* **113**(2), 231–263.

Allen, H., Baicker, K., Finkelstein, A., Taubman, S., Wright, B. J. & Group, O. H. S. (2010), 'What the oregon health study can tell us about expanding medicaid', *Health Affairs* **29**(8), 1498–1506.

Allen, H., Baicker, K., Taubman, S., Wright, B. & Finkelstein, A. (2013), 'The oregon health insurance experiment: when limited policy resources provide research opportunities', *Journal of health politics, policy and law* **38**(6), 1183–1192.

Angrist, J. D. (2004), 'Treatment effect heterogeneity in theory and practice', *The economic journal* **114**(494), C52–C83.

Athey, S. & Imbens, G. W. (2019), 'Machine learning methods that economists should know about', *Annual Review of Economics* **11**, 685–725.

Athey, S., Tibshirani, J., Wager, S. et al. (2019), 'Generalized random forests', *The Annals of Statistics* **47**(2), 1148–1178.

Athey, S. & Wager, S. (2021), 'Policy learning with observational data', *Econometrica* **89**(1), 133–161.

Baicker, K., Taubman, S. L., Allen, H. L., Bernstein, M., Gruber, J. H., Newhouse, J. P., Schneider, E. C., Wright, B. J., Zaslavsky, A. M. & Finkelstein, A. N. (2013), 'The oregon experiment—effects of medicaid on clinical outcomes', *New England Journal of Medicine* **368**(18), 1713–1722.

Billings, J., Parikh, N. & Mijanovich, T. (2000), 'Emergency room use: the new york story', *Issue Brief (Commonw Fund)* **434**, 1–12.

Chen, C., Scheffler, G. & Chandra, A. (2011), 'Perspective: Massachusetts' health care reform and emergency department utilization', *N Engl J Med.* **365**, e25.

Chernozhukov, V., Chetverikov, D., Demirer, M., Duflo, E., Hansen, C., Newey, W. K. & Robins, J. (2018), 'Double/debiased machine learning for treatment and structural parameters', *The Econometrics Journal* **21**(1), C1–C68.

Deaton, A. (2010), 'Instruments, randomization, and learning about development', *Journal of economic literature* **48**(2), 424–55.

Dehejia, R. H. (2005), 'Program evaluation as a decision problem', *Journal of Econometrics* **125**(1-2), 141–173.

Einav, L., Finkelstein, A., Ryan, S. P., Schrimpf, P. & Cullen, M. R. (2013), 'Selection on moral hazard in health insurance', *American Economic Review* **103**(1), 178–219.

Fink, G., McConnell, M. & Vollmer, S. (2014), 'Testing for heterogeneous treatment effects in experimental data: false discovery risks and correction procedures', *Journal of Development Effectiveness* **6**(1), 44–57.

Finkelstein, A. (2013), 'Oregon health insuranceexperiment public use data', Available at http://www.nber.org/oregon/data.html.

Finkelstein, A. N., Taubman, S. L., Allen, H. L., Wright, B. J. & Baicker, K. (2016), 'Effect of medicaid coverage on ed use—further evidence from oregon's experiment', *N Engl J Med* **375**(16), 1505–1507.

Finkelstein, A., Taubman, S., Wright, B., Bernstein, M., Gruber, J., Newhouse, J. P., Allen, H., Baicker, K. & Group, O. H. S. (2012), 'The oregon health insurance experiment: evidence from the first year', *The Quarterly journal of economics* **127**(3), 1057–1106.

Heckman, J. J., Smith, J. & Clements, N. (1997), 'Making the most out of programme evaluations and social experiments: Accounting for heterogeneity in programme impacts', *The Review of Economic Studies* **64**(4), 487–535.

Hirano, K. & Porter, J. R. (2009), 'Asymptotics for statistical treatment rules', *Econometrica* **77**(5), 1683–1701.

Imbens, G. W., Angrist, J. D. et al. (1994), 'Identification and estimation of local average treatment effects', *Econometrica* **62**(2), 467–475.

Kitagawa, T. & Tetenov, A. (2018), 'Who should be treated? empirical welfare maximization methods for treatment choice', *Econometrica* **86**(2), 591–616.

Kowalski, A. E. (2018), Reconciling seemingly contradictory results from the oregon health insurance experiment and the massachusetts health reform, Working Paper 24647, National Bureau of Economic Research.
URL: *http://www.nber.org/papers/w24647*

Lewbel, A., Dong, Y. & Yang, T. T. (2012), 'Comparing features of convenient estimators for binary choice models with endogenous regressors', *Canadian Journal of Economics/Revue canadienne d'économique* **45**(3), 809–829.

List, J. A., Shaikh, A. M. & Xu, Y. (2019), 'Multiple hypothesis testing in experimental economics', *Experimental Economics* **22**(4), 773–793.

Manski, C. F. (2004), 'Statistical treatment rules for heterogeneous populations', *Econometrica* **72**(4), 1221–1246.

Miller, S. (2012), 'The effect of insurance on emergency room visits: an analysis of the 2006 massachusetts health reform', *Journal of public Economics* **96**(11-12), 893–908.

Moore, B. J. & Liang, L. (2020), Costs of emergency department visits in the united states, 2017, *in* 'Healthcare Cost and Utilization Project (HCUP) Statistical Briefs [Internet]', Agency for Healthcare Research and Quality (US).

Ndumele, C. & Wallace, J. (2020), 'Improving Auto-Assignment in Medicaid Managed Care'.
URL: *https://onepercentsteps.com/policy-briefs/improving-auto-assignment-in-medicaid-managed-care/*

Sexton, J. & Laake, P. (2009), 'Standard errors for bagged and random forest estimators', *Computational Statistics & Data Analysis* **53**(3), 801–811.

Sommers, B. D., Blendon, R. J., Orav, E. J. & Epstein, A. M. (2016), 'Changes in utilization and health among low-income adults after medicaid expansion or expanded private insurance', *JAMA internal medicine* **176**(10), 1501–1509.

Sommers, B. D. & Simon, K. (2017), 'Health insurance and emergency department use-a complex relationship', *The New England journal of medicine* **376**(18), 1708–1711.

Stoye, J. (2009), 'Minimax regret treatment choice with finite samples', *Journal of Econometrics* **151**(1), 70–81.

Stoye, J. (2012), 'Minimax regret treatment choice with covariates or with limited validity of experiments', *Journal of Econometrics* **166**(1), 138–156.

Tang, N., Stein, J., Hsia, R. Y., Maselli, J. H. & Gonzales, R. (2010), 'Trends and characteristics of us emergency department visits, 1997-2007', *Jama* **304**(6), 664–670.

Taubman, S. L., Allen, H. L., Wright, B. J., Baicker, K. & Finkelstein, A. N. (2014), 'Medicaid increases emergency-department use: evidence from oregon's health insurance experiment', *Science* **343**(6168), 263–268.

Wager, S. & Athey, S. (2018), 'Estimation and inference of heterogeneous treatment effects using random forests', *Journal of the American Statistical Association* **113**(523), 1228–1242.

Woodworth, L. (2020), 'Swamped: Emergency department crowding and patient mortality', *Journal of health economics* **70**, 102279.

Wooldridge, J. M. (2010), *Econometric analysis of cross section and panel data*, MIT press.

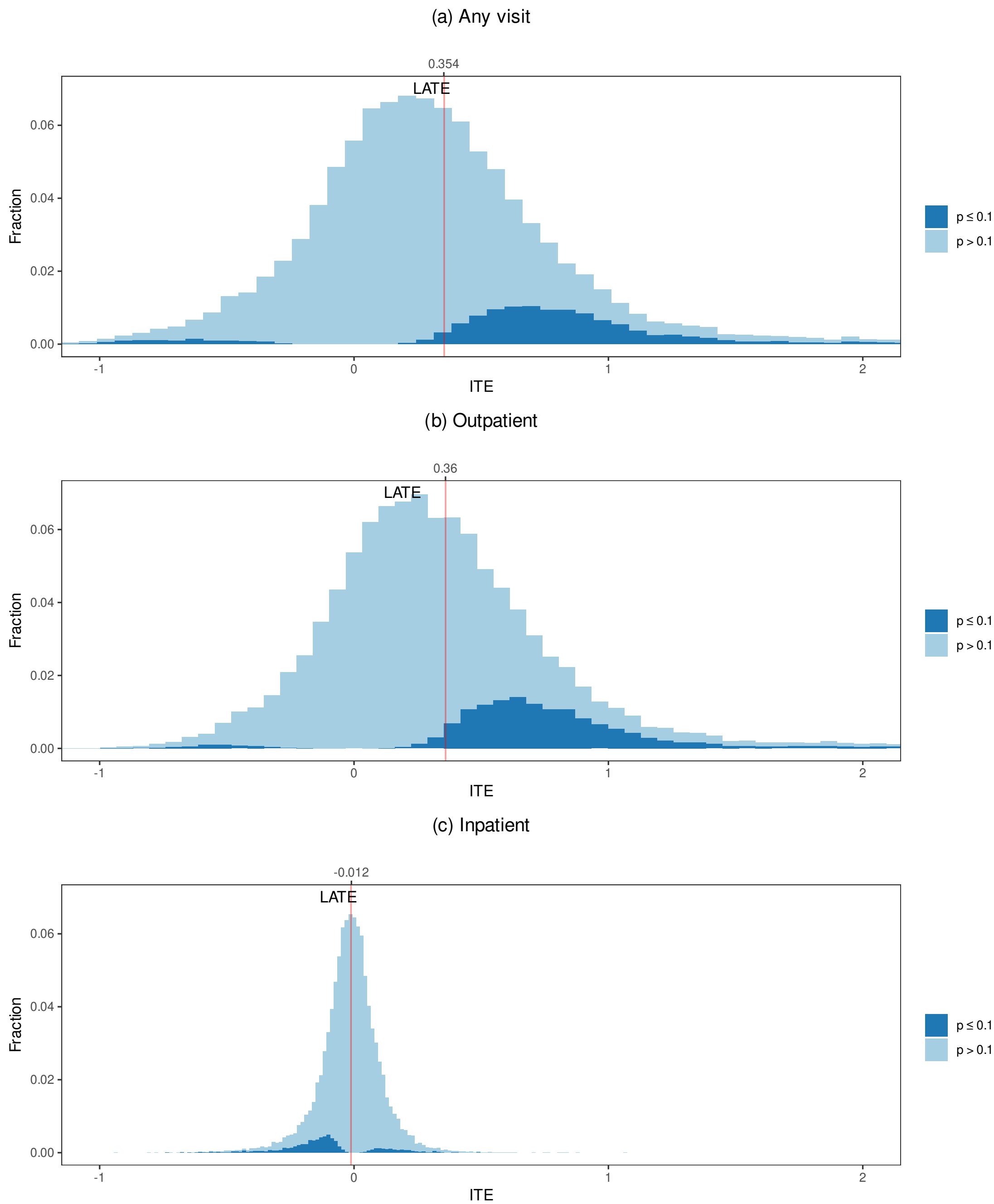
Yardley, W. (2008), 'Drawing lots for health care', *New York Times* p. A12.

Young, A. (2019), 'Channeling fisher: Randomization tests and the statistical insignificance of seemingly significant experimental results*', *The Quarterly Journal of Economics* **134**(2), 557–598.

Zhao, Y., Zeng, D., Rush, A. J. & Kosorok, M. R. (2012), 'Estimating individualized treatment rules using outcome weighted learning', *Journal of the American Statistical Association* **107**(499), 1106–1118.

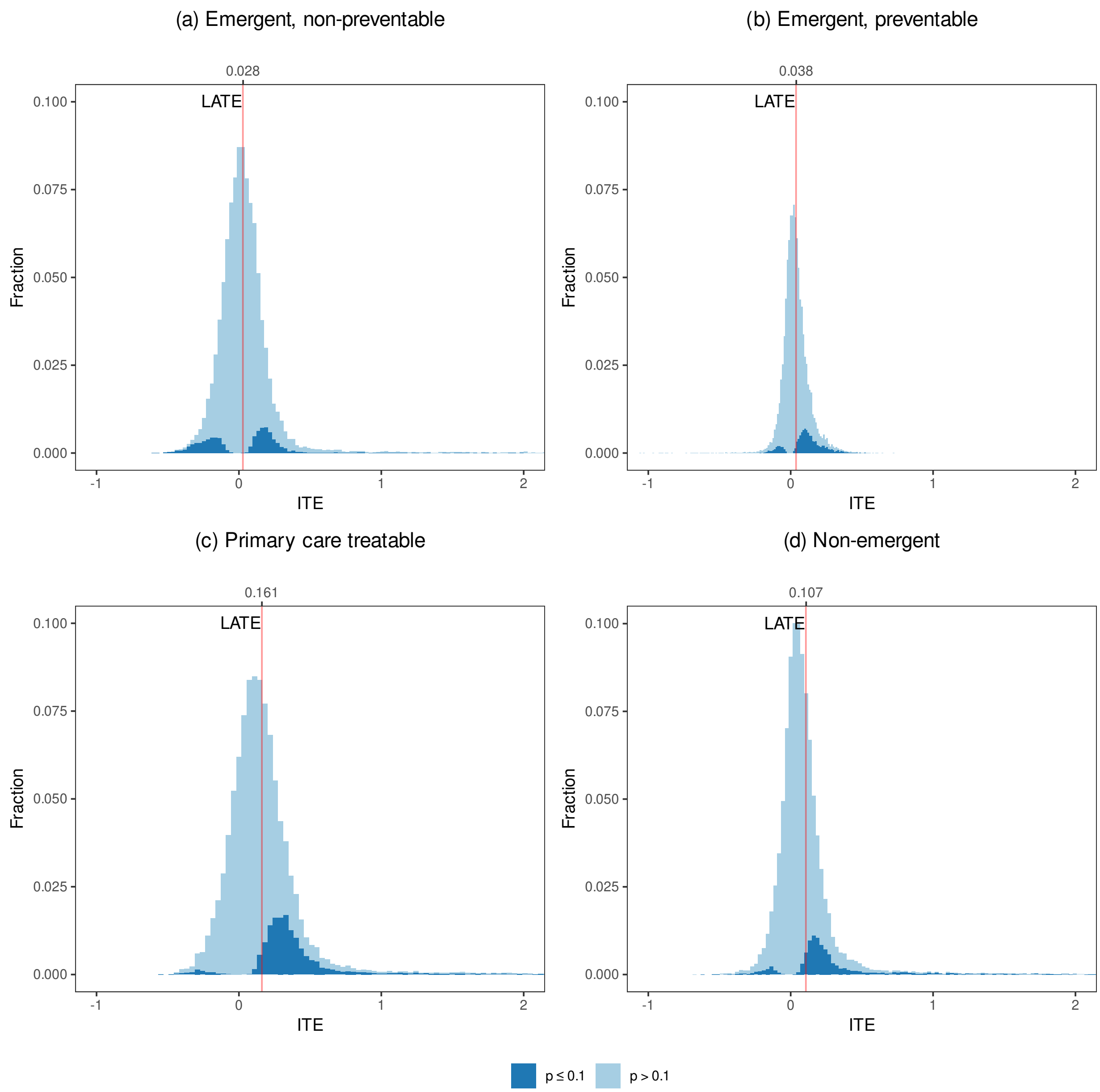
A. Supplementary Material for Medicaid Coverage Effects

Figure A.1: Distribution of individualized treatment effects of Medicaid on the number of ED visits



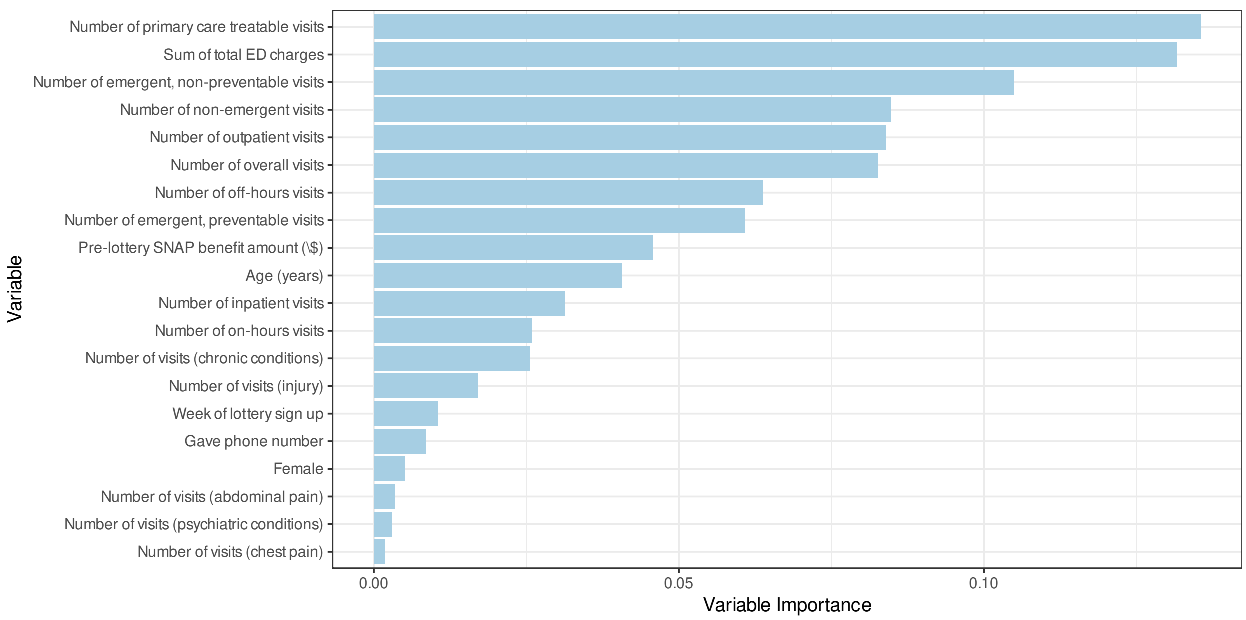
Notes: This figure plots the individualized treatment effects of Medicaid on the number of overall ED visit (panel a), the number of outpatient ED visits (panel b), and the number of inpatient ED use (panel c) based on generalized random forests. The darker shade denotes statistical significance at the 10% level. The red vertical line indicates the local average treatment effect. The baseline sample consists of 24,615 individuals in the Taubman et al. (2014) sample with non-missing information on pre-lottery emergency department utilization and SNAP/TANF receipt. The estimates displayed exclude less than half a percentile at the top and bottom of the distribution, resulting in the axes corresponding approximately to the percentile range [0.5%, 99.5%]. Bin size is chosen according to the Freedman-Diaconis rule.

Figure A.2: Distribution of individualized treatment effects of Medicaid on the number of ED visits by type of condition



Notes: This figure plots the individualized treatment effects of Medicaid by type of ED visit based on generalized random forests for the number of emergent, non-preventable visits (panel a), the number of emergent, preventable visits (panel b) , the number of primary care treatable visits (panel c), and the number of non-emergent visits (panel d). Measures of the type of ED visit are based on Billings et al.'s (2000) algorithm described in Taubman et al. (2014). The number of visits of each type are then obtained by summing the probabilities across all visits for an individual. The darker shade denotes statistical significance at the 10% level. The red vertical line indicates the local average treatment effect. The baseline sample consists of 24,615 individuals in the Taubman et al. (2014) sample with non-missing information on pre-lottery emergency department utilization and SNAP/TANF receipt. The estimates displayed exclude less than half a percentile at the top and bottom of the distribution, resulting in the axes corresponding approximately to the percentile range [0.5%, 99.5%]. Bin size is chosen according to the Freedman-Diaconis rule.

Figure A.3: Variable importance scores in growing causal forest (Number of visits)



Notes This figure shows the variable importance scores of the top 20 characteristics in growing the generalized random forests used to estimate the individualized treatment effects of Medicaid for the number of overall ED visit. The variable importance measure is a simple weighted sum of the proportion of times a variable is used in a tree splitting step at each depth in growing the forest. The scores roughly capture how important a variable is for driving treatment effect heterogeneity. The baseline sample consists of 24,615 individuals in the Taubman et al. (2014) sample with non-missing information on pre-lottery emergency department utilization and SNAP/TANF receipt.

Table A.1: Variable importance for all variables in growing causal forest (overall ED use)

Any visit		Number of visits	
Variable	Importance	Variable	Importance
Pre-lottery SNAP benefit amount ($)	0.18	Number of primary care treatable visits	0.14
Age (years)	0.18	Sum of total ED charges	0.13
Sum of total ED charges	0.09	Number of emergent, non-preventable visits	0.10
Number of emergent, non-preventable visits	0.09	Number of non-emergent visits	0.08
Number of primary care treatable visits	0.08	Number of outpatient visits	0.08
Week of lottery sign up	0.05	Number of overall visits	0.08
Number of non-emergent visits	0.05	Number of off-hours visits	0.06
Female	0.04	Number of emergent, preventable visits	0.06
Number of emergent, preventable visits	0.02	Pre-lottery SNAP benefit amount ($)	0.05
Number of visits (chronic conditions)	0.01	Age (years)	0.04
Number of inpatient visits	0.01	Number of inpatient visits	0.03
Pre-lottery SNAP recipient	0.01	Number of on-hours visits	0.03
Number of outpatient visits	0.01	Number of visits (chronic conditions)	0.03
Number of off-hours visits	0.01	Number of visits (injury)	0.02
Number of on-hours visits	0.01	Week of lottery sign up	0.01
Number of overall visits	0.01	Gave phone number	0.01
Signed up self for lottery	0.01	Female	0.01
Gave phone number	0.01	Number of visits (abdominal pain)	0.00
Number of visits (injury)	0.01	Number of visits (psychiatric conditions)	0.00
English as preferred language	0.00	Number of visits (chest pain)	0.00
Pre-lottery TANF benefit amount ($)	0.00	Pre-lottery TANF benefit amount ($)	0.00
Number of visits (psychiatric conditions)	0.00	Number of visits (back pain)	0.00
Number of visits (mood disorders)	0.00	Number of visits (skin conditions)	0.00
Number ambulatory-care-sensitive visits	0.00	Pre-lottery SNAP recipient	0.00
Number of visits (skin conditions)	0.00	Number ambulatory-care-sensitive visits	0.00
Number of visits (abdominal pain)	0.00	Number of visits (headache)	0.00
Number of visits (back pain)	0.00	Number of visits (mood disorders)	0.00
Pre-lottery TANF recipient	0.00	Signed up self for lottery	0.00
Number of visits (chest pain)	0.00	Pre-lottery TANF recipient	0.00
Number of visits (headache)	0.00	English as preferred language	0.00
Provided P.O. box address	0.00	Provided P.O. box address	0.00

Notes: This table shows the top variable importance scores of all characteristics for growing the generalized random forests used to estimate the ITE of Medicaid coverage for overall ED visits. The variable importance measure is a simple weighted sum of the proportion of times a variable is used in a splitting step at each depth in growing the causal forest, thus, capturing how important a variable is for driving treatment effect heterogeneity. The sample consists of 24,615 individuals in the Taubman et al. (2014) sample with non-missing information on pre-lottery emergency department utilization and SNAP/TANF receipt.

Table A.2: Group average treatment effects of Medicaid on the number of ED visits

	Panel A: GRF estimates			Panel B: Linear estimates			
Group	GATE	SE	p-value	GATE	SE	p-value	% N
Female:	0.27	0.16	0.08	0.23	0.18	0.21	0.55
Male:	0.43	0.17	0.01	0.46	0.18	0.01	0.45
Gave phone number: No	-0.06	0.33	0.86	-0.26	0.38	0.50	0.13
Gave phone number: Yes	0.42	0.12	0.00	0.42	0.14	0.00	0.87
English as preferred language: No	0.14	0.17	0.41	0.05	0.18	0.78	0.14
English as preferred language: Yes	0.39	0.13	0.00	0.36	0.14	0.01	0.86
First week sign-up: No	0.38	0.14	0.01	0.48	0.16	0.00	0.62
First week sign-up: Yes	0.29	0.19	0.12	0.13	0.21	0.54	0.38
Pre-lottery SNAP recipient: No	0.21	0.15	0.16	0.17	0.18	0.34	0.46
Pre-lottery SNAP recipient: Yes	0.45	0.17	0.01	0.34	0.17	0.04	0.54
Pre-lottery TANF recipient: No	0.34	0.11	0.00	0.32	0.13	0.01	0.98
Pre-lottery TANF recipient: Yes	0.93	1.21	0.44	1.46	2.44	0.55	0.02
Age ≥ 50: No	0.43	0.14	0.00	0.38	0.16	0.02	0.75
Age ≥ 50: Yes	0.05	0.20	0.81	0.25	0.21	0.22	0.25
Two+ household members on lottery list: No	0.32	0.14	0.02	0.27	0.15	0.07	0.80
Two+ household members on lottery list: Yes	0.52	0.16	0.00	0.74	0.23	0.00	0.20
Any pre-lottery ED visit No	0.22	0.09	0.01	0.25	0.08	0.00	0.69
Any pre-lottery ED visit: Yes	0.67	0.32	0.03	0.45	0.31	0.15	0.31
Any pre-lottery on-hours ED visit: No	0.19	0.09	0.04	0.26	0.09	0.00	0.77
Any pre-lottery on-hours ED visit: Yes	0.87	0.41	0.03	0.67	0.42	0.11	0.23
Any pre-lottery off-hours ED visit: No	0.22	0.09	0.02	0.21	0.09	0.03	0.81
Any pre-lottery off-hours ED visit: Yes	0.95	0.45	0.04	0.56	0.45	0.21	0.19
Any pre-lottery emergent, non-preventable ED visit: No	0.22	0.09	0.02	0.28	0.09	0.00	0.87
Any pre-lottery emergent, non-preventable ED visit: Yes	1.15	0.63	0.07	0.69	0.63	0.27	0.13
Any pre-lottery emergent, preventable ED visit: No	0.25	0.09	0.01	0.32	0.09	0.00	0.92
Any pre-lottery emergent, preventable ED visit: Yes	1.50	0.96	0.12	1.31	0.90	0.15	0.08
Any pre-lottery primary care treatable ED visit: No	0.18	0.09	0.04	0.26	0.09	0.00	0.81
Any pre-lottery primary care treatable ED visit: Yes	1.03	0.46	0.03	0.72	0.47	0.13	0.19
Any pre-lottery non-emergent ED visit: No	0.27	0.09	0.00	0.34	0.09	0.00	0.86
Any pre-lottery non-emergent ED visit: Yes	0.82	0.61	0.18	0.26	0.63	0.68	0.14

Notes: This table reports the group average treatment effects of Medicaid based on generalized random forests (GRF) in Panel A and those based on the linear IV method's subsample analysis in Panel B. The first row reproduces the overall average effect. The baseline sample consists of 24,615 individuals in the Taubman et al. (2014) sample with non-missing information on pre-lottery emergency department utilization and SNAP/TANF receipt.

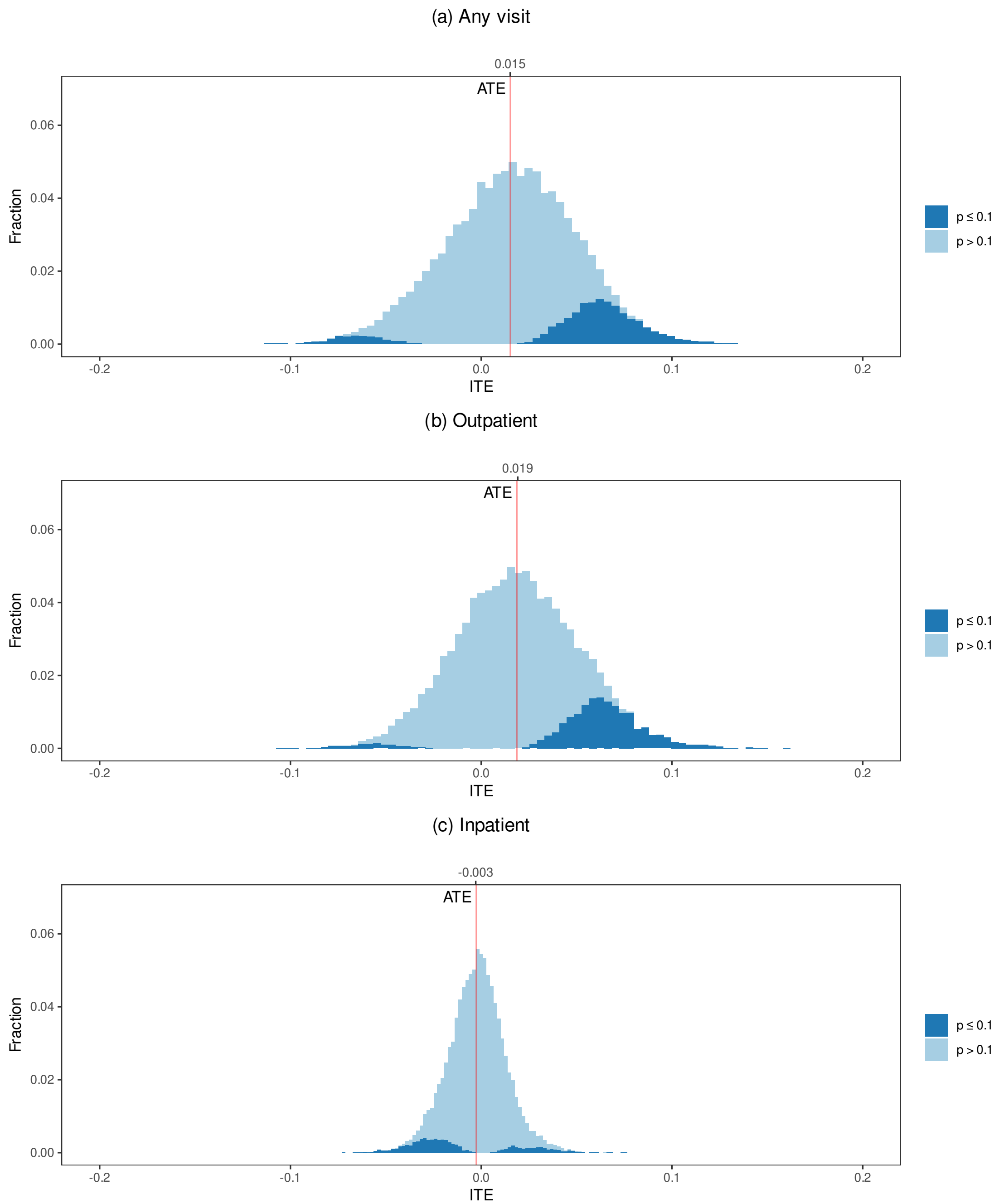
Table A.3: Characteristics of individuals who increased and decreased ED use (Number of visits)

Variable	Increased ED use	Decreased ED use	Difference
Lottery list characteristics			
Age (years)	38.54	42.88	-4.34***
Gave phone number	0.86	0.89	-0.03***
English as preferred language	0.87	0.84	0.03***
Female	0.53	0.58	-0.05***
Week of lottery sign up	1.59	1.56	0.03
Provided P.O. box address	0.02	0.03	-0.01***
Signed up self for lottery	0.89	0.92	-0.03***
Pre-lottery SNAP recipient	0.59	0.36	0.23***
Pre-lottery SNAP benefit amount ($)	1497.29	817.99	679.30***
Pre-lottery TANF recipient	0.02	0.02	0.00***
Pre-lottery TANF benefit amount ($)	96.79	94.84	1.95
Pre-lottery ED usage			
Number of overall visits	0.90	0.30	0.60***
Number of inpatient visits	0.10	0.06	0.04***
Number of outpatient visits	0.80	0.25	0.55***
Number of on-hours visits	0.52	0.19	0.33***
Number of off-hours visits	0.38	0.11	0.27***
Number of emergent, non-preventable visits	0.18	0.06	0.12***
Number of emergent, preventable visits	0.07	0.03	0.04***
Number of primary care treatable visits	0.32	0.08	0.24***
Number of non-emergent visits	0.18	0.07	0.11***
Number ambulatory-care-sensitive visits	0.05	0.03	0.02***
Number of visits (chronic conditions)	0.15	0.07	0.08***
Number of visits (injury)	0.20	0.05	0.15***
Number of visits (skin conditions)	0.06	0.01	0.05***
Number of visits (abdominal pain)	0.04	0.01	0.03***
Number of visits (back pain)	0.04	0.02	0.02***
Number of visits (chest pain)	0.02	0.01	0.01***
Number of visits (headache)	0.03	0.01	0.02***
Number of visits (mood disorders)	0.03	0.01	0.02***
Number of visits (psychiatric conditions)	0.07	0.03	0.04***
Sum of total ED charges	1027.60	387.86	639.74***
N	18581.00	6018.00	24599

Notes: This table reports the means of individual characteristics and pre-randomization ED use for those estimated to increase and decrease ED use upon receiving Medicaid coverage based on the causal forest CATE estimates. ED use is measured as the number of ED visits. Panel A reports the means for the full sample while Panel B is limited to effects significant at the 10% level. The sample consists of 24,615 individuals in the Taubman et al. (2014) sample with non-missing information on pre-lottery emergency department utilization and SNAP/TANF receipt.

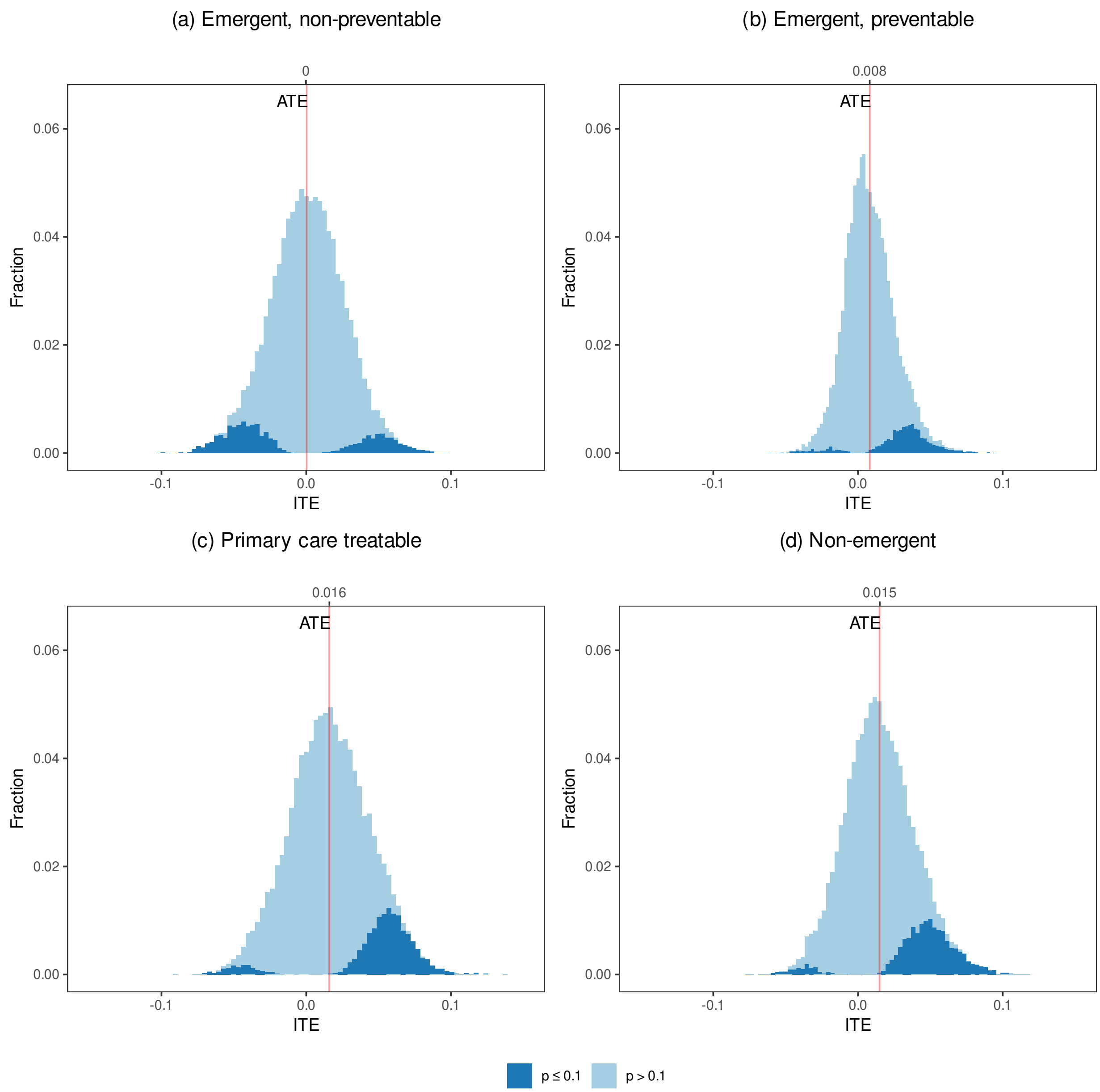
B. Supplementary Material for Intent-to-treat effects

Figure B.1: Distribution of individualized treatment effects of winning the lottery on the propensity of ED use



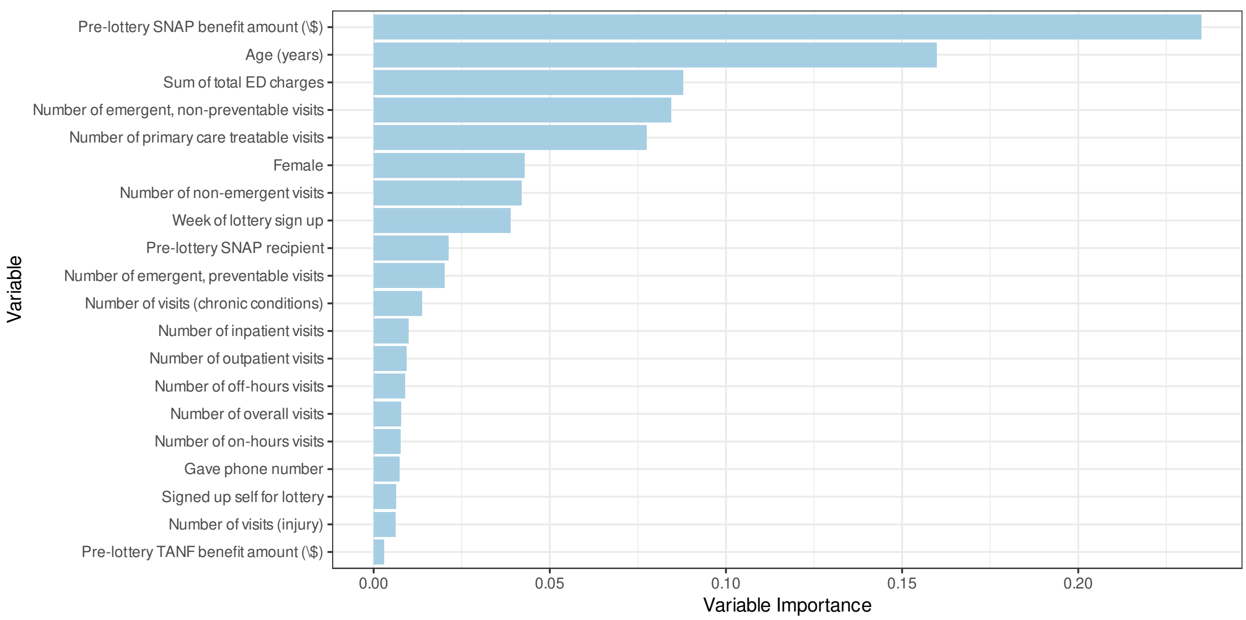
Notes: This figure plots the individualized treatment effects of winning the lottery (and being invited to apply for Medicaid) on any overall ED visit (panel a), any outpatient ED visits (panel b), and any inpatient ED use (panel c) based on generalized random forests. The darker shade denotes statistical significance at the 10% level. The red vertical line indicates the local average treatment effect. The baseline sample consists of 24,615 individuals in the Taubman et al. (2014) sample with non-missing information on pre-lottery emergency department utilization and SNAP/TANF receipt. The estimates displayed exclude less than half a percentile at the top and bottom of the distribution, resulting in the axes corresponding approximately to the percentile range [0.5%, 99.5%]. Bin size is chosen according to the Freedman-Diaconis rule.

Figure B.2: Distribution of individualized treatment effects of winning the lottery on the propensity of ED use by type of condition



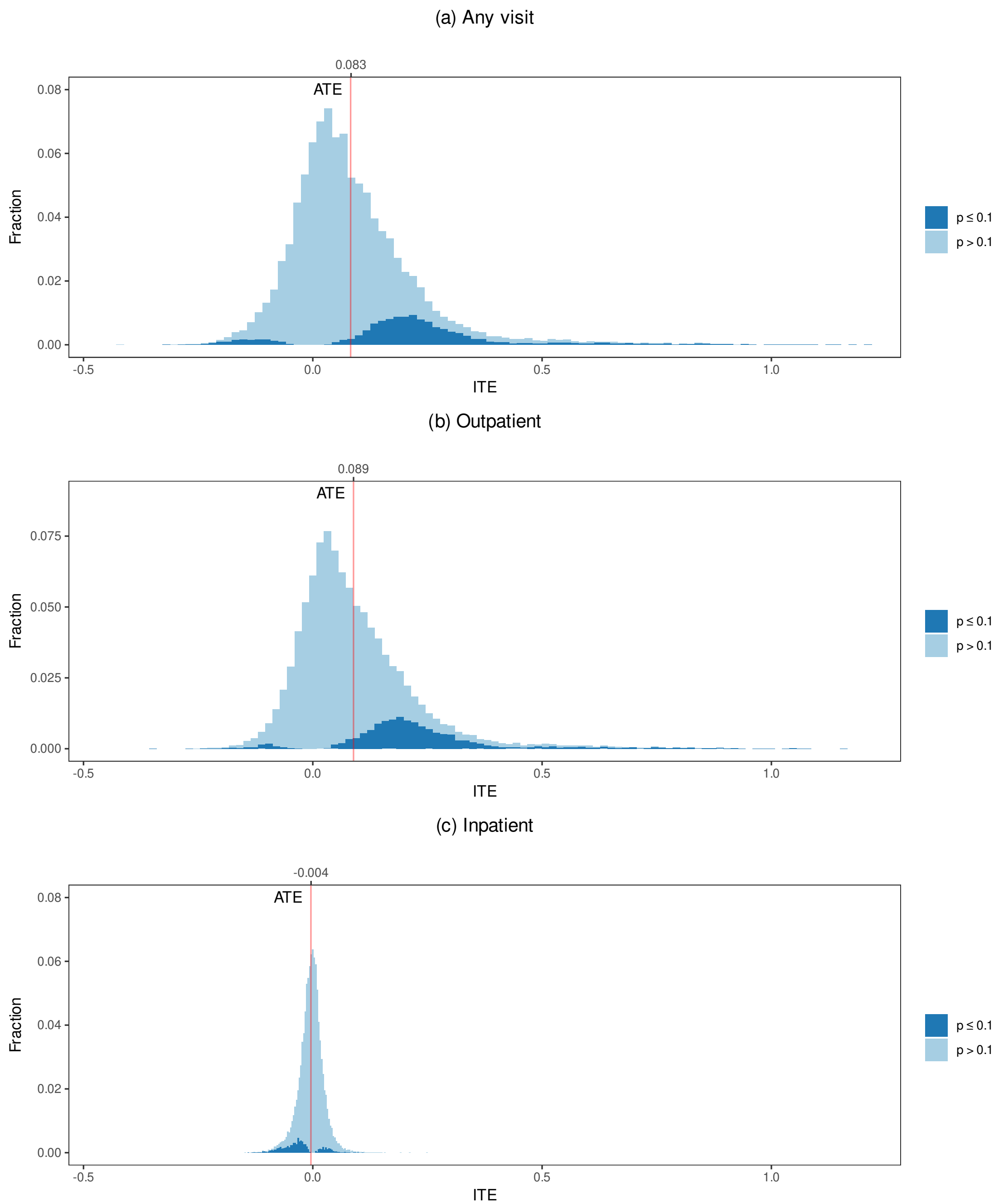
Notes: This figure plots the individualized treatment effects of winning the lottery (and being invited to apply for Medicaid) by type of ED visit based on generalized random forests for any emergent, non-preventable visit (panel a), any emergent, preventable visit (panel b), any primary care treatable visit (panel c), and any non-emergent visit (panel d). Measures of the type of ED visit are based on Billings et al.'s (2000) algorithm described in Taubman et al. (2014). We use these measures to construct binary indicators of ED visits by type of condition as described in the main text. The darker shade denotes statistical significance at the 10% level. The red vertical line indicates the local average treatment effect. The baseline sample consists of 24,615 individuals in the Taubman et al. (2014) sample with non-missing information on pre-lottery emergency department utilization and SNAP/TANF receipt. The estimates displayed exclude less than half a percentile at the top and bottom of the distribution, resulting in the axes corresponding approximately to the percentile range [0.5%, 99.5%]. Bin size is chosen according to the Freedman-Diaconis rule.

Figure B.3: Variable importance scores in growing causal forest (Any visit)



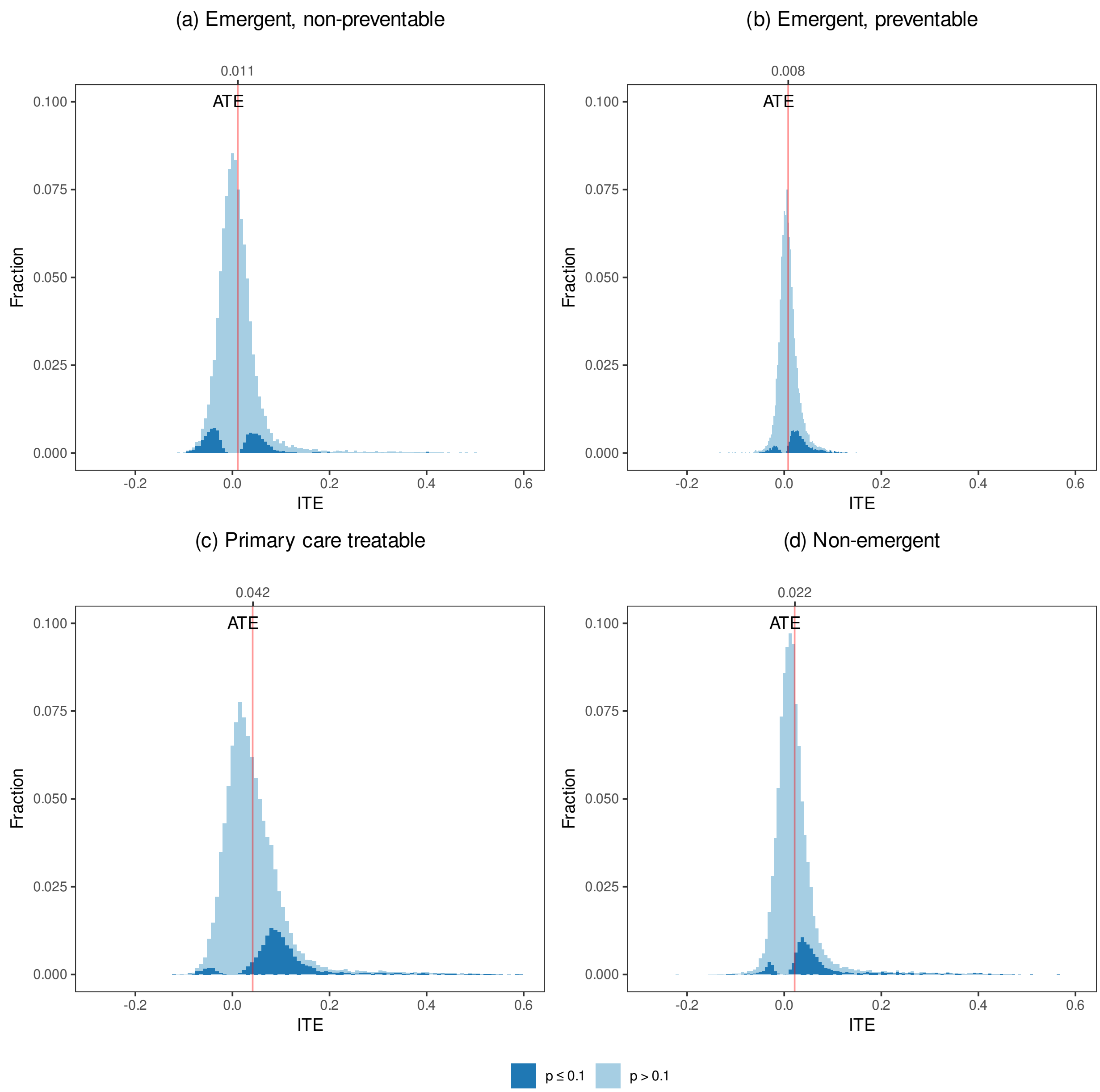
Notes This figure shows the estimates of variable importance for the top 20 characteristics used in growing the generalized random forests in estimating the CATE of winning the lottery (and being invited to apply for Medicaid) for any overall ED visit. The variable importance measure is a simple weighted sum of the proportion of times a variable is used in a splitting step at each depth in growing the causal forest, thus, capturing how important a variable is for driving treatment effect heterogeneity. The sample consists of 24,615 individuals in the Taubman et al. (2014) sample with non-missing information on pre-lottery emergency department utilization and SNAP/TANF receipt.

Figure B.4: Distribution of individualized treatment effects of winning the lottery on the number of ED visits



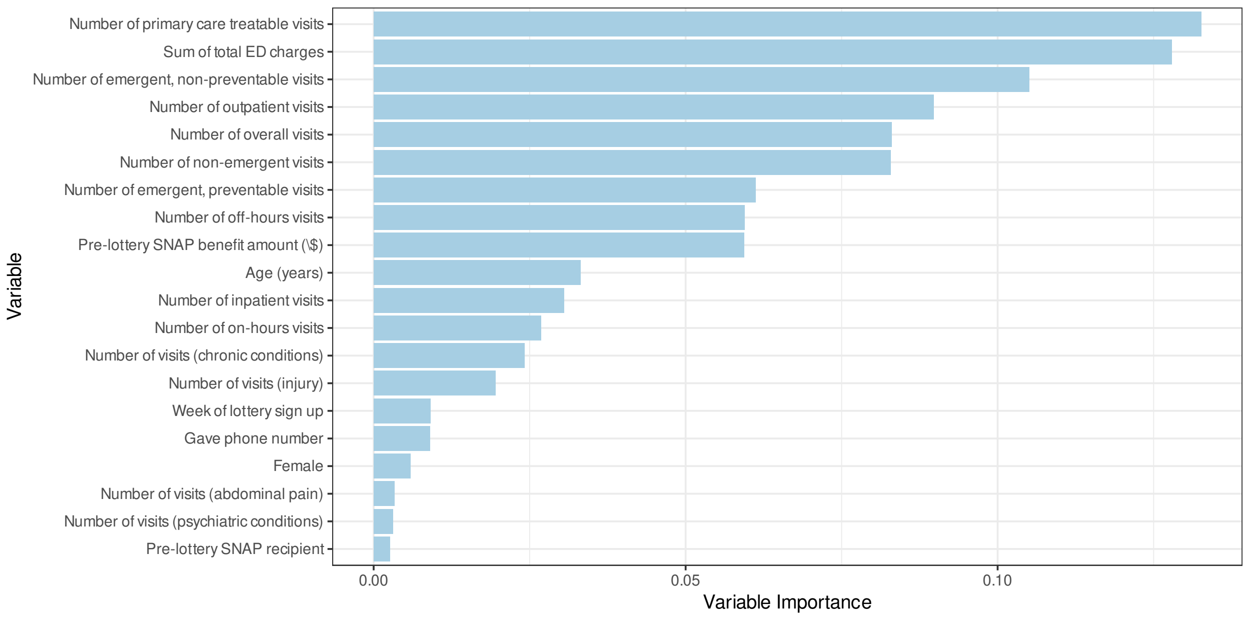
Notes: This figure plots the individualized treatment effects of winning the lottery (and being invited to apply for Medicaid) on the number of overall ED visit (panel a), the number of outpatient ED visits (panel b), and the number of inpatient ED use (panel c) based on generalized random forests. The darker shade denotes statistical significance at the 10% level. The red vertical line indicates the local average treatment effect. The baseline sample consists of 24,615 individuals in the Taubman et al. (2014) sample with non-missing information on pre-lottery emergency department utilization and SNAP/TANF receipt. The estimates displayed exclude less than half a percentile at the top and bottom of the distribution, resulting in the axes corresponding approximately to the percentile range [0.5%, 99.5%]. Bin size is chosen according to the Freedman-Diaconis rule.

Figure B.5: Distribution of individualized treatment effects of winning the lottery on the number of ED visits by type of condition



Notes: This figure plots the individualized treatment effects of winning the lottery (and being invited to apply for Medicaid) by type of ED visit based on generalized random forests for the number of emergent, non-preventable visit (panel a), the number of emergent, preventable visit (panel b), the number of primary care treatable visit (panel c), and the number of non-emergent visit (panel d). Measures of the type of ED visit are based on Billings et al.'s (2000) algorithm described in Taubman et al. (2014). The darker shade denotes statistical significance at the 10% level. The red vertical line indicates the local average treatment effect. The baseline sample consists of 24,615 individuals in the Taubman et al. (2014) sample with non-missing information on pre-lottery emergency department utilization and SNAP/TANF receipt. The estimates displayed exclude less than half a percentile at the top and bottom of the distribution, resulting in the axes corresponding approximately to the percentile range [0.5%, 99.5%]. Bin size is chosen according to the Freedman-Diaconis rule.

Figure B.6: Variable importance scores in growing causal forest (Number of visits)



Notes This figure shows the estimates of variable importance for the top 20 characteristics used in growing the generalized random forests in estimating the CATE of winning the lottery (and being invited to apply for Medicaid) for the number of overall ED visits. The variable importance measure is a simple weighted sum of the proportion of times a variable is used in a splitting step at each depth in growing the causal forest, thus, capturing how important a variable is for driving treatment effect heterogeneity. The sample consists of 24,615 individuals in the Taubman et al. (2014) sample with non-missing information on pre-lottery emergency department utilization and SNAP/TANF receipt.

Table B.1: Treatment effect estimates of winning the lottery on ED use

	GRF estimates			Linear estimates		
Variable	ATE	SE	*p*-value	ATE	SE	*p*-value
Extensive margin						
Any overall visit	0.015	0.006	0.008	0.017	0.006	0.004
Any inpatient visit	-0.003	0.003	0.430	-0.003	0.003	0.424
Any outpatient visit	0.019	0.006	0.001	0.020	0.006	0.001
Any emergent, non-preventable visit	0.000	0.005	0.940	0.002	0.005	0.731
Any emergent, preventable visit	0.008	0.004	0.026	0.010	0.004	0.010
Any primary care treatable visit	0.016	0.005	0.002	0.017	0.005	0.001
Any non-emergent visit	0.015	0.005	0.001	0.016	0.005	0.001
Intensive margin						
Number of overall visits	0.083	0.027	0.002	0.093	0.026	0.000
Number of inpatient visits	-0.004	0.006	0.519	-0.004	0.006	0.516
Number of outpatient visits	0.089	0.024	0.000	0.096	0.024	0.000
Number of emergent, non-preventable visits	0.011	0.008	0.162	0.010	0.008	0.224
Number of emergent, preventable visits	0.008	0.004	0.056	0.009	0.004	0.034
Number of primary care treatable visits	0.042	0.011	0.000	0.042	0.011	0.000
Number of non-emergent visits	0.022	0.008	0.007	0.026	0.008	0.002

Notes: This table reports the estimates of winning the lottery on ED use based on generalized random forests and a linear model. The sample consists of 24,615 individuals in the Taubman et al. (2014) sample with non-missing information on pre-lottery emergency department utilization and SP/TANF receipt.

Table B.2: Empirical quantiles of the distribution of individualized treatment effects of winning the lottery on ED use

Variable	ATE	Min	25%	50%	75%	Max
Extensive margin						
Any overall visit	0.015	-0.112	-0.007	0.017	0.039	0.159
Any inpatient visit	-0.003	-0.072	-0.012	-0.002	0.007	0.076
Any outpatient visit	0.019	-0.107	-0.003	0.018	0.041	0.159
Any emergent, non-preventable visit	0.000	-0.103	-0.016	0.001	0.018	0.096
Any emergent, preventable visit	0.008	-0.061	-0.003	0.007	0.019	0.095
Any primary care treatable visit	0.016	-0.089	-0.003	0.016	0.035	0.137
Any non-emergent visit	0.015	-0.077	-0.001	0.014	0.031	0.116
Intensive margin						
Number of overall visits	0.083	-0.427	0.000	0.061	0.145	1.205
Number of inpatient visits	-0.004	-0.148	-0.018	-0.003	0.011	0.248
Number of outpatient visits	0.089	-0.347	0.009	0.065	0.146	1.154
Number of emergent, non-preventable visits	0.011	-0.116	-0.015	0.004	0.026	0.576
Number of emergent, preventable visits	0.008	-0.270	-0.003	0.006	0.018	0.238
Number of primary care treatable visits	0.042	-0.118	0.003	0.031	0.067	0.686
Number of non-emergent visits	0.022	-0.218	-0.002	0.014	0.034	0.563

Notes: This table reports selected quantiles of the individualized treatment effects of winning the lottery (and being invited to apply for Medicaid) on ED use based on generalized random forests. The first column reports the average effect (intent-to-treat effect). The baseline sample consists of 24,615 individuals in the Taubman et al. (2014) sample with non-missing information on pre-lottery emergency department utilization and SNAP/TANF receipt.

Table B.3: GATE estimates of winning the lottery on the propensity of ED use

	Panel A: GRF estimates			Panel B: Linear estimates			
Group	GATE	SE	p-value	GATE	SE	p-value	% N
ATE	0.02	0.01	0.01	0.02	0.01	0.00	100.00
Female:	0.00	0.01	0.72	0.01	0.01	0.34	0.55
Male:	0.03	0.01	0.00	0.03	0.01	0.00	0.45
Gave phone number: No	0.01	0.02	0.54	0.01	0.02	0.66	0.13
Gave phone number: Yes	0.02	0.01	0.01	0.02	0.01	0.00	0.87
English as preferred language: No	0.01	0.01	0.59	0.00	0.01	0.73	0.14
English as preferred language: Yes	0.02	0.01	0.01	0.02	0.01	0.01	0.86
First week sign-up: No	0.01	0.01	0.06	0.02	0.01	0.01	0.62
First week sign-up: Yes	0.02	0.01	0.06	0.01	0.01	0.16	0.38
Pre-lottery SNAP recipient: No	0.00	0.01	0.94	0.00	0.01	0.62	0.46
Pre-lottery SNAP recipient: Yes	0.03	0.01	0.00	0.02	0.01	0.01	0.54
Pre-lottery TANF recipient: No	0.02	0.01	0.01	0.02	0.01	0.00	0.98
Pre-lottery TANF recipient: Yes	-0.01	0.04	0.78	0.01	0.04	0.82	0.02
Age $\geq$ 50: No	0.02	0.01	0.00	0.02	0.01	0.00	0.75
Age $\geq$ 50: Yes	-0.01	0.01	0.67	0.00	0.01	0.77	0.25
Two+ household members on lottery list: No	0.01	0.01	0.06	0.01	0.01	0.06	0.80
Two+ household members on lottery list: Yes	0.03	0.01	0.02	0.04	0.01	0.00	0.20
Any pre-lottery ED visit No	0.01	0.01	0.07	0.02	0.01	0.02	0.69
Any pre-lottery ED visit: Yes	0.02	0.01	0.05	0.02	0.01	0.09	0.31
Any pre-lottery on-hours ED visit: No	0.01	0.01	0.03	0.02	0.01	0.01	0.77
Any pre-lottery on-hours ED visit: Yes	0.02	0.01	0.11	0.02	0.01	0.13	0.23
Any pre-lottery off-hours ED visit: No	0.01	0.01	0.04	0.01	0.01	0.03	0.81
Any pre-lottery off-hours ED visit: Yes	0.02	0.01	0.08	0.02	0.01	0.16	0.19
Any pre-lottery emergent, non-preventable ED visit: No	0.01	0.01	0.07	0.01	0.01	0.03	0.79
Any pre-lottery emergent, non-preventable ED visit: Yes	0.03	0.01	0.02	0.03	0.01	0.07	0.21
Any pre-lottery emergent, preventable ED visit: No	0.01	0.01	0.02	0.02	0.01	0.00	0.90
Any pre-lottery emergent, preventable ED visit: Yes	0.03	0.02	0.15	0.03	0.02	0.13	0.10
Any pre-lottery primary care treatable ED visit: No	0.01	0.01	0.10	0.01	0.01	0.03	0.74
Any pre-lottery primary care treatable ED visit: Yes	0.03	0.01	0.02	0.03	0.01	0.05	0.26
Any pre-lottery non-emergent ED visit: No	0.01	0.01	0.04	0.02	0.01	0.01	0.86
Any pre-lottery non-emergent ED visit: Yes	0.03	0.02	0.07	0.02	0.02	0.18	0.14

Notes: This table reports the GATE estimates of winning the lottery based on generalized random forests in Panel A and the linear model GATEs in Panel B. The overall effect is reproduced in the first row. The sample consists of 24,615 individuals in the Taubman et al. (2014) sample with non-missing information on pre-lottery emergency department utilization and SNAP/TANF receipt.

Table B.4: GATE estimates of winning the lottery on the number of ED visits

	Panel A: GRF estimates			Panel B: Linear estimates			
Group	GATE	SE	*p*-value	GATE	SE	*p*-value	% N
ATE	0.08	0.03	0.00	0.08	0.03	0.01	100.00
Female:	0.05	0.03	0.19	0.05	0.04	0.21	0.55
Male:	0.13	0.04	0.00	0.12	0.05	0.01	0.45
Gave phone number: No	-0.04	0.08	0.62	-0.06	0.09	0.50	0.13
Gave phone number: Yes	0.10	0.03	0.00	0.10	0.03	0.00	0.87
English as preferred language: No	0.03	0.03	0.43	0.01	0.04	0.78	0.14
English as preferred language: Yes	0.09	0.03	0.00	0.09	0.04	0.01	0.86
First week sign-up: No	0.09	0.03	0.00	0.11	0.04	0.00	0.62
First week sign-up: Yes	0.06	0.05	0.17	0.03	0.06	0.54	0.38
Pre-lottery SNAP recipient: No	0.04	0.03	0.17	0.03	0.03	0.34	0.46
Pre-lottery SNAP recipient: Yes	0.12	0.04	0.01	0.11	0.05	0.04	0.54
Pre-lottery TANF recipient: No	0.08	0.03	0.00	0.08	0.03	0.01	0.98
Pre-lottery TANF recipient: Yes	0.07	0.27	0.78	0.17	0.29	0.55	0.02
Age $\geq$ 50: No	0.10	0.03	0.00	0.09	0.04	0.02	0.75
Age $\geq$ 50: Yes	0.03	0.05	0.50	0.07	0.06	0.23	0.25
Two+ household members on lottery list: No	0.07	0.03	0.02	0.07	0.04	0.07	0.80
Two+ household members on lottery list: Yes	0.12	0.03	0.00	0.15	0.05	0.00	0.20
Any pre-lottery ED visit No	0.05	0.02	0.01	0.06	0.02	0.00	0.69
Any pre-lottery ED visit: Yes	0.15	0.07	0.04	0.12	0.09	0.15	0.31
Any pre-lottery on-hours ED visit: No	0.05	0.02	0.01	0.06	0.02	0.00	0.77
Any pre-lottery on-hours ED visit: Yes	0.19	0.09	0.04	0.17	0.11	0.12	0.23
Any pre-lottery off-hours ED visit: No	0.05	0.02	0.02	0.05	0.02	0.03	0.81
Any pre-lottery off-hours ED visit: Yes	0.22	0.11	0.04	0.16	0.12	0.21	0.19
Any pre-lottery emergent, non-preventable ED visit: No	0.06	0.02	0.01	0.07	0.02	0.00	0.87
Any pre-lottery emergent, non-preventable ED visit: Yes	0.23	0.14	0.11	0.18	0.17	0.28	0.13
Any pre-lottery emergent, preventable ED visit: No	0.06	0.02	0.00	0.08	0.02	0.00	0.92
Any pre-lottery emergent, preventable ED visit: Yes	0.30	0.22	0.17	0.36	0.25	0.15	0.08
Any pre-lottery primary care treatable ED visit: No	0.05	0.02	0.03	0.06	0.02	0.00	0.81
Any pre-lottery primary care treatable ED visit: Yes	0.24	0.11	0.03	0.19	0.13	0.13	0.19
Any pre-lottery non-emergent ED visit: No	0.07	0.02	0.00	0.08	0.02	0.00	0.86
Any pre-lottery non-emergent ED visit: Yes	0.14	0.14	0.30	0.07	0.16	0.68	0.14

Notes: This table reports the GATE estimates of winning the lottery based on generalized random forests in Panel A and the GATE based on a linear model in Panel B. The overall effect is reproduced in the first row. The sample consists of 24,615 individuals in the Taubman et al. (2014) sample with non-missing information on pre-lottery emergency department utilization and SNAP/TANF receipt.

Table B.5: Characteristics of individuals who increased and decreased ED use upon winning the lottery

Variable	Increased ED use	Decreased ED use	Difference
Lottery list characteristics			
Age (years)	38.81	41.32	-2.51***
Gave phone number	0.87	0.87	0.00
English as preferred language	0.87	0.84	0.03***
Female	0.51	0.63	-0.12***
Week of lottery sign up	1.57	1.61	-0.04**
Provided P.O. box address	0.02	0.03	-0.01***
Signed up self for lottery	0.89	0.92	-0.03***
Pre-lottery SNAP recipient	0.62	0.35	0.27***
Pre-lottery SNAP benefit amount ($)	1613.82	715.39	898.43***
Pre-lottery TANF recipient	0.02	0.01	0.01***
Pre-lottery TANF benefit amount ($)	110.21	65.85	44.36***
Pre-lottery ED usage			
Number of overall visits	0.87	0.52	0.35***
Number of inpatient visits	0.09	0.07	0.02***
Number of outpatient visits	0.78	0.45	0.33***
Number of on-hours visits	0.50	0.32	0.18***
Number of off-hours visits	0.37	0.21	0.16***
Number of emergent, non-preventable visits	0.18	0.10	0.08***
Number of emergent, preventable visits	0.07	0.05	0.02***
Number of primary care treatable visits	0.30	0.18	0.12***
Number of non-emergent visits	0.18	0.11	0.07***
Number ambulatory-care-sensitive visits	0.05	0.04	0.01***
Number of visits (chronic conditions)	0.14	0.12	0.02***
Number of visits (injury)	0.20	0.10	0.10***
Number of visits (skin conditions)	0.05	0.03	0.02***
Number of visits (abdominal pain)	0.04	0.02	0.02***
Number of visits (back pain)	0.04	0.02	0.02***
Number of visits (chest pain)	0.02	0.02	0.00***
Number of visits (headache)	0.03	0.02	0.01***
Number of visits (mood disorders)	0.02	0.03	-0.01
Number of visits (psychiatric conditions)	0.06	0.05	0.01
Sum of total ED charges	1002.57	615.50	387.07***
N	16871	7742	24613

Notes: This table reports the means of individual characteristics and pre-randomization ED use for those estimated to increase and decrease ED use upon winning the lottery based on the causal forest CATE estimates. ED use is measured as the propensity to use the ED. Panel A reports the means for the full sample while Panel B is limited to individuals with effects significant at the 10% level. The sample consists of 24,615 individuals in the Taubman et al. (2014) sample with non-missing information on pre-lottery emergency department utilization and SNAP/TANF receipt.

Table B.6: Characteristics of individuals who increased and decreased ED use upon winning the lottery (Number of visits)

Variable	Increased ED use	Decreased ED use	Difference
Lottery list characteristics			
Age (years)	38.52	42.89	-4.37***
Gave phone number	0.87	0.88	-0.01***
English as preferred language	0.87	0.84	0.03***
Female	0.53	0.59	-0.06***
Week of lottery sign up	1.59	1.55	0.04
Provided P.O. box address	0.02	0.03	-0.01***
Signed up self for lottery	0.89	0.92	-0.03***
Pre-lottery SNAP recipient	0.59	0.36	0.23***
Pre-lottery SNAP benefit amount ($)	1502.03	813.33	688.70***
Pre-lottery TANF recipient	0.02	0.02	0.00**
Pre-lottery TANF benefit amount ($)	95.68	98.23	-2.55
Pre-lottery ED usage			
Number of overall visits	0.90	0.32	0.58***
Number of inpatient visits	0.10	0.06	0.04***
Number of outpatient visits	0.80	0.26	0.54***
Number of on-hours visits	0.52	0.21	0.31***
Number of off-hours visits	0.38	0.12	0.26***
Number of emergent, non-preventable visits	0.18	0.07	0.11***
Number of emergent, preventable visits	0.07	0.03	0.04***
Number of primary care treatable visits	0.32	0.08	0.24***
Number of non-emergent visits	0.18	0.08	0.10***
Number ambulatory-care-sensitive visits	0.05	0.03	0.02***
Number of visits (chronic conditions)	0.15	0.08	0.07***
Number of visits (injury)	0.20	0.06	0.14***
Number of visits (skin conditions)	0.06	0.01	0.05***
Number of visits (abdominal pain)	0.04	0.01	0.03***
Number of visits (back pain)	0.04	0.02	0.02***
Number of visits (chest pain)	0.02	0.01	0.01***
Number of visits (headache)	0.03	0.01	0.02***
Number of visits (mood disorders)	0.03	0.02	0.01***
Number of visits (psychiatric conditions)	0.07	0.03	0.04***
Sum of total ED charges	1022.87	411.30	611.57***
N	18494.00	6105.00	24599

Notes: This table reports the means of individual characteristics and pre-randomization ED use for those estimated to increase and decrease ED use upon winning the lottery based on the causal forest CATE estimates. ED use is measured as the number of total visits. Panel A reports the means for the full sample while Panel B is limited to individuals with effects significant at the 10% level. The sample consists of 24,615 individuals in the Taubman et al. (2014) sample with non-missing information on pre-lottery emergency department utilization and SNAP/TANF receipt.

Table B.7: Variable importance for all variables in growing causal forest (overall ED use)

Any visit		Number of visits	
Variable	Importance	Variable	Importance
Pre-lottery SNAP benefit amount ($)	0.23	Number of primary care treatable visits	0.13
Age (years)	0.16	Sum of total ED charges	0.13
Sum of total ED charges	0.09	Number of emergent, non-preventable visits	0.11
Number of emergent, non-preventable visits	0.08	Number of outpatient visits	0.09
Number of primary care treatable visits	0.08	Number of overall visits	0.08
Female	0.04	Number of non-emergent visits	0.08
Number of non-emergent visits	0.04	Number of emergent, preventable visits	0.06
Week of lottery sign up	0.04	Number of off-hours visits	0.06
Pre-lottery SNAP recipient	0.02	Pre-lottery SNAP benefit amount ($)	0.06
Number of emergent, preventable visits	0.02	Age (years)	0.03
Number of visits (chronic conditions)	0.01	Number of inpatient visits	0.03
Number of inpatient visits	0.01	Number of on-hours visits	0.03
Number of outpatient visits	0.01	Number of visits (chronic conditions)	0.02
Number of off-hours visits	0.01	Number of visits (injury)	0.02
Number of overall visits	0.01	Week of lottery sign up	0.01
Number of on-hours visits	0.01	Gave phone number	0.01
Gave phone number	0.01	Female	0.01
Signed up self for lottery	0.01	Number of visits (abdominal pain)	0.00
Number of visits (injury)	0.01	Number of visits (psychiatric conditions)	0.00
Pre-lottery TANF benefit amount ($)	0.00	Pre-lottery SNAP recipient	0.00
English as preferred language	0.00	Number of visits (chest pain)	0.00
Number of visits (mood disorders)	0.00	Pre-lottery TANF benefit amount ($)	0.00
Number of visits (psychiatric conditions)	0.00	Number of visits (skin conditions)	0.00
Number ambulatory-care-sensitive visits	0.00	Number of visits (back pain)	0.00
Number of visits (abdominal pain)	0.00	Number of visits (headache)	0.00
Number of visits (skin conditions)	0.00	Number of visits (mood disorders)	0.00
Pre-lottery TANF recipient	0.00	Number ambulatory-care-sensitive visits	0.00
Number of visits (chest pain)	0.00	Signed up self for lottery	0.00
Number of visits (back pain)	0.00	Pre-lottery TANF recipient	0.00
Number of visits (headache)	0.00	English as preferred language	0.00
Provided P.O. box address	0.00	Provided P.O. box address	0.00

Notes: This table shows the top variable importance scores of all characteristics for growing the generalized random forests used to estimate the ITE of winning the lottery for overall ED visits. The variable importance measure is a simple weighted sum of the proportion of times a variable is used in a splitting step at each depth in growing the causal forest, thus, capturing how important a variable is for driving treatment effect heterogeneity. The sample consists of 24,615 individuals in the Taubman et al. (2014) sample with non-missing information on pre-lottery emergency department utilization and SNAP/TANF receipt.