Evaluating the Heterogeneous Impacts of Indonesia's National Health Insurance Scheme using Causal Machine Learning

Vishalie Shah

PhD
University of York
Economics
March 2023

Abstract

Targeting policies efficiently requires an understanding of who is expected to benefit the most and the least from each of these policies. It is well understood that policy impacts are not uniform across a given population, but heterogeneous. The increasing availability of rich, granular data has enabled researchers to estimate heterogeneous treatment effects, which capture the extent to which differences in observable characteristics can modify policy (or treatment) effects. This has initiated a shift away from static policymaking towards a personalised, data-driven approach. Recent developments for estimating heterogeneous treatment effects have resulted in approaches that use causal machine learning to search for effect modification in a flexible, structured way. This thesis explores some of these approaches within the context of an impact evaluation of Indonesia's national health insurance policy, the Jaminan Kesehatan Nasional (JKN), using non-randomised data. Throughout, the focus is on the subsidised component of JKN that targets low-income individuals. The first chapter reviews the current state of methods for estimating treatment effect heterogeneity, focusing on three causal machine learning approaches, and demonstrates their application to a case study evaluation of JKN on inpatient health care utilisation. The second chapter expands on this work with a more detailed evaluation on both inpatient and outpatient health care utilisation, using a combination of predictive and causal machine learning methods. The third chapter uses estimates of treatment effect heterogeneity to learn optimal policy rules that efficiently assign different modalities of JKN to the population, in a way that reduces the risk of catastrophic health expenditures; another target of JKN. The results provide evidence of varying policy impacts for both outcome measures. In particular, rural households benefit relatively less from the policy compared to urban households, which is consistent with previous evidence about geographical variations in health care accessibility. This work suggests that more careful considerations of policy impacts beyond the average could improve policymaking.

Contents

List of Tables

List of Figures

Acknowledgements

First of all, I would like to express my deepest gratitude to my supervisors, Noemi Kreif and Andrew Jones. Their countless suggestions and contributions have invaluably benefitted this work, and perhaps more significantly, their unwavering dedication to this project has been influential in getting to this point. They have been a limitless source of inspiration and have no doubt shaped my development as a researcher.

I would also like to thank my thesis advisor, Marc Suhrcke, for his valuable comments, and the Department of Economics and Related Studies for funding this research.

I am grateful to the health economics community at York. In particular, I would like to thank members of the HEDG group for all of the helpful insights and discussions. The weekly meetings have provided a constant form of reassurance, especially during the pandemic when the PhD journey felt lonelier than usual. I would also like to thank Julia Hatamyar for sharing many useful tips along the way, and helping me to navigate this fast-evolving field.

My thanks also extend to Mark van der Laan and Ivana Malenica for very kindly hosting me during my research visit to U.C. Berkeley.

Lastly, a special thanks go to my family for their continued encouragement over the years. This thesis is dedicated to my husband, Viraj, without whose endless support reaching this milestone would not have been possible.

Declaration

I declare that this thesis is a presentation of original work and that I am the primary author. This work has not previously been presented for an award at this, or any other, university. All sources are acknowledged in the Bibliography.

I am the principal author for all chapters, having formulated the research questions, assembled the datasets, constructed the variables, defined the empirical strategy, undertaken the statistical analyses, written the scientific code, interpreted and discussed the results, and written up the chapters. I state the contributions of my co-authors to each chapter below.

Chapter 1 is co-authored with Noemi Kreif and Andrew M. Jones. A version of this chapter is published in the Handbook of Research Methods and Applications in Empirical Microeconomics, 2021, Volume 30. Issue 9. pp. 438-489.

For Chapter 1, Noemi Kreif and Andrew M. Jones provided guidance on the identification of the important research questions for the narrative review, advised on the empirical strategy for the case study, and provided comments on the written drafts.

Chapter 2 is co-authored with Noemi Kreif. Previous versions of this chapter were presented and discussed at the 14th International Health Economics Association (iHEA) Annual Conference (Virtual, 2021), the 7th Indonesian Health Economics Association (InaHEA) Biennial Scientific Meeting (Bandung/Virtual, 2021) and the 8th European Health Economics Association (EuHEA) Conference (Rotterdam/Virtual, 2021).

For Chapter 2, Noemi Kreif advised on the refinement of the research question, the choice of methodology and the empirical strategy, and provided comments on the written drafts.

Chapter 3 is co-authored with Taufik Hidayat, Andrew M. Jones and Noemi Kreif.

For Chapter 3, Noemi Kreif and Andrew Jones supported the refinement of the research question, advised on the choice of methodology and the empirical strategy,

and provided comments on the written drafts. Taufik Hidayat provided guidance on the selection of the targeting variables to include in the empirical model.

The data that support the findings of this thesis are the 2017 National Socioeconomic Survey (SUSENAS) and the 2018 Village Potential Statistics (PODES), which were purchased from Badan Pusat Statistik (Statistics Indonesia). The purchase was funded by the Medical Research Council grant (MR/T04487X/1, PI: Noemi Kreif). Taufik Hidayat advised on how to assemble the data into a usable format.

Introduction

Learning causal relationships from data is one of the fundamental objectives of applied econometric research. Understanding the likely impact of an action is critical for strategic and informed decision making. This is particularly important within the context of health and social policy, where decision makers face the challenge of improving population outcomes, typically under binding budget constraints. To maximise the intended impacts of a new policy, or to enhance an existing one, they are confronted with a myriad of questions: Is the policy effective overall? Who benefits the most and the least? Does the policy improve existing inequalities in outcomes? Is there a better allocation strategy than the current assignment that increases net benefits?

Although subtle, these questions are inherently causal in nature and can be answered through impact evaluation, which is founded on well-established causal inference methodology, such as the potential outcomes framework (Holland, 1986; Imbens and Rubin, 2015; Rubin, 1974). The goal is to conduct inference on causal effects by estimating the impact of counterfactual policies (that is, estimating the impact of a policy change compared to an alternative, or "baseline"). To derive the causal effect requires defining a target estimand (that is, the causal parameter of interest, for example, the average treatment effect), and making the necessary assumptions to be able to identify the estimand using observed data. The most effective way to identify causal parameters is through randomising the policy assignment, often referred to as the "gold standard" of study designs. Randomisation reduces or eliminates the preferential assignment of policy to observations, also known as selection (or confounding) that can create bias in treatment effect estimates. In reality, most impact evaluations of health and social policy rely on non-randomised, observational data, which is less costly and more readily available (Athey and Imbens, 2017). There are a set of assumptions that must be satisfied in order to draw the same causal conclusions from observational data that would be possible in a randomised study (Athey, 2015). Once the causal parameters have been adequately identified, the

focus turns towards the estimation task.

Classical methods for parameter estimation are rooted in economic theory, and usually involve selecting a statistical model (usually parametric) that best describes the functional form of the relationship between the outcome y and a set of predictors x. For estimating average treatment effects, these tend to be regression and matching methods, where the modelling of y as a function of x is necessary to adjust for confounding (Imbens and Wooldridge, 2009). The target causal parameter is estimated by finding the values that best fit the data in terms of the objective function (for example, the sum of squared errors or the likelihood function). Ultimately, the emphasis is on producing good estimators of the causal parameter, rather than the best model. Statistical theory is used to report standard errors and confidence intervals for the estimated parameters. Behind the scenes, many different model specifications may be tested, but these are rarely reported in practice due to concerns about multiple hypothesis testing (Athey, 2019). This practice of testing different models, however, may actually lead to ad-hoc model selection and biased estimates of the target parameter.

In recent years, there have been significant advances in the scope of methods available for causal effect estimation. One of the key drivers behind this development is the so-called "big-data" revolution (Einav and Levin, 2014). As datasets continue to grow larger and become more high-dimensional (that is, more variables than observations), more powerful tools are required to conduct meaningful analyses (Varian, 2014). In particular, the increasingly important role of machine learning in economics is becoming more apparent (Athey, 2019; Athey and Imbens, 2019; Taylor et al., 2014). Machine learning (or "supervised" machine learning, which is most relevant here) is a set of algorithms that are designed for prediction tasks (that is, generating predictions of y given x). This differs from classical econometric methods which instead focus on parameter estimation (that is, generating estimates of parameters that explain the relationship between x and y) (Mullainathan and Spiess, 2017). Machine learning algorithms are not overly concerned about the underlying model that generates the data, and rely solely on the input data for model selection. Rather than specifying and estimating a single model, the algorithm usually finds

the best model, among different specifications of the same model and alternative models that maximises a well-defined loss function (Athey, 2019).

There is a well-established literature describing the benefits of machine learning and its potential uses in empirical economics – see Mullainathan and Spiess (2017) and Athey and Imbens (2019) for a review. A few relevant examples are highlighted here. First, the focus of machine learning is on fitting flexible models that achieve good out-of-sample predictions, rather than in-sample goodness-of-fit. It does this using model validation techniques on unused data, which enables unbiased model comparisons. Second, machine learning enables a data-driven approach to variable selection in sparse models (that is, models where only a small number of parameters are non-zero), which in most cases is superior to a manual, theoretically-motivated selection process, particularly in high-dimensional settings. Third, machine learning algorithms are generally designed to be computationally efficient and are able to scale well to large datasets. In summary, despite the differences in objectives between machine learning and econometric methods, there are scenarios in which machine learning can replace classical approaches altogether (for example, prediction problems), and others in which machine learning can provide additional support (for example, causal parameter estimation) (Athey, 2017; Kleinberg et al., 2015; Kreif and DiazOrdaz, 2019). The latter is particularly true for datasets with a large number of variables because for any estimation strategy, it is important to be able to fit a flexible model that accounts for this feature of the data.

The focus of this thesis is on one area within the causal inference literature in which machine learning has proven to be particularly useful; the estimation of heterogeneous treatment effects, which captures the extent to which differences in observed characteristics can modify treatment effects within a given population. In observational studies, heterogeneous treatment effects can be estimated via subgroup analyses, which compare policy impacts across different subpopulations. For example, health policies often aim to target disadvantaged populations, so an assessment of heterogeneity that isolates treatment effects for subgroups of varying socioeconomic status could generate valuable information for policymaking (Mackenbach, 2003). The problem is that, until recently, subgroup analyses have typically relied

on ad-hoc interactions between the treatment and pre-specified effect modifiers in the statistical model, which in the advent of high-dimensional data is impractical and, more importantly, prone to statistical bias (Kreif et al., 2022; Petticrew et al., 2012). The recent adoption of machine learning in the econometrics toolbox has presented more principled approaches to subgroup discovery that capture heterogeneity in treatment effects through flexible, data-adaptive models (Athey et al., 2019b; Künzel et al., 2019a). The argument is that machine learning methods can systematically search over the entire covariate space, thus reducing arbitrary subgroup analyses (VanderWeele et al., 2019).

The overarching aim of this thesis is to explore how recent methodological developments in heterogeneous treatment effect estimation can be applied to evaluations of health policies using observational data. In recent years, a growing number of applied econometric studies have been adopting these methods for policy evaluation, particularly in the field of labour economics (Brand et al., 2021; Davis and Heller, 2017; Knaus et al., 2022). In the health setting, the existing evidence base focuses on exploring heterogeneity for personalised medicine, which tends to rely on randomised evaluations of clinical interventions (Dahabreh et al., 2016; Kent et al., 2018; Kosorok and Laber, 2019). There are limited studies exploring the heterogeneous impacts of system-level health policies, where in most cases, there is no randomisation and so additional statistical measures are necessary to control for confounding bias (Kreif et al., 2022). This work hopes to fill the research gap by firstly demonstrating to health researchers the potential advantages of incorporating causal machine learning into subgroup analyses, and secondly contributing to the empirical evidence base through rigorous evaluations of Indonesia's national health insurance programme, the Jaminan Kesehatan Nasional (JKN), for which heterogeneous impacts are yet to be fully explored using data-driven approaches. Throughout this thesis, common challenges faced by health researchers are highlighted and addressed, including settings with rare outcomes (for example, health care utilisation outcomes) and the potential concerns that arise when the selection of study participants is not random.

Chapter 1 provides a comprehensive review into the current state of methods for

estimating treatment effect heterogeneity using machine learning. The primary objective is to introduce the reader to the recently established and rapidly evolving field of causal machine learning, and to highlight some of the key methodological developments that are gaining popularity among applied researchers. A particular emphasis is given to three proposed algorithms that have been specifically designed to estimate heterogeneous treatment effects: the "X-learner" (Künzel et al., 2019a), the "R-learner" (Nie and Wager, 2021) and "causal forests" (Athey et al., 2019b). The chapter concludes by introducing the JKN health insurance policy for a demonstrative application of the three methods to a real-world policy evaluation problem. A step-by-step instruction guides the reader on how to implement these methods in `R`. Since the case study example is mainly illustrative, specific details of the policy setting are deferred to later chapters. The key takeaways from this chapter are the relative strengths and limitations of alternative methods for estimating treatment effect heterogeneity, depending on the study design. For example, in non-randomised settings with a large number of potential effect modifiers, an implementation of the R-learner via the causal forests algorithm could be a good option.

Over the past decade, health care has become a policy priority in many developing countries. The inclusion of universal health coverage (UHC) in the UN's health-related Sustainable Development Goals (SDG target 3.8) has initiated a collective push towards policy reforms that improve access to effective, affordable and equitable health care (Filho et al., 2020; Ghebreyesus, 2017). Put simply, UHC ensures that everyone has access to high-quality health care, as and when required, without enduring financial hardship (Garrett et al., 2009; Moreno-Serra and Smith, 2012; World Health Organization, 2010). Cost barriers are a major deterrent to accessing health services, and globally, 100 million people fall into poverty each year as a result of health-related expenditure (Kim et al., 2017; Wiseman et al., 2018). Accordingly, many governments have implemented national health insurance policies to reduce the financial burden of health care, and to accelerate progress towards UHC (Lagomarsino et al., 2012).

Indonesia has successfully moved in this direction with the implementation of the JKN in 2014, achieving 83% coverage by mid-2019 (Prabhakaran et al., 2019). The

policy mandates all individuals to be enrolled into JKN, with those in formal employment enrolled into the contributory scheme via their employers, and those classified as poor and near-poor enrolled into the fully-subsidised scheme (PBI) (Maulana et al., 2022). The subsidised version is further divided into two schemes, depending on whether the subsidy is funded by the subnational (PBI-APBD) or national (PBI-APBN) government. Given that there are notable geographic and socioeconomic disparities in disease burden and access to health care across Indonesia, the impacts of JKN are likely heterogeneous (Agustina et al., 2019). Previous studies have found this to be the case, with policy impacts on health care utilisation and out-of-pocket health spending varying across subgroups stratified by location, wealth and health care accessibility (Anindya et al., 2020; Erlangga et al., 2019a; Kreif et al., 2022; Maulana et al., 2022; Nugraheni et al., 2020). However, only one of these studies, which explores the impacts of JKN on maternal health utilisation and infant mortality, supplements theoretical hypotheses about potential effect modifiers with a data-driven approach to subgroup discovery (Kreif et al., 2022). The current evidence base exploring varying policy impacts on other important objectives of JKN – utilisation of other health care services and financial protection from out-of-pocket health spending – using similar data-driven approaches is nonexistent. Chapters 2 and 3 aim to fill this gap.

Evaluating the impact of health insurance policies is complicated given well-founded concerns that some individuals may self-select into insurance. In the usual setting, the insured and the uninsured populations may differ in terms of characteristics that are unobservable to the researcher (Wagstaff, 2010). Both chapters take certain steps to minimise bias stemming from this issue. First, enrollees of the contributory scheme are removed from the analysis since they are fundamentally different from the rest of the population; eligibility is determined by employment status. The remaining population comprises the uninsured and those insured by the subsidised scheme, who display similar characteristics. The eligibility criteria for subsidised JKN is based on proxy means testing models that aim to identify the poorest households using a vector of demographic and socioeconomic covariates. A large proportion of the uninsured population comprises low- and middle-income individuals who are eligible for subsidised insurance but have not yet enrolled (Agustina et al., 2019;

Dartanto et al., 2020b). Restricting the sample to these two populations confronts some of the selection concerns. Second, to further reduce bias due to unobservable factors that may still influence selection into subsidised insurance, a rich and diverse set of individual-, household- and district-level characteristics are included as controls in the causal model, to capture, as much as possible, the selection process into insurance. All of the studies rely on cross-sectional data from the 2017 version of the National Socioeconomic Survey (SUSENAS).

Chapter 2 uses machine learning to evaluate the average and heterogeneous impacts of being enrolled into the subsidised component of JKN, compared to being uninsured, on health care access and utilisation, measured by the demand for outpatient and inpatient care. A predictive algorithm, the super learner, is used for prediction tasks, and a causal algorithm, causal forests, is used for estimating treatment effect heterogeneity (Athey et al., 2019b; van der Laan et al., 2007). Compared to other evaluations of JKN that explore utilisation outcomes, this study accounts for the mass points at zero by fitting two-part models that decompose the outcome distribution into zero and non-zero counts. Intuitively, this decomposition offers additional insights into the separate processes driving the decision to seek health care and the quantity of care consumed (Pohlmeier and Ulrich, 1995). A combination of theory-based and data-driven approaches are used to identify effect modifiers, and treatment effect heterogeneity is summarised using innovative methods from the recent causal machine learning literature (Athey and Wager, 2019; Chernozhukov et al., 2018b; Kennedy, 2020; Knaus et al., 2021; Semenova and Chernozhukov, 2021). The results find beneficial average impacts of subsidised health insurance on health care utilisation, and evidence of heterogeneity in impacts, meaning that some individuals increase their consumption more than others as a response to health insurance. Effect modifiers that drive this variation in effects consist of variables selected from theory (for example, age, wealth and health care accessibility) and data using machine learning methods (for example, education level, marital status and technology usage). Using the data to identify population subgroups with heterogeneous treatment effects is a novel contribution to the evidence base on the impacts of JKN and UHC schemes more generally. In addition, this study confronts some of the challenges commonly faced by health policy researchers: an abundance of zeros in the

outcome and the need to control for confounding under a selection on observables framework.

Another active area of research within the causal inference literature is personalised policymaking (Athey and Wager, 2021; Bertsimas and Kallus, 2020; van der Laan and Luedtke, 2015; Zhou et al., 2022). There is increasing interest in using data to learn "rules" that efficiently allocate policy based on observed data. The problem of learning policy rules is intrinsically linked to the estimation of heterogeneous treatment effects (Bembom and van der Laan, 2007). A decision making process that leverages upon these estimates through the evaluation of past and current programme performance can inform better future policies in the sense of maximising expected outcomes (Kitagawa and Tetenov, 2018; Manski, 2004). Learning policy rules is an optimisation task that can be solved using causal machine learning, and certain predictive algorithms, including the super learner, have been adapted for this purpose (Luedtke and van der Laan, 2016b; Montoya et al., 2022).

Chapter 3 uses machine learning to learn an optimal policy rule that allocates households to PBI-APBD or PBI-APBN among the eligible subsidised population under both unconstrained and resource-constrained settings. Despite a high rate of insurance coverage, there remains a large uninsured population, of whom most are eligible for the subsidised scheme. A targeted enrolment strategy has been proposed, which would involve assigning the uninsured to APBD or APBN (Dutta et al., 2020; World Bank, 2020). A policy rule that optimises the assignment strategy could support this process. In this study, the optimisation uses an objective function where net benefits correspond to a reduction in households' risk of incurring catastrophic health expenditure; a measure of the financial protectiveness of insurance. The adapted version of the super learner is used to find the best rule, among different candidate specifications of the estimated rule that minimises the risk of catastrophic health expenditure. The optimal policy rules are evaluated using robust estimators of the expected outcomes under the rule, and households that are counterfactually assigned to APBD and APBN are characterised using descriptive analyses. The results suggest that potential gains from policy learning are possible since there is evidence of heterogeneity in policy impacts. The optimal policy rules

achieve better expected outcomes (that is, a reduction in the risk of incurring catastrophic expenditures) compared to the actual assignment and the static rule that assigns APBN to everyone. Geography, particularly the urban-rural distinction, is one of the main differentiating factors between the assignment under the optimal policy rules and the actual assignment, which is known to be associated with the availability of health services and a determinant of health spending. This study is one of the first to learn optimal policy rules from evaluating a system-level health policy, and demonstrates the potential for realising welfare gains, if causal impacts (specifically, heterogeneous impacts) can be estimated reasonably well. It is also one of the few evaluations of social programmes that incorporates resource constraints into the policy learning problem, which is an often overlooked but important consideration (Bhattacharya and Dupas, 2012). Lastly, the study offers some insights into the performance of different machine learning algorithms, which may be useful for those interested in the computational aspects of optimisation problems.

The thesis concludes with a discussion of the three chapters, including a summary of the findings and their implications for policy and research, the main limitations, and potential directions for future work.

Chapter 1

Machine Learning for Causal Inference: Estimating Heterogeneous Treatment Effects

This chapter presents some of the recent developments in the literature for estimating heterogeneous treatment effects that can be attributed to differences in observable population characteristics. We compare traditional approaches for exploring treatment effect heterogeneity with newer approaches that incorporate machine learning methods. We introduce some of the key concepts used in machine learning and explain how these methods are being applied to estimate causal parameters using the potential outcomes framework of causal inference. These include the average treatment effect (ATE) and the conditional average treatment effect (CATE), that captures heterogeneity in treatment effects. We review three algorithms in particular for estimating the CATE that are popular among applied researchers – the X-learner, the R-learner and causal forests – and demonstrate their application to a case study evaluation of Indonesia's national health insurance programme, the JKN, using non-randomised data. We evaluate the average and heterogeneous impacts of the subsidised component of JKN on the utilisation of inpatient health care. Throughout, we provide guidance on how to implement these methods in `R`.

1.1 Introduction

Until recently, most of the causal inference literature has focused on evaluating the average impact of a change in policy, or *treatment*, on a population of interest. The corresponding causal estimand of interest is the average treatment effect (ATE) (Imbens and Rubin, 2015; Imbens and Wooldridge, 2009). However, underlying this average effect is substantial variation in how individuals respond to treatment, suggesting that treatment effects are not in fact uniform, but heterogeneous (Athey and Imbens, 2019). Variation in treatment response can be explained by differences in the background characteristics of respondents, as well as differences in the features of treatment. Some members of the population may respond positively to treatment, while others may require a higher *dose* to achieve the same effect. It could also be the case that although the policy impact on the population as a whole may be limited, certain subgroups of the population may still reap some benefits. For example, in the context of health policy, the beneficial impact of a national health insurance programme may be greater for vulnerable populations than those who can afford health care. Identifying these subgroups can be extremely informative for understanding causal mechanisms, personalising decision making, and guiding policy. It can provide decision makers with important insights into the distributional impacts of policy, including which subpopulations display above- or below-average treatment effects. The estimates of heterogeneous treatment effects, that is, the conditional (on covariates) average treatment effects (CATEs), can be used as key inputs into the design of optimal policy rules that give policymakers the tools required to determine which individuals or subpopulations would benefit the most from treatment, and at what dose (Athey, 2019; Imai and Ratkovic, 2013; Imai and Strauss, 2011).

Variability across units of analysis plays an important role across all empirical social research. In causal inference, heterogeneity can arise in two distinct forms. First, the existence of variation in the response to treatment, either for individual observations or for different strata of the population, based on differences in their observed characteristics. Second, the existence of unobserved or unmeasured differences between units of observation that are also correlated with the observed and included covari-

ates (also known as *unobserved heterogeneity*). Omitted variables are commonly a feature of causal inference studies, where selection into treatment cannot always be observed. If assignment to treatment is correlated with an omitted variable, it will lead to biased estimates of the treatment effect, as a result of endogeneity. There is a large volume of literature that explores unobserved heterogeneity and selection bias, in particular the role of microeconometric tools, such as instrumental variables analysis and panel data regression methods, that are used to consistently estimate the treatment effect (Heckman, 2001). The focus of this chapter is on the former definition of heterogeneity. That is, we are interested in estimating the variation in treatment effects for population subgroups, based on their observed characteristics.

Methods for estimating heterogeneity in treatment effects have gained popularity across a variety of research disciplines, most notably in clinical research (Foster et al., 2011). Early work on effect heterogeneity focused on detecting quantitative interactions between treatment effects and patient subgroups, often defined by baseline characteristics (Bonetti and Gelber, 2004; Gail and Simon, 1985). This type of subgroup analysis is a popular tool for estimating treatment effects for subpopulations that share similar characteristics. For example, researchers are often interested in how gender plays a role in treatment response. In this scenario, subgroup analysis is fairly straightforward and simply involves separately estimating treatment effects for men and women. Although this type of analysis allows for simple interactions between the treatment and covariates, there are several limitations in using this method for identifying heterogeneity. First, it can lead to *cherry-picking*, whereby researchers pre-specify covariates of interest that are favourable to existing results (Assmann et al., 2000; Cook et al., 2004). Second, it can result in reduced sample sizes, invalid p-values (due to multiple hypothesis testing) and *false discovery*, especially if statistical inference is not correctly adjusted for (Assmann et al., 2000; Brookes et al., 2004, 2001; List et al., 2019; Wang et al., 2007). Third, it can be impractical to use with high-dimensional data, where more complex covariate-treatment interactions are possible. Given these limitations, finding true heterogeneity through subgroup selection can be a challenging task.

There is a rapidly growing causal inference literature that combines applied econo-

metrics with machine learning for the improved estimation of treatment effects. Machine learning lies at the intersection of computer science and statistics, using learning-based algorithms to make predictions from data. We focus on the branch of machine learning known as *supervised learning*, since its prediction algorithms are particularly useful for solving causal problems (Athey and Imbens, 2017). Supervised learning is based on a predictive model for some outcome Y, as a function of covariate vector X. The idea is to select an algorithm (or a combination of algorithms) that produces optimal predictions of Y given new values of X. Although machine learning is primarily geared towards prediction, economists have started to carefully tune and adapt these methods to effectively answer research questions that are causal in nature (Athey, 2017; Athey and Imbens, 2019; Kleinberg et al., 2015; Mullainathan and Spiess, 2017; Varian, 2014). While prediction tools cannot directly assess causality, they can support classical tools for causal inference in producing more accurate estimates of causal effects (Lechner, 2018; Varian, 2014). This is especially helpful when data originate from non-randomised experiments, where selection bias is an issue (Heckman et al., 1998). For example, machine learning is increasingly used in doubly robust estimation; a method that combines outcome and treatment modelling to adjust for selection bias in causal effect estimates (Robins et al., 2007; van der Laan and Robins, 2003). Incorporating flexible machine learning tools into doubly robust estimation can support the de-biasing of estimators by reducing the risk of model misspecification in the outcome and treatment models (Chernozhukov et al., 2018a; van der Laan and Rose, 2011).

In recent years, machine learning has been incorporated into methods for exploring heterogeneity in causal effects. It offers a practical solution for situations where the analyst has access to a potentially large number of covariates to form subgroups, and limited knowledge on which of these are relevant for heterogeneity (Chernozhukov et al., 2018b). In these high-dimensional settings, correctly specifying a parametric regression for the outcome may prove challenging, particularly if the true model is described by complex interactions. Machine learning algorithms offer a flexible, non-parametric approach to subgroup analysis by selecting covariates of interest in a data-adaptive way, as opposed to *a priori*. These tools can often provide new insights into subpopulations that have not previously been studied (Athey and

Imbens, 2015). Further, some algorithms are able to maintain good performance with high-dimensional data and complex covariate-treatment interactions (Knaus et al., 2021; Powers et al., 2018). One of the more popular machine learning methods for estimating heterogeneous effects is the causal forests estimator, which uses a tree-like structure to control for observed covariates, and estimates the CATE function within each leaf (Athey et al., 2019b). Other notable advances in machine learning methods for exploring treatment effect heterogeneity include algorithms based on the least absolute shrinkage and selection operator (LASSO) (Tian et al., 2014), Bayesian causal forests (Hahn et al., 2020), boosting (Powers et al., 2018), neural networks (Shalit et al., 2017) and ensemble methods (Künzel et al., 2019a; Nie and Wager, 2021).

The aim of this chapter is to provide an insight into causal machine learning, and to highlight some of the recent developments in the literature, in particular for estimating heterogeneous treatment effects. We focus our attention on three promising algorithms: the X-learner (Künzel et al., 2019a), the R-learner (Nie and Wager, 2021) and causal forests (Athey et al., 2019b). We select the former algorithm given its flexibility and intuitive design, and the latter two algorithms given their popularity among applied researchers and ability to control for confounding bias in observational studies. All three algorithms enable statistical inference and construction of valid confidence intervals. Further, each algorithm can easily be implemented through its associated `R` package. We describe each of these methods in turn, and additionally demonstrate their application to an empirical case study.

The structure of this chapter is as follows. First, we introduce the notation and assumptions required for the identification of the causal estimands of interest, the ATE and the CATE, under the potential outcomes framework. Following this, we provide an overview of some key concepts within the area of supervised machine learning. We continue with a review of causal machine learning methods for estimating the ATE, and subsequently explore in detail our three selected algorithms for estimating the CATE. Next, we introduce our impact evaluation case study and share our methods and results. We conclude with a brief overview of some empirical applications of the described methods, and highlight some recent, notable develop-

ments in the literature that extend the methods to other settings. Where possible, we identify the software packages required to implement the methods in `R`.

1.2 Methods of causal inference and machine learning

1.2.1 The causal framework

We define our parameters of interest using the potential outcomes framework of causal inference (Imbens and Rubin, 2015; Rubin, 1974). Suppose we have access to an independent and identically distributed sample of observations, $i = 1, \ldots, N$, and observe (X_i, D_i, Y_i), where X_i is a vector of individual covariates, $D_i \in \{0, 1\}$ is a binary treatment[1], and Y_i is the outcome of interest. We denote $Y_i(d)$ as the potential outcome that would be observed if unit i was assigned to level d of the treatment. The observed outcome Y_i can be written in terms of the potential outcomes as $Y_i = Y_i(1)(D_i) + Y_i(0)(1 - D_i)$. Individual level treatment effects (ITEs) are defined as $\tau_i = Y_i(1) - Y_i(0)$, and the ATE is the expected difference in the potential outcomes under treatment and control for the entire population:

$$ATE = \tau = E[Y_i(1) - Y_i(0)]. \tag{1.1}$$

For a target population of interest, we can estimate the ATE for the treated (ATT) and the controls (ATC).

Since we are interested in estimating how treatment effects vary across the population according to a given covariate profile $X = x$, we require the CATE function, defined as the expected difference in the potential outcomes, as a function of $X = x$ (Imai and Ratkovic, 2013; Imai and Strauss, 2011):

$$\text{CATE} = \tau(x) = E[Y(1) - Y(0)|X = x], \tag{1.2}$$

[1]We consider a binary treatment but for dose-response estimation, a continuous treatment is required.

where $\mu_1(x) = E[Y(1)|X = x]$ and $\mu_0(x) = E[Y(0)|X = x]$ are the conditional expectations of the potential outcomes under treatment and control respectively. Note that the ATE is the expectation of the CATE function, $\tau = E[\tau(X)]$.

CATEs can be evaluated for each individual i using the individual covariate vector X_i to give ITEs, and for population subgroups g that share the the same covariate profile $X = x$, to give group average treatment effects (GATEs). Similar to ATEs, CATEs can also be evaluated for a target population of interest, such as the treated (CATT) and the controls (CATC).

The fundamental problem of causal inference is that for each individual i, we observe either $Y_i(1)$ or $Y_i(0)$, but not both (Holland, 1986). Therefore, $\tau(x)$ cannot be identified without imposing a set of assumptions:

Assumption 1 (Unconfoundedness)

$$(Y_i(0), Y_i(1)) \perp D_i|X_i$$

This assumption, also known as *selection on observables*, states that conditional on the observed covariates, the potential outcomes are independent of treatment status (Rosenbaum and Rubin, 1983). It assumes away the existence of unobserved factors that affect treatment and are also associated with the outcome. However, the plausibility of this assumption must be carefully assessed through rigorous data collection and expert knowledge about potential confounders.

Assumption 2 (Overlap)

$$0 < e(X_i) \equiv P(D_i = 1|X_i = x) < 1, \text{ for all } x \text{ in the support of } X_i$$

This assumption states that the propensity score $e(x)$, defined as the conditional probability of being treated given the observed confounders, must be bounded away from zero and one. This ensures that each unit of observation has a non-zero probability of being selected into treatment. Individuals who are either ineligible for treatment, or are automatically selected into the treatment or control groups, must be removed from the target population.

These two assumptions are jointly referred to as *strong ignorability* (Rosenbaum and Rubin, 1983). If these assumptions hold, then the conditional expectation of the potential outcomes is equal to the conditional expectation of the observed outcomes (Imbens and Wooldridge, 2009). The CATE in (1.2) can then be identified as:

$$\begin{aligned}\tau(x) &= E[Y(1)|X_i = x] - E[Y(0)|X_i = x] \\ &= E[Y_i|X_i = x, D_i = 1] - E[Y_i|X_i = x, D_i = 0] \quad \text{(by Assumption 1)} \\ &= \mu_1(x) - \mu_0(x), \end{aligned} \tag{1.3}$$

where $\mu_1(x)$ and $\mu_0(x)$ are functions of the observed data (also known as the *counterfactual response functions*). The conditional expectation function of the observed outcome under the treatment actually received (also known as the outcome regression) is denoted by $\mu_D(x)$. The ATE can be identified as the expectation of the difference between the response functions under treatment and control: $E[\mu_1(x) - \mu_0(x)]$.

The definition of the CATE in (1.3) suggests that the estimator of $\tau(x)$ can be constructed as:

$$\hat{\mu}_1(x) = \frac{1}{N_1(x)} \sum_{i:X_i=x}^{D_i=1} Y_i$$
$$\hat{\mu}_0(x) = \frac{1}{N_0(x)} \sum_{i:X_i=x}^{D_i=0} Y_i,$$

where $N_d(x)$ is the number of observations under treatment $D = d$ with covariate profile $X = x$. However, this is a crude construction of $\hat{\tau}(x)$ since there will not exist many observations with exactly the same covariate profile. Although this procedure is not followed in practice, it provides an insight into the intuition behind methods for estimating heterogeneous treatment effects. The notion that $\hat{\tau}(x)$ is the difference between the response functions highlights how the causal problem can be transformed into a prediction problem, for which machine learning can be used.

1.2.2 An introduction to supervised machine learning

Supervised machine learning focuses primarily on prediction tasks, that is, predictions of Y_i as a function of X_i, or D_i as a function of X_i. For example, given a dataset in which the outcomes Y_i and the covariates X_i are already labelled (that is, assigned to each unit of observation), the goal of the algorithm is to estimate a model that makes "good" predictions based on the values of Y_i and X_i in the input data (also known as *training* data). The estimated model is then used to predict outcomes for new values of X_i in the remaining, unseen data (also known as *test* data) (Athey and Imbens, 2017). Data are often divided into training and testing (hold-out) samples; the training data are used to "train" (or fit) the model and the test data are used to test the model's predictive power. This idea of splitting a dataset into subsamples for the purpose of evaluating model performance, commonly known as *cross-validation*, has been a feature of the statistics literature for many years (Allen, 1974; Stone, 1974). The objective of cross-validation is to gain an insight into how the selected model will generalise to an independent dataset, which is crucial for predictive algorithms.

Evaluating model performance is usually defined in terms of a loss function, for example, the mean squared error (MSE) or the sum of squared residuals (SSR). The researcher specifies a loss function, for example, the MSE, as an input and searches for a function that minimises the MSE on new observations of X_i from the test data, not including the observations used to fit the model. A frequently used loss function is the *test MSE*, defined as the MSE for observations in the test data. The ultimate goal of prediction algorithms is to achieve good *out-of-sample* predictions by minimising the test MSE. It is straightforward to build an estimator that works well in-sample, but the difficulty lies in ensuring this same estimator is generalisable to unseen data. At this point, we introduce the concept of the bias-variance trade-off. Suppose we estimate a non-linear model with many higher-order terms and vary the degree of the polynomial M in our estimator of predicted outcome Y. When M is low, the estimator is likely to be "underfitting" the data. The simple model is unable to explain the patterns in the data, resulting in a poor training fit and poor generalisation. In other words, there is high *bias* in the predictive model. As M

increases, the model becomes more flexible, thus passing through more data points and subsequently improving the goodness-of-fit. When M is high, the estimator is likely to be "overfitting" the data. The complex model captures more noise in the data and although there is a near perfect training fit, it will fail to generalise on new data. In this case, there is high *variance* in the predictive model. The solution is to find a balance between bias and variance such that it minimises the total prediction error (that is, the expected test MSE):

$$E[(Y - \hat{Y})^2] = Var(\hat{Y}) + [Bias(\hat{Y})]^2 + Var(\varepsilon), \tag{1.4}$$

where Var(ε) is the irreducible error resulting from noise in the outcome itself.

There is a two-step process in finding the optimal balance between bias and variance. The first step is *regularisation*, which penalises models for over-complexity. Regularisation techniques reduce the variance of the model, so that it captures less noise in the data and improves generalisation. The key question is how to select an appropriate level of regularisation; in other words, how to *tune* the algorithm so that it does not underfit or overfit the data. This leads us to the second step, *empirical tuning*, which uses subsamples of the training data to compare and select the level of regularisation that achieves the best performance (that is, the smallest MSE). This procedure can be made more efficient through a process called *K-fold cross validation*, which requires the following steps:

1. Select a prediction algorithm.
2. Randomly divide the dataset into K mutually exclusive subsets of equal size, $k = 1, \ldots, K$. Start with $k = 1$.
3. Select a value for the tuning parameter, $\gamma \in \{\gamma_1, \ldots, \gamma_m\}$.
4. Fit the model for γ on $K - k$ subsets of the data.
5. Test the model for γ on subset k and evaluate the associated loss (for example, test MSE).
6. Return to step 4 and iterate this process over all K folds.

7. Compute the average test MSE for γ over all K folds.

8. Go back to step 3, choosing a different value for γ, and repeat this process.

After K-fold cross-validation, there will exist m values of the tuning parameter and the average test MSE. The optimal tuning parameter will correspond to the one with the lowest average test MSE. The most commonly used values of K are 5, 10 and the sample size minus 1 (also known as *leave-one-out*). In ideal circumstances, a completely separate subset of the training data, known as the *validation data*, will be used to test the predictive power of the algorithm selected through cross-validation. It is important for the training and validation data to be drawn from the same distribution.

Another important concept in machine learning is ensembling, which combines several individual machine learning algorithms, called *base learners*, into one optimal predictive model, through a reduction in bias and variance. The rationale is that a combination of algorithms can outperform a single algorithm in improving model accuracy and goodness-of-fit (Athey et al., 2019a; Surowiecki, 2005; Varian, 2014). Most ensembles use base learners of the same type (that is, homogeneous learners), but there exist ensembles that use base learners of different types (that is, heterogeneous learners) (Zhou, 2012). A simple example of a homogeneous learner is random forests, which combines predictions across several regression trees (Breiman, 2001). An example of a more complex, heterogeneous learner is the winning entry of the famous Netflix Prize competition, which found that combining predictions from many different algorithms led to the greatest improvement in root mean squared error (RMSE) (Bell et al., 2010).

There are three main types of ensemble algorithms: *bagging*, *boosting* and *stacking*. Bagging (short for bootstrap aggregation) aims to reduce variance by drawing random, repeated samples from the training data with replacement. The prediction of the outcome is obtained by averaging across the predictions in the individual bootstrap samples. Boosting aims to reduce bias and follows a similar process to bagging, except that each base learner is estimated sequentially, using information from previously estimated learners. Each learner in the sequence places more weight

on observations with a large prediction error – that is, observations for which the learner incorrectly predicted Y_i from X_i – caused by previous learners in the sequence. The final prediction is obtained by either a vote or a weighted sum of the boosted learners. Boosting and bagging tend to combine homogeneous base learners using deterministic algorithms. Stacking is an alternative type of ensemble method that combines heterogeneous base learners using what is known as a *meta-learner*. The meta-learner takes as inputs the outputs of the base learners, and generates an ensemble prediction. Stacking aims to improve predictive power by finding the optimal combination of base learners. The *super learner*, proposed by van der Laan et al. (2007), is an example of a stacking algorithm that uses K-fold cross-validation to train the meta-learner. It involves evaluating the base learners on the same k-fold split of the training data, and using the out-of-fold predictions to train the meta-learner on how to generate an optimal weighted combination of the predictions. More specifically, the meta-learner selects the combination of predictions that minimises the cross-validated MSE.

1.2.3 Machine learning for estimating average treatment effects

Much of empirical economics is dominated by the study of causal relationships; the effect of treatment D on outcome Y. Researchers are often interested in the counterfactual impact of a change in policy (or treatment) on a given population. They first define a causal estimand of interest (also known as the *target parameter*), for example the ATE, and carefully consider the assumptions required for identification. The causal estimand is then mapped to an estimator via the identifying assumptions. This converts the causal inference problem into an estimation task. Many estimators, especially in settings with observational data, involve estimating parameters that are not of primary interest, but are necessary for estimating the target parameter. Such parameters, often referred to as *nuisance parameters*, enable researchers to obtain unbiased effect estimates by carefully adjusting for confounders under certain causal assumptions. Examples of nuisance parameters include the outcome regression and the propensity score, which are both prediction tasks used to

identify causal effects. These parameters are estimated using an objective function, most commonly the SSR or the likelihood function. They are subsequently used as inputs into the estimating model for the target parameter.

The machine learning literature is less concerned with causality and more with developing algorithms for prediction. Although machine learning can also produce familiar outputs such as regression coefficients, the ability to construct valid confidence intervals (for the majority of algorithms) is currently not possible (Athey and Imbens, 2019). Machine learning is more focused on achieving out-of-sample performance, and for many prediction problems, this is more important than being able to conduct inference. If machine learning cannot produce interpretable estimates of the target parameter, then how can it be applied to problems of causal inference? It can be used to improve the prediction component of estimation tasks, such as the nuisance parameters – the outcome regression $\mu_D(x)$ and the propensity score $e(x)$.

The propensity score was introduced by Rosenbaum and Rubin (1983) as a means of reducing confounding bias in treatment effect estimation using observational data by balancing the distribution of observed confounders between treatment and control groups. Weighting or matching individuals based on their propensity score can produce unbiased estimates of the ATE. The literature on estimating $\mu_D(x)$ and $e(x)$ is vast (Angrist and Pischke, 2008; Greene, 2000; Wooldridge, 2010). In most cases, the nuisance parameters are estimated using parametric models, making them sensitive to model misspecification. Parametric models are usually selected based on theory and expert knowledge. They require strong assumptions about their functional form and can lead to biased treatment effect estimates, if misspecified. For example, there is sometimes a lack of theory to guide the choice of the covariate vector X_i.

Doubly robust estimation that combines the outcome and treatment models is a popular method for reducing the impact of functional form misspecification. The idea is to exploit propensity score matching or weighting prior to further regression adjustment, as a means of fully controlling for confounding bias (Abadie and Imbens, 2006, 2011; Imbens and Wooldridge, 2009). Doubly robust estimators utilise both nuisance parameters and have the special property of being consistent if at least

one of the two parameters is correctly specified (Bang and Robins, 2005; Robins and Rotnitzky, 1995). When both models are correctly specified, doubly robust estimators are semi-parametrically efficient and asymptotically normal (Neugebauer and van der Laan, 2005). Although the nuisance parameters can be estimated using traditional regression models, applying machine learning to nuisance parameter estimation is becoming increasingly popular (Lee et al., 2010; Westreich et al., 2010). These tools enable a more flexible approach to model specification, estimating and comparing many alternative algorithms using K-fold cross-validation, to select the one that minimises the pre-defined loss function.

The problem with directly applying *off-the-shelf* machine learning methods to estimate τ is that the *ground truth* is never observed. In other words, either $Y_i(1)$ or $Y_i(0)$ is observed for each individual. Without the ground truth, validation techniques (such as K-fold cross-validation) cannot be used to compare the performance of different prediction algorithms (Künzel et al., 2019b). Therefore, applying off-the-shelf algorithms to causal problems, without any form of adjustment, can produce biased estimates of the treatment effect. There are two sources of bias in the estimator for τ: *regularisation bias* and bias from overfitting. As mentioned previously, machine learning algorithms use regularisation to prevent over-complexity of the model and to reduce overfitting. Although this decreases the variance of the model, it introduces bias and slower convergence. One approach to overcoming regularisation bias is to use *orthogonalisation* by fitting separate treatment and outcome models, and then regressing the residual of the outcome model on the residual of the treatment model. The estimated τ is then free of regularisation bias (provided that the unconfoundedness assumption is satisfied) since the associations between X_i and D_i, and Y_i and X_i (conditional on D_i) have been partialled out. The other source of bias in the treatment effect arises from flexible machine learning algorithms overfitting the data. When $\mu_D(x)$ is overfitted, the model may capture some of the noise from the error term. This is an issue if the estimation error from $\mu_D(x)$ is associated with the error from $e(x)$, resulting in bias. One approach to tackling bias from overfitting is to use *sample splitting*. This involves partitioning the data into multiple subsamples, fitting the nuisance models on the first subsample, and estimating τ on the second subsample. However, although this reduces bias, this type

of sample splitting can reduce the statistical efficiency and power of the estimator. A better approach is to use *K-fold cross-fitting* that estimates the parameters K times across the K folds, and averages the estimates to obtain a single estimate of $\hat{\tau}$ (Athey and Wager, 2021; Chernozhukov et al., 2018a; Jacob, 2020; Newey and Robins, 2018).

The most powerful doubly robust estimators, such as the targeted maximum likelihood estimator (TMLE) (van der Laan and Rose, 2011; van der Laan and Rubin, 2006) and the double machine learning (DML) estimator (Chernozhukov et al., 2017, 2018a), suggest incorporating machine learning into the estimation of the nuisance parameters. The TMLE uses a two-step approach to de-bias the estimate of the target causal parameter. The first step involves estimating the nuisance parameters (van der Laan et al. (2007) propose to use the super learner), and the second step updates the initial estimates of the outcome regression in a way that is targeted to the causal parameter (for example, the ATE) by correcting for bias in the initial estimator. In technical terms, it does this by adjusting the parametric fluctuation in the initial estimator until the point at which the fluctuation is zero, thus solving the efficient influence curve estimating equation (Gruber and van der Laan, 2010). The TMLE is asymptotically linear which enables statistical inference via asymptotically consistent confidence intervals. The DML estimator combines the residuals from both the outcome regression and the propensity score model to form a new residual-on-residual regression that is based on the partially linear model by Robinson (1988). The estimator usually employs K-fold cross-fitting, so that the nuisance parameters are estimated on one part of the data, and the predictions used to construct the estimate of the target parameter are obtained using the remaining data. This process is repeated K times, and the average estimate of the target parameter across the K folds is the final DML estimate. The standard errors are based on the influence function from semi-parametric statistical theory (Chernozhukov et al., 2018a). In summary, the TMLE and DML estimator both use machine learning to improve the estimation of the target causal parameter without having to rely on an underlying causal model, while also being able to generate valid methods for statistical inference (for example, standard errors) (Díaz, 2020). More generally, they are examples of how machine learning can successfully complement econometric theory

for causal parameter estimation.

1.3 Machine learning for estimating heterogeneous treatment effects

We earlier described that under the assumptions of strong ignorability, the CATE $\tau(x)$ can be identified as the difference between $\hat{\mu}_1(x)$ and $\hat{\mu}_0(x)$. The literature on methods for estimating CATEs is rapidly expanding, and many recently proposed methods incorporate machine learning into the estimation task using, for example, the least absolute shrinkage and selection operator (LASSO) (Imai and Ratkovic, 2013; Tian et al., 2014), tree-based algorithms (regression trees (Athey and Imbens, 2016), random forests (Athey et al., 2019b; Foster et al., 2011; Wager and Athey, 2018), and Bayesian additive regression trees (BARTs) (Hahn et al., 2020)), neural networks (Shalit et al., 2017), boosting (Nie and Wager, 2021; Powers et al., 2018), and meta-learners (Künzel et al., 2019a; Nie and Wager, 2021). These are some of the prominent developments in the literature, but this list is not exhaustive and is continually evolving.

The majority of recent work on heterogeneous treatment effects is based on settings with randomised data. This means that when treated and control groups are likely to be *balanced*, the differences in treatment effects attributable to the observed covariates X_i are interpreted as heterogeneous (Ben-Michael et al., 2020). However, in many fields of economics, obtaining randomised data is often unfeasible or ethically impossible, and non-randomised, observational data are more readily available and used. Most studies conducted using observational data often follow a selection on observables identification strategy for estimating causal effects (either ATEs or CATEs), in which the included set of covariates X_i control for all potential confounders, provided that the assumption of strong ignorability holds. Causal effects conditional on covariates can then be estimated via a non-parametric outcome regression (Hahn, 1998; Rosenbaum and Rubin, 1983, 1984). However, this is not a practical approach in high-dimensional settings where the covariate vector X_i is rich and detailed. Although the unconfoundedness assumption becomes more

plausible as the number of included covariates increases (Rosenbaum, 2002; Rubin, 2009), the behaviour of non-parametric estimators quickly deteriorates (that is, the convergence rate slows) given the sparsity of data in high-dimensional settings – a sparse statistical model is one in which only a relatively few number of parameters are relevant (Stone, 1980). This poor performance is often referred to as the *curse of dimensionality*. There is a fast growing causal machine learning literature for estimating heterogeneous treatment effects that additionally adjusts for selection bias. These methods use a selection on observables framework, and impose additional measures to control for any residual selection bias arising from the observed covariates.

In the next section, we describe three approaches to heterogeneous treatment effect estimation that have gathered attention in recent years: meta-learners, the R-learner and causal forests. We explore these methods in detail and identify those which are promising for estimating heterogeneous treatment effects under various settings. For methods that do not adjust for confounding, we propose a pre-processing step to ensure that treatment and control groups are balanced prior to estimating treatment effects. See Figure 1.8 for a conceptual diagram showing the causal framework for estimating CATEs using the three selected machine learning methods.

1.3.1 Meta-learners

We start by introducing meta-learners to the problem of CATE estimation in the binary treatment case. As described earlier, meta-learners leverage upon information from several heterogeneous base learners to generate an optimal, ensemble prediction. They can be used to decompose the CATE estimation into multiple regression problems that can be estimated with any regression or supervised machine learning algorithm. The choice of base learner can largely influence predictive performance, therefore subject knowledge is highly important. For example, BARTs are appropriate for small datasets with a global structure (for example, sparsity or linearity that applies to the entire dataset), whereas random forests are more suited to larger datasets with a more local structure (for example, sparsity or linearity that applies to portions of the entire dataset). In this section, we discuss two commonly used

meta-learners for estimating heterogeneous treatment effects: the *T-learner* and the *S-learner*, described by Künzel et al. (2019a), and explain some of their strengths and limitations. We then explore a more promising meta-learner, the *X-learner* from Künzel et al. (2019a), that aims to solve some of the challenges encountered by the T- and S-learners, and offers a more favourable approach to CATE estimation in various settings. The T-, S- and X-learners can be implemented using the `causalToolbox`[2] package in `R`.

T-learner

The T-learner (where "T" means two) is the most intuitive meta-learner for estimating $\tau(x)$. It involves a two-step process where first, base learners are used to predict $\hat{\mu}_1(x)$ and $\hat{\mu}_0(x)$ on treated and control subsamples respectively, and second, the CATE is estimated by taking the difference between the predicted values:

$$\hat{\tau}_T(x) = \hat{\mu}_1(x) - \hat{\mu}_0(x) \tag{1.5}$$

This approach has been studied using linear regression (Athey and Imbens, 2016) and tree-based algorithms (Foster et al., 2011) as the base learners. When used with trees, this method is often referred to as the *Two Tree* (TT) estimator.

Despite the simplicity of this approach, the T-learner generally performs poorly in terms of producing unbiased estimates of $\tau(x)$. This is mainly due to the fact that the predicted response functions $\hat{\mu}_1(x)$ and $\hat{\mu}_0(x)$ are trained separately, not together. In other words, the function does not take into account the controls when predicting potential outcomes for the treated, and vice versa. This is especially problematic in study designs where the number of treated and control observations is unbalanced. In most observational studies using administrative or survey data, the number of control observations exceeds the number of treated observations.[3] This disparity can lead to biased treatment effect estimates, if the fitted response functions for treated and controls vary in terms of complexity. For example, the

[2]https://github.com/soerenkuenzel/causalToolbox

[3]In our empirical application, the number of control and treated observations is fairly equal, which is an exception to the norm in non-randomised studies.

fitted control function may be more prone to overfitting, while the fitted treated function may be more prone to underfitting. As a result, their difference may not be a good estimator for $\tau(x)$ (Künzel et al., 2019a).

Another issue with the T-learner is that if the base learners are regularised to solve overfitting, this could unintentionally generate regularisation bias in the CATE estimates. Nie and Wager (2021) demonstrate this point with an example where the LASSO algorithm is used to estimate $\hat{\mu}_1(x)$ and $\hat{\mu}_0(x)$ in the following linear model: $Y_i(d) = X_i\beta_{(d)} + \varepsilon_i(d)$. If the treated and control groups are fitted with two separate LASSO functions:

$$\hat{\beta}_d = \text{argmin}_{\beta(w)} \left\{ \sum_{i:D_i=d} (Y_i - X_i\beta_d)^2 + \lambda_d \parallel \beta_d \parallel_1 \right\}, \tag{1.6}$$

where $\lambda_{(d)}$ is the tuning parameter that penalises the flexibility of the model. The CATE estimator is then $\hat{\tau}(x) = x(\hat{\beta}_1 - \hat{\beta}_0)$. Since $\hat{\beta}_1$ and $\hat{\beta}_0$ are separately regularised towards zero (that is, the regression shrinks the coefficients towards zero), $\hat{\beta}_1 - \hat{\beta}_0$ is regularised away from zero as a result. This means that $\hat{\tau}(x)$ may display heterogeneity even when the true $\tau(x)$ is near or equal to zero. Therefore, the T-learner performs well when there exists heterogeneity in treatment effects. When the treatment effect is simpler, the T-learner performs less well since it is unable to identify and replicate a common behaviour in the treated and control response functions.

S-learner

Closely related to the T-learner is the S-learner (where "S" means single). The key difference is that the treatment D_i is not given any special status, and is considered as just another covariate in the vector X_i. Therefore, instead of estimating separate response functions for treated and controls, CATE estimation is performed in a single step using a combined response function for all observations:

$$\hat{\tau}_S(x) = \hat{\mu}(x, 1) - \hat{\mu}(x, 0). \tag{1.7}$$

The estimated CATE $\hat{\tau}(x)$ is constructed as the difference in the expected outcomes when treatment status changes from 0 to 1. Any heterogeneity picked up by the S-learner is therefore driven by the interaction between D_i and X_i. This approach has been studied with regression trees (Athey and Imbens, 2015) and BARTs (Green and Kern, 2012; Hill, 2011) as the base learners. When used with regression trees, this method is referred to as the *Single Tree* (ST) algorithm.

Since treatment is not given any special status, the base learner used to estimate the response function is not obliged to pick up heterogeneity if it does not exist (that is, when $\tau(x) = \tau$). For this reason, base learners that use regularisation techniques to control model complexity (such as the LASSO and tree-based methods) could end up ignoring the treatment variable if it is not a strong predictor of the outcome. This works well when there exists little or no heterogeneity in treatment effects. However, when heterogeneity does exist, and the treatment indicator is not strongly predictive of the outcome, the S-learner could unintentionally shrink the treatment effects towards zero. In addition, by pooling the data across treated and controls, the S-learner avoids fitting different functions of X_i for $Y_i(1)$ and $Y_i(0)$. This is particularly problematic when effect heterogeneity is strong, and the response functions for the treated and controls are very different. The S-learner does not perform as well in these circumstances.

X-learner

There are challenges associated with both the T-learner and S-learner. First, they are unable to adapt to structural properties of the CATE, if known. For example, prior knowledge about the sparseness or smoothness (that is, less "noise") of the underlying treatment effect could determine the choice of regression or adaptive estimator. Second, they do not perform well in unbalanced study designs due to regularisation bias, as described earlier. Künzel et al. (2019a) propose the X-learner, which builds on the T- and S-learners, and addresses some of the above concerns. It uses information from the control group to better predict treated outcomes, and vice versa. In particular, it can adapt to the structural nature of the CATE, which is useful given that the treatment effect is often linear or constant. The premise

behind the X-learner is to estimate $\hat{\mu}_1$ and $\hat{\mu}_0$ in such a way that, regardless of any variation in the complexity of the fitted response functions for treated and controls, the difference remains a good estimator for $\tau(x)$. The X-learner can be implemented in four stages:

1. Estimate the response functions $\mu_1(x)$ and $\mu_0(x)$ using any non-parametric regression or supervised machine learning algorithm (referred to as base learners of the first stage).

2. Impute the treatment effects based on the response functions from the first stage. The control-outcome estimator (for treated observations) is subtracted from the observed treated outcomes, and the observed control outcomes are subtracted from the treatment-outcome estimator (for control observations) to give the imputed treatment effects:

$$\tilde{\tau}_i^{\,1} = Y_i(1) - \hat{\mu}_0(X_i(1)) \tag{1.8}$$

$$\tilde{\tau}_i^{\,0} = \hat{\mu}_1(X_i(0)) - Y_i(0), \tag{1.9}$$

   where $Y_i(0)$ and $Y_i(1)$ are the ith observed control and treated observations, and $X_i(0)$ and $X_i(1)$ are the associated covariate vectors.

3. Estimate $\tau(x)$ for treated and controls separately by regressing the imputed treatment effects from the second stage $\tilde{\tau}_i$ on the covariate vector X_i. Use any regression or supervised machine learning algorithm (base learners of the second stage):

$$\hat{\tau}_1(x) = \mathbb{E}[\tilde{\tau}_i^{\,1} | X = x] \tag{1.10}$$

$$\hat{\tau}_0(x) = \mathbb{E}[\tilde{\tau}_i^{\,0} | X = x]. \tag{1.11}$$

4. Define the X-learner CATE estimate as the weighted average of $\hat{\tau}_0(x)$ and $\hat{\tau}_1(x)$ from step 3:

$$\hat{\tau}_X(x) = g(x)\hat{\tau}_0(x) + (1 - g(x))\hat{\tau}_1(x), \tag{1.12}$$

   where $g \in [0, 1]$ is a weighting function chosen by the analyst to minimise

the variance of $\hat{\tau}(x)$. Künzel et al. (2019a) recommend using the estimated propensity score $\hat{e}(x)$ as a potential value for $g(x)$, explaining that in study designs with more control observations than treated, $\hat{\tau}(x)$ will be similar to $\hat{\tau}_1(x)$ since $\hat{e}(x)$ will be small. Alternatively, the value of $g(x)$ can also be set to 1 and 0 to appropriately weight the treated and control observations in unbalanced study designs. The X-learner produces an estimate of the ITE based on the respondent's covariate profile X_i. The estimated ITEs can be aggregated to the ATE for the entire population, or the GATE for a pre-specified subgroup of interest.

The X-learner has several advantages compared to the T- and S-learners. It performs particularly well in unbalanced study designs, or when there is prior knowledge about the structural form of the treatment effect and response functions. It can also reduce any errors inherent in the estimated response functions by including the additional imputation step in the CATE construction. Weighting the CATE estimate by the propensity score can further help to stabilise these errors. See Figure 1.9 for a workflow of the X-learner.

In general, when the treatment effect is zero or constant, the X-learner performs better than the T-learner. The best performing meta-learner in this scenario, however, is the S-learner since pooling the treated and control observations is more appropriate when estimating the response function. When there is substantial heterogeneity in treatment effects, and the response functions for treated and controls are very different, the T- and X-learners will perform well since they avoid pooling the data. Although there is no best performing meta-learner given any situation, the X-learner has the overall best performance according to simulations performed by Künzel et al. (2019a). Further, since the X-learner does not make any parametric assumptions about the CATE, approximate confidence intervals can be constructed through bootstrapped samples of the training and test data.

It should be noted that the estimates of $\tau(x)$ from the T-, S- and X-learners are not doubly robust. In a selection on observables framework where confounding adjustment is crucial, additional pre-processing steps would be required to ensure that the distribution of covariates are similar between the treated and controls.

1.3.2 R-learner

Nie and Wager (2021) propose a framework for estimating CATEs that is based on the partially linear model, which is a semi-parametric model containing parametric and non-parametric elements. It was originally formalised by Robinson (1988) but has since become popular in the causal machine learning literature for semi-parametric estimation (Athey et al., 2019b; Chernozhukov et al., 2018a). The partially linear model is written as:

$$Y_i = \mu_0(X_i) + D_i\tau(X_i) + \varepsilon, \tag{1.13}$$

where the shape of $\mu_0(X_i)$ is unspecified. Under the unconfoundedness assumption:

$$E[\varepsilon_i|X_i, D_i] = 0, \text{ where } \varepsilon_i = Y_i - \mu_0(X_i) - D\tau(X_i). \tag{1.14}$$

The CATE function $\tau(x)$ can be rewritten in terms of the conditional mean outcome $m(X_i) = E[Y_i|X_i = x] = \mu_0(X_i) + e(X_i)\tau(X_i)$, and in centred form as:

$$Y_i - m(X_i) = \tau(X_i)(D_i - e(X_i)) + \varepsilon_i. \tag{1.15}$$

This decomposition is referred to as Robinson's transformation and can be used to estimate treatment effect heterogeneity by incorporating modern machine learning tools.

If there is prior knowledge that $\tau(x)$ is constant for some neighbourhood of X_i (that is, $\tau(x) = \tau$ for all $x \in \mathcal{X}$), the partially linear model can be solved locally over $\mathcal{X}$ in three steps:

1. Predict $\hat{e}(X_i)$ and $\hat{m}(X_i)$ using any supervised machine learning algorithm
2. Obtain $\hat{V}_i = D_i - \hat{e}(X_i)$ and $\hat{U}_i = Y_i - \hat{m}_i(X_i)$
3. Regress $\hat{V}_i$ on $\hat{U}_i$ (that is, a residual-on-residual regression) over $\mathcal{X}$ to estimate $\tau(x)$:

$$\hat{\tau}(x) = \frac{\sum_{i:X_i \in \mathcal{N}(x)}(Y_i - \hat{m}^{-i}(X_i))(D_i - \hat{e}^{-i}(X_i)}{\sum_{i:X_i \in \mathcal{N}(x)}(D_i - \hat{e}^{-i}(X_i))^2}, \tag{1.16}$$

where $\hat{m}^{-i}(X_i)$ and $\hat{e}^{-i}(X_i)$ are *leave-one-out* predictions, that is, a form of K-fold cross-fitting where the training data is split into $K = N$ subsamples and predictions are made for each fold k (in this case, $k = i$) using the remaining data. This estimator is semi-parametrically efficient for $\tau(x)$ under the unconfoundedness assumption (Chernozhukov et al., 2018a; Robinson, 1988). However, it also relies on the strong assumption that $\tau(x)$ is homogeneous within a given neighbourhood of X_i, which is often unknown. There are different approaches to selecting the neighbourhood function, including k-nearest neighbours and kernel averaging. A common approach is to use the random forests algorithm to generate kernel weights by averaging tree-based neighbourhoods. We explore tree-based methods in more detail in the next section.

The CATE estimator in (1.16) is designed for estimating locally constant treatment effects. However, Nie and Wager (2021) show that it can be used in combination with the partially linear model to motivate a loss function $L_n(\tau(x))$ that captures a global estimate of heterogeneous treatment effects:

$$\tau(x) = \text{argmin}_{\tau'} \left\{ E[(Y_i - m(X_i) - (D_i - e(X_i))(\tau'(X_i))^2] \right\}. \tag{1.17}$$

The loss function is equivalent to minimising the squared error of the CATE function in (1.15). However, it is dependent on the unknown quantities $m(x)$ and $e(x)$, thus making it unfeasible. The solution is to use a two-step estimator, referred to as the *R-learner* (in recognition of Robinson (1988) and the focus on residualisation):

1. Estimate $m(x)$ and $e(x)$ using any supervised machine learning algorithm.

2. Estimate $\tau(x)$ via a plug-in version of the loss function:

$$\hat{L}_n(\tau(x)) = \frac{1}{n} \sum_{i=1}^{n} ((Y_i - \hat{m}^{-i}(X_i)) - (D_i - \hat{e}^{-i}(X_i))\tau(X_i))^2, \tag{1.18}$$

   where the squared loss $\hat{L}_n(\tau(x))$ is referred to as the *R-loss*.

3. Optimise $\hat{L}(\tau(x))$ using any supervised machine learning algorithm.

4. Tune the algorithm using cross-validation on $\hat{L}(\tau(x))$:

$$\hat{L}_n(\tau(x)) = \frac{1}{n}\{\sum_{i=1}^{n}((Y_i - \hat{m}^{-i}(X_i)) - (D_i - \hat{e}^{-i}(X_i))\tau(X_i))^2 + \alpha_n(\tau(x)\}, \quad (1.19)$$

where $\alpha_n(\tau(x))$ is a regulariser (or tuning parameter) that controls the complexity of the R-loss. Cross-fitting should be used to address any bias from regularisation.

The R-learner has certain advantages over other methods that estimate $\tau(x)$ in a single estimation step (Powers et al., 2018; Shalit et al., 2017; Wager and Athey, 2018). By separating the estimator into various tasks, the structure of the loss function is able to control for any correlation between $e(x)$ and $m(x)$, prior to estimating $\tau(x)$. This separation of tasks also provides more flexibility in terms of the choice of machine learning algorithm used for optimising the R-Loss. The optimisation task is essentially an empirical minimisation problem, which can be solved using various off-the-shelf algorithms, that is, algorithms that do not need to be modified to control for confounding. Therefore, the choice of algorithm will depend solely on its ability to optimise the R-Loss on unseen test data. Since the R-learner is based on semi-parametric efficiency and orthogonality, Nie and Wager (2021) show that if the treatment effect function $\tau(x)$ is simpler than the nuisance functions $m(x)$ and $e(x)$, then $\hat{\tau}(x)$ from (1.19) may converge faster than $\hat{m}(x)$ and $\hat{e}(x)$. In other words, if $\hat{m}(x)$ and $\hat{e}(x)$ are $o(n^{-1/4})$ consistent for $m(x)$ and $e(x)$ respectively in RMSE, then the rate of convergence for $\hat{\tau}(x)$ may be faster and will depend only on the complexity of the treatment effect function.

The R-learner can be implemented using the `rlearner`[4] package in `R`. Although it can be used in combination with any base learner with optimal predictive performance, its use with tree-based causal machine learning methods has sparked a separate, rapidly growing strand of literature linking tree-based methods, causal inference and orthogonalisation.

[4]https://github.com/xnie/rlearner

1.3.3 Tree-based methods

In this section, we explore the family of tree-based machine learning methods that can be used to estimate $\tau(x)$. The literature on heterogeneous treatment effect estimation using tree-based algorithms is continuously evolving. We provide here a few examples of the more notable contributions to the field, starting with the work by Hill (2011), Green and Kern (2012) and Hill and Su (2013) on Bayesian regression trees (BART) (Chipman et al., 2010); an ensemble method that uses the prior to regularise the fitted CATE function. Hahn et al. (2020) further adapt the BART model for problems of causal inference by incorporating the estimated propensity score function into the response model, thus inducing a covariate dependent prior on the estimated CATE function. Foster et al. (2011) apply random forests to estimate effects in treated and control groups separately, taking the differences as predicted values and subsequently fitting regression or classification trees to find covariates strongly associated with treatment effects. Wager and Athey (2018) and Athey et al. (2019b) develop causal estimators based on classification and regression trees (CARTs), providing the first inferential theory for CATEs using the infinitesimal jackknife – a method of estimating confidence intervals and standard errors using variance estimates from the random forests (Efron, 2014; Wager et al., 2014). This contribution moves forest-based methods away from the group of so-called *black box* machine learning estimators, towards those established in causal inference methodology for rigorous asymptotic analysis (Wager and Athey, 2018). The popularity of tree-based algorithms stems from their ability to flexibly model covariate-treatment interactions, even in high-dimensional settings. They also possess attractive asymptotic properties, which we discuss further later in the section.

We begin with a description of the Classification and Regression Tree (CART); a classic yet powerful decision tree algorithm that provides a foundation for tree-based algorithms. Causal trees are an adaptation of the conventional CART algorithm, developed by economists interested in causal inference as opposed to prediction. We describe the framework for estimating CATEs using causal trees. Next, we move onto ensemble tree-based methods, known as forests, which tend to outperform individual trees. Random forests and causal forests represent ensembles of CARTs

and causal trees respectively. We show that by combining the adaptive nature of forest-based algorithms with the orthogonalisation of the R-learner described earlier, causal forests can generate CATE estimates that are robust to confounding bias.

Classification and Regression Trees (CARTs)

CARTs were formalised by Breiman et al. (1984) to aid prediction tasks. The objective is to estimate $m(x)$ using observed covariates and outcomes in a training sample. The algorithm selects an individual covariate and threshold value that minimise the in-sample loss function, and splits the sample into two regions. This process is repeated iteratively. There are two key features of the CART: the initial construction of the tree and the use of cross-validation to select the optimal tree depth.

We use Hastie et al. (2009) to describe the CART framework using the same notation as before, but now introducing p as the number of covariates in the vector X. The CART partitions X into M regions (also known as *terminal nodes* or *leaves*), R_1, $R_2, \ldots, R_M$. The parts of the tree connecting the nodes are referred to as *branches*. Let x_j be a splitting variable, s be a split threshold and let $R_1(j,s) = \{X|X_j \leq s\}$ and $R_2(j,s) = \{X|X_j > s\}$. The algorithm seeks the pair (j,s) that minimises the SSE:

$$\min_{j,s}[\sum_{x_i \in R_1(j,s)} (Y_i - \hat{\delta}_1)^2 + \sum_{x_i \in R_2(j,s)} (Y_i - \hat{\delta}_2)^2], \tag{1.20}$$

where $\hat{\delta}_1$ and $\hat{\delta}_2$ are the estimated conditional means for training observations in R_1 and R_2 respectively. This splitting process is repeated recursively until a tree-like structure is formed. Each region R_m contains either a single observation or a group of observations with similar values of Y. The outcome for unit i in region R_m with covariate profile $X_i = x$ is estimated as the average outcome for all units in the same region. This splitting process, known as *recursive binary partitioning*, is a *top-down*, *greedy* algorithm: top-down meaning that the splitting process begins at the top of the tree; and greedy in that at each split decision, the optimal split is selected at that particular point (as opposed to choosing a split that could produce a better tree in a future split).

Tree depth is a tuning parameter that controls the model's complexity. A larger tree could overfit the data and may not work well out of sample. Alternatively, a smaller tree could underfit the data and may not capture all of the important patterns. The objective is to control the complexity of the model to prevent overfitting, whilst also finding an optimal tree size that maximises predictive performance, that is, optimising the bias-variance trade-off. There are two ways to approach this using regularisation. The first approach is to build the tree such that the decrease in SSR at each split exceeds some specified threshold. However, the problem is that a worthless split higher up in the tree could be succeeded by a much better split lower down. The second, more preferred option is to grow a large tree, stopping only when a minimum node size is realised, and then to *prune* it back using a tuning parameter α to create a subtree $T \subset T_0$. T is obtained by collapsing any number of non-terminal tree nodes. For each value of α, there exists a subtree $T \subset T_0$ that minimises the loss function:

$$L_\alpha(T) = \sum_{m=1}^{|T|} \sum_{x_i \in R_m} (Y_i - \hat{\delta}_m)^2 + \alpha|T|, \tag{1.21}$$

where $\hat{\delta}_m = \frac{1}{N_{R_m}} \sum_{x_i \in R_m} Y_i$, $N_{R_m} = \#(x_i \in R_m)$, and $|T|$ is the number of terminal nodes. Adding a penalty term avoids overfitting since we only consider splits where the improvement in goodness-of-fit is above some threshold. The tuning parameter $\alpha \geq 0$ controls the trade-off between tree size and goodness-of-fit. Larger values of α lead to smaller trees, and vice versa for smaller values of α. The optimal value is selected through K-fold cross-validation, and the subtree that corresponds with the chosen value of α is subsequently used.

Causal trees

Causal trees were developed by Athey and Imbens (2015) as a way of adapting CARTs to problems of causal inference. Conventional machine learning methods, such as CARTs, rely on the ground truth being observed to conduct regularisation for tuning the objective function, which we know is not possible due to the fundamental problem of causal inference. Causal trees modify CARTs in a way that enables causal

parameter estimation. They differ from CARTs in two ways. First, their objective is to estimate heterogeneous treatment effects, as opposed to predicting outcomes. Their resulting leaf estimates R_m represent the difference in the response functions (that is, CATEs) rather than mean outcomes. The CATE estimator $\hat{\tau}(x)$ for any $x \in R_m$ can be constructed as:

$$\hat{\tau}(x) = \frac{1}{|i : D_i = 1, X_i \in R_m|} \sum_{i:D_i=1, X_i \in R_m} Y_i - \frac{1}{|i : D = 0, X_i \in R_m|} \sum_{i:D=0, X_i \in R_m} Y_i \tag{1.22}$$

Second, causal trees are estimated using *honest* methods. This means that the training data used to construct the tree and to estimate the leaf-specific effects is partitioned, so that separate training and estimation samples are used for each task respectively (Athey and Imbens, 2016). Honest methods differ from *adaptive* methods, which are more commonly used in conventional machine learning, and where the same sample of data is used for both tasks. We explain in more detail the concepts of adaptive and honest estimation in the following section. Athey and Imbens (2016) propose an honest approach for constructing unbiased estimates of the MSE of the causal effect of the treatment.

Adaptive and honest trees We start by introducing some new notation. Let $\mathcal{S}^{tr}$ denote the training sample, $\mathcal{S}^{te}$ the testing sample, and $\mathcal{S}^{est}$ the independent estimation sample, where $\mathcal{S}^{est} \subseteq \mathcal{S}^{tr}$. Let Π represent a partitioning of the covariate vector X_i, π be an algorithm that partitions the full sample N into M regions, and $R_m(x, \Pi)$ denote the leaf $R_m \in \Pi$ such that $x \in R_m$.

We define the loss function for minimising the MSE_δ in adaptive CARTs as:

$$L^A(\delta, \pi) = E_{\mathcal{S}^{te}, \mathcal{S}^{tr}} \left[\frac{1}{N^{te}} \sum_{i \in \mathcal{S}^{te}} (Y_i - \hat{\delta}(X_i; \mathcal{S}^{tr}, \pi(\mathcal{S}^{tr})))^2 \right], \tag{1.23}$$

where N^{te} is the number of observations in the testing sample. Note that the same training sample $\mathcal{S}^{tr}$ is used to partition the tree and estimate the conditional means. The tree is evaluated over $\mathcal{S}^{te}$.

Athey and Imbens (2016) challenge the validity of adaptive methods, highlighting

that performing both tasks on the training sample can lead to overfitting. While this is not so much of an issue for CARTs, this can pose certain problems for causal trees since we cannot observe the counterfactual. The estimated CATEs may not reflect true heterogeneity, but noise idiosyncratic to the sample. For example, if there are extreme values of Y_i in the training sample, they could appear in both the tree construction and the estimation tasks. If these extreme values are placed into the same leaf as other extreme values by the estimation algorithm π, the conditional means in S^{tr} would be more extreme (that is, either higher or lower) than they would on a separate estimation sample. This could result in poor coverage probabilities in confidence intervals. The authors suggest that honest methods can overcome the challenges of adaptive methods in two ways. First, by imposing a separation in the data used to grow the tree and to predict $\hat{\tau}$. The loss function for minimising MSE in honest causal trees can therefore be defined as:

$$L^H(\tau, \pi) = E_{\mathcal{S}^{est}, \mathcal{S}^{est}, \mathcal{S}^{tr}} \left[\frac{1}{N^{te}} \sum_{i \in \mathcal{S}^{te}} (\tau_i - \hat{\tau}(X_i; \mathcal{S}^{est}, \pi(\mathcal{S}^{tr})))^2 \right]. \tag{1.24}$$

The tree structure is constructed using $\mathcal{S}^{tr}$, and the predictions $\hat{\tau}$ are obtained using $\mathcal{S}^{est}$. Second, by modifying the splitting function to incorporate the fact that $\mathcal{S}^{est}$ will produce unbiased leaf estimates (thus reducing one aspect of overfitting), and that a larger tree will increase the variance of these estimates. The honest splitting criterion for causal trees aims to minimise the expectation of $\text{MSE}_\tau(\Pi)$, $\widehat{\text{EMSE}}_\tau(\Pi)$, over the test and estimation samples:

$$\widehat{\text{EMSE}}_\tau(\mathcal{S}^{tr}, N^{est}, \Pi) \equiv -\frac{1}{N^{tr}} \sum_{i \in \mathcal{S}^{tr}} \hat{\tau}^2(X_i; \mathcal{S}^{tr}, R) + \left(\frac{1}{N^{tr}} + \frac{1}{N^{est}} \right) \cdot \sum_{\ell \in R} \left(\frac{S^2_{\mathcal{S}^{tr}_{treat}(\ell)}}{p} + \frac{S^2_{\mathcal{S}^{tr}_{control}(\ell)}}{1-p} \right), \tag{1.25}$$

where p is the probability of being in the treatment group, and $S^2_{\mathcal{S}^{tr}_{treat}(\ell)}$ and $S^2_{\mathcal{S}^{tr}_{control}(\ell)}$ are the within-leaf variances of the training sample for treated and control observations. The honest criterion is based on the assumption that $\hat{\tau}$ is constant within each leaf and consists of two terms. The first term rewards partitions that find strong heterogeneity. The second term penalises partitions that create variance in

the within-leaf estimates. A key difference between the honest criterion for CARTs and causal trees is that in the case of CARTs, both terms select covariates that predict heterogeneity in outcomes, whereas in the case of causal trees, both terms select different types of covariates that maximise heterogeneity in treatment effects. Therefore, the greater the heterogeneity in the leaf estimates, or the lower the within-leaf variance, the greater the improvement in $\widehat{\text{EMSE}}_\tau$.

Although the honest splitting criterion is an unbiased estimator for EMSE_τ, repeatedly using the same training sample for constructing the tree can generate some bias. This is because splits higher up in the tree tend to place observations with extreme values into the same leaf. Therefore, after the initial splits, the within-leaf variance of observations in the training sample tends to be lower than the variance would be on an unseen, independent sample. In other words, as the tree is grown deeper, $\widehat{\text{EMSE}}_\tau$ is likely to overstate goodness-of-fit. Cross-validation can help to reduce this bias by evaluating each partition Π using the cross-validation sample $\mathcal{S}^{tr,cv}$ instead of $\mathcal{S}^{tr,tr}$.

Forests

A major limitation of CARTs and causal trees is their instability to small changes in the training data. Leaf estimates are therefore susceptible to high variability, or *noise*. Breiman (2001) suggests that instead of searching for a single optimal tree, averaging across a number of unpruned trees can produce better results. Since individual, unpruned trees have low bias and high variance, ensembling helps to reduce the variance of leaf estimates, as well as smoothing decision boundaries (Bühlmann and Yu, 2002).

Random forests, introduced by Breiman (2001), are an ensemble method that make predictions by averaging across B unpruned and *decorrelated* trees. This decorrelation is attained using two approaches: 1) the forest is constructed using bagging, in that for each tree, $b = 1, \ldots, B$, a bootstrap sample $\mathcal{S}_b$ is drawn from the training data without replacement; and 2) the tree is grown using recursive partitioning, but at each tree node, a random subset m of p covariates is considered for the split

decision, where $m \leq p$. This procedure decorrelates trees that would otherwise split on similar covariates that are (most likely) strong predictors of the outcome (Amit and Geman, 1997). The random forest predicts $\hat{m}(X)$ as the average of B tree predictions $\hat{m}_b(x)$:

$$\hat{m}(x) = \frac{1}{B} \sum_{b=1}^{B} \sum_{i=1}^{n} \frac{Y_i \mathbf{1}(X_i \in L_b(x), i \in S_b)}{|i : X_i \in L_b(x), i \in S_b|}, \tag{1.26}$$

where $L_b(x)$ denotes the set of training examples falling into the same leaf as $X_i = x$ (Athey et al., 2019b).

It has been well studied that random forests can be viewed as a type of adaptive neighbourhood with weights (Lin and Jeon, 2006). This idea was first introduced by Hothorn et al. (2004) for survival analysis and Meinshausen (2006) for quantile regression. They show that this adaptive neighbourhood can be used to weight a set of neighbours for a given test point $X_i = x$, which can subsequently be used to solve a local moment equation for a target parameter of interest (Lin and Jeon, 2006). Although neighbourhood weights are usually obtained by kernel functions, this method does not perform well in high dimensions. Athey et al. (2019b) suggest that forest-based algorithms can be used instead, by averaging neighbourhoods generated across B trees:

$$w_{bi}(x) = \frac{1}{B} \sum_{b=1}^{B} \frac{\mathbf{1}(X_i \in \ell_b(x))}{|L_b(x)|}, \quad w_i(x) = \frac{1}{B} \sum_{b=1}^{B} w_{bi}(x), \tag{1.27}$$

where w_i are the weights that sum to 1. The random forest weighting function works by initially giving equal weight to observations in the training sample that fall into the same leaf as the test point $X_i = x$, and zero weight to the remaining observations. The forest then averages the weightings across all B trees and calculates the frequency with which the i-th training example falls into the same leaf as x. The weights $w_i(x)$ are larger for observations where $E[Y_i|X = X_i]$ is similar to $E[Y_i|X = x]$ (Meinshausen, 2006). These weights can be used to solve for $m(x)$:

$$\hat{m}(x) = \frac{1}{B} \sum_{b=1}^{B} \sum_{i=1}^{n} Y_i \frac{\mathbf{1}(X_i \in \ell_b(x))}{|L_b(x)|} = \sum_{i=1}^{n} Y_i w_{bi}(x). \tag{1.28}$$

Athey et al. (2019b) develop an estimator, which they term *generalised random forests* (GRFs), that can estimate any target parameter of interest $\theta(x)$, as a solution to a local estimating equation of the form:

$$E[\psi_{\theta(x),v(x)}(O_i)|X_i = x] = 0, \tag{1.29}$$

where ψ is a score function, $v(x)$ is an optimal nuisance parameter and O_i is the observable data. Their approach uses the adaptive neighbourhood with weights $w_i(x)$ (obtained using forest-based algorithms) to fit $\theta(x)$ via an empirical version of (1.29)

$$(\hat{\theta}(x), \hat{v}(x)) \in \text{argmin}_{\theta,v} \left\{ \left\| \sum_{i=1}^{n} w_i(x)\psi_{\theta,v}(O_i) \right\|_2 \right\}, \tag{1.30}$$

which simplifies to $\sum_{i=1}^{n} w_i(x)\psi_{\theta(x),v(x)}(O_i) = 0$ when (1.30 has a unique root. GRFs retain several features of the core random forests algorithm, including greedy recursive partitioning, bagging and decorrelation, but no longer make predictions by averaging estimates across an ensemble of trees. Instead, they make predictions over a weighted average of an ensemble of trees. The forest is also grown using honest methods, in that tree construction and estimation of the various parameters are performed on separate subsamples. The main difference in the tree building process is that GRFs search for splits that maximise heterogeneity in $\theta(x)$ across the leaves. Once the forest has been constructed, the GRF algorithm generates predictions using the following steps:

1. Each test point is dropped down each tree to determine which leaf it falls into
2. The neighbouring training points are weighted by the frequency with which the training point falls into the same leaf as the test point
3. Estimates of θ are generated using the weighted list of neighbours and the trained ensemble of trees (that is, a forest).

For prediction problems, $\hat{\theta}(x)$ is the average outcome of the neighbours of test point x. For causal problems, $\hat{\theta}(x)$ is the treatment effect, constructed using the outcomes and treatment status of the neighbours of test point x. The GRF estimator is

asymptotically normal, therefore confidence intervals for $\hat{\theta}(x)$ can be constructed using plug-in or bootstrap methods.

When applied to causal problems, GRFs can be referred to as *causal forests*. For CATE estimation, the aim is to solve the local estimating equation in (1.29), where $\theta(x) = \xi \cdot \tau(x)$ for a given contrast ξ and $O_i = (Y_i, D_i)$. Given the forest weights $w_i(x)$ obtained from the adaptive neighbourhood, we can similarly solve for $\hat{\tau}(x)$ and $\hat{v}$ by minimising the expression in (1.30 The induced estimator $\hat{\tau}(x)$ for $\tau(x)$ is denoted by:

$$\hat{\tau}(x) = \frac{\sum_{i=1}^{n} w_i(x)(D_i - \bar{D}_w)(Y_i - \bar{Y}_w)}{\sum_{i=1}^{n} w_i(x)(D_i - \bar{D}_w^2)}, \tag{1.31}$$

where $\bar{D}_w = \sum w_i(x) D_i$ and $\bar{Y}_w = \sum w_i(x) Y_i$.

Each tree is grown using recursive partitioning, starting at a parent node $P \subseteq X$, in which the solution to the estimating equation is denoted by:

$$(\hat{\tau}_P, \hat{v}_P)(\mathcal{S}) \in \text{argmin}_{\theta,v} \left\{ \left\| \sum \{i \in \mathcal{S} : X_i \in P\} \psi_{\tau,v}(O_i) \right\|_2 \right\}, \tag{1.32}$$

where $\mathcal{S}$ is a sample of data. The parent node P is partitioned into two children $(R_1, R_2) \subseteq X$ in a recursive manner. The aim is to choose the splits that minimise the prediction error of the node, but this minimisation is infeasible since $\tau(x)$ is unknown and is only identified through a local moment condition. Athey et al. (2019b) instead propose maximising the following criterion:

$$\max \triangle(R_1, R_2) := \frac{N_{R_1} N_{R_2}}{N_p^2} (\hat{\tau}_{R_1}(\mathcal{S}) - \hat{\tau}_{R_2}(\mathcal{S})^2, \tag{1.33}$$

where N_{R_1},N_{R_2} and N_P are the number of observations in the respective child and parent nodes. However, maximising (1.33) while also solving for $\hat{\tau}_{R_1}$ and $\hat{\tau}_{R_2}$ in each potential child node can be computationally demanding. Therefore, the authors propose to maximise an approximate criterion $\tilde{\triangle}(R_1, R_2)$ constructed using gradient-based approximations of $\hat{\tau}_{R_1}$ and $\hat{\tau}_{R_2}$, defined as:

$$\tilde{\tau}_C = \tilde{\tau}_P - \frac{1}{|\{i : X_i \in X\}|} \sum_{\{i:X_i \in C\}} \zeta^T A_P^{-1} \psi_{\hat{\tau}_P, \hat{v}_P}(O_i), \tag{1.34}$$

where $\hat{\tau}_P$ and $\hat{v}_P$ are the estimators at the parent node obtained from (1.32), and ζ is

a vector that selects the τ-coordinate from (τ, v). A_P is any consistent estimate for the gradient of the expectation of the ψ-function $\nabla E[\psi_{\hat{\tau}_P, \hat{v}_P}(O_i)|X_i \in P]$, denoted as:

$$\begin{aligned} A_P &= \frac{1}{|\{i : X_i \in P|\}} \sum_{\{i:X_i \in P\}} \nabla \psi_{\hat{\tau}_P, \hat{v}_P}(O_i) \\ &= \frac{1}{|\{i : X_i \in P|\}} \sum_{\{i:X_i \in P\}} (W - \bar{W}_P)^{\otimes 2}, \end{aligned} \tag{1.35}$$

where $\bar{W}_P$ is an average taken over the parent P. The recursive partitioning scheme is now reduced to a two-step process. The first step is the *labelling* step that constructs the pseudo-outcomes ρ_i (that are required for the splits):

$$\begin{aligned} \rho_i &= -\zeta^\top A_P^{-1} \psi_{\hat{\tau_P}(\tau), \hat{v_P}(\tau)}(O_i) \\ &= \zeta^\top A_P^{-1} (D_i - \bar{D}_P)(Y_i - \bar{Y}_P - (D_i - \bar{D}_P)\hat{\tau}_P), \end{aligned} \tag{1.36}$$

where $\bar{Y}_P$ is an average taken over the parent P. The second step is the *regression* step that performs a standard CART regression split on ρ_i, thus splitting P into R_1 and R_2 so as to maximise the following approximate criterion:

$$\tilde{\triangle}(R_1, R_2) = \sum_{m=1}^{2} \frac{1}{|\{i : X_i \in R_m|\}} \left(\sum_{i:X_i \in R_m} \rho_i \right)^2. \tag{1.37}$$

This extensive partitioning continues recursively until the minimum node size is reached, which is either specified a priori or selected from the data via an internal cross-validation procedure.

In their paper, Athey et al. (2019b) present the theory used to construct asymptotically valid confidence intervals for $\tau(x)$ on $\hat{\tau}(x)$ using the delta method. They also propose that the causal forests estimator described above can be improved by *local centring* (motivated by the R-learner in (1.19)) that further address confounding concerns. As previously discussed, the R-learner can be solved using any machine learning algorithm, and causal forests have become a popular solution. The idea is to orthogonalise the forest by partialling out the effect of X_i on Y_i and D_i using Robinson's transformation (Robinson, 1988). As before, the weights are quasi-

automatically obtained from the forest, and a gradient-based approximation is used to compute the psuedo-outcomes for recursive partitioning. The main difference is that D_i and Y_i in (1.36) are now replaced by $\hat{V}_i = D_i - \hat{e}^{-i}(X_i)$ and $\hat{U}_i = Y_i - \hat{m}^{-i}(X_i)$:

$$\rho_i = (\hat{V}_i - \bar{D}_P)(\hat{U}_i - \bar{Y}_P - (\hat{V}_i - \bar{D}_P)\hat{\beta}_P)/Var_P(\hat{V}_i). \tag{1.38}$$

The causal forest implements the orthogonalised forest in a few steps. First, two separate regression forests are fitted to obtain estimates of the nuisance parameters $\hat{e}(X_i)$ and $\hat{m}(X_i)$. Second, these first-stage forests are used to make *out-of-bag* predictions for $\hat{e}(X_i)$ and $\hat{m}(X_i)$, where out-of-bag observations are those that have not been included in the bagged training data (Breiman, 1996a). Third, the residualised treatment $W_i - \hat{e}(X_i)$ and outcome $Y_i - \hat{m}(X_i)$ are constructed, and the causal forest is then trained on these residuals. There are various tuning parameters in the causal forest that can be chosen by cross-validation on the R-Loss in (1.18) (see Appendix 1.B for a description of some of the available tuning parameters). The forests are trained using different values of $\alpha_n(\tau(x))$ for each parameter, and the ones that make out-of-bag estimates of the objective minimised in (1.18) as small as possible are selected. See Figure 1.10 for a causal forest workflow, motivated by the R-learner.

In recent years, a number of algorithms for estimating $\theta(x)$ have been proposed that build upon the GRF framework and are robust to confounding. In particular, Oprescu et al. (2019) introduce *orthogonal random forests* that combines the causal forest algorithm of GRFs with the orthogonality concept from DML. The objective is to estimate $\theta(x)$ at a rate that is robust to the nuisance parameter estimation error. Unlike previous estimators, this approach enables non-parametric estimation of the target parameter on a low-dimensional set of covariates, while also controlling for a high-dimensional set of confounders. The treatment effect estimates are asymptotically normal and enable valid inference.

GRFs can be implemented using the `grf`[5] package in `R`. See Appendix 1.B for details on some of the main training and tuning parameters that can be specified by the user (for a full list, see the `grf` help page).

[5]https://github.com/grf-labs/grf

1.4 Case study: the impact of health insurance on health care service utilisation in Indonesia

We demonstrate the application of the X-learner, R-learner and causal forests algorithms to a policy impact evaluation of a national health insurance programme in Indonesia, the Jaminan Kesahatan Nasional (JKN). Using cross-sectional, individual-level data from Indonesia's national socio-economic household survey (the SUSENAS) in 2017, we evaluate the ATE and CATE, where *treatment* corresponds to being enrolled into the subsidised insurance scheme that targets poor and disabled populations. We are interested in estimating the effects of health insurance on health care utilisation, which we measure using a continuous variable for the total length of inpatient stay (in days) at any public or private medical facility in the previous one year. We extract 80 variables from the SUSENAS dataset to construct the covariate vector X_i, which contains both individual and household characteristics (see Appendix 1.A for a full list). The vector X_i includes two types of variables: *confounders*, which are associated with both the treatment and outcome; and *effect modifiers*, by which the treatment effects vary, as captured by the CATE function. We refer to the existing literature on health insurance and health care utilisation, both in Indonesia and other low- and middle-income countries, to guide our selection of confounders in X_i. As per the impact evaluation of Erlangga et al. (2019a), we control for characteristics that are determinants of health care-seeking behaviour, such as age, sex, marital status, education and the urban-rural distinction. We additionally control for a number of socio-economic factors including employment status and type, poverty status (based on the poverty line by province obtained from the Indonesian Central Bureau of Statistics), monthly household consumption expenditure per capita, dwelling characteristics, use of technology, and measures of social security coverage (Erlangga et al., 2019a; Mulyanto et al., 2019a,b). Finally, we adjust for geographical differences by region (Mulyanto et al., 2019c). We do not include other measures of health care need that are potential mediators in the causal pathway between the treatment and the outcome. An example of a confounder that we include in our model is whether the household is located in an urban or rural

area. Urban residents are more likely to seek care compared to rural residents given easier access to health care facilities. They are also more likely to be beneficiaries of health insurance (OECD, 2019).

We anticipate that the impact of health insurance on health care utilisation varies across the population conditional on certain effect modifiers. To motivate our understanding of potential covariates by which treatment effects may vary, we again refer to the existing literature to support our selection of population subgroups among which we will explore heterogeneity. Erlangga et al. (2019a) perform various subgroup analyses in their impact evaluation of JKN on health care utilisation. They explore variation in treatment effects stratified by certain variables: quintiles of the asset index, urban versus rural location, and supply-side factors. Their results show that the policy increased the probability of inpatient utilisation among the richer quintiles, those living in rural areas, and areas with a high density of health care facilities. Our dataset does not include supply-side variables so we cannot control for these. However, we can estimate effect modification for populations grouped by household wealth (measured by monthly household consumption expenditure per capita, for example), and the urban-rural distinction.

The final dataset contains 912,812 observations, of which 475,930 are in the subsidised treated group and 436,882 are in the uninsured control group. Table 1.3 presents descriptive statistics of selected variables in X that describe the characteristics of the treated and control populations. Compared to the controls, the treated are more likely to be older, educated, in employment and married. Their home is more likely to be self-owned, in a rural location and without basic facilities. They also tend to have lower monthly consumption expenditure per capita and are defined as poor relative to the provincial poverty line. They are also more likely to be beneficiaries of social security assistance compared to the controls. Geographically, there is some regional variation in where the treated and controls are located. Finally, on comparing the baseline difference in the outcome, the average length of inpatient stay for the treated is more than double that for the controls, even though the absolute values are very small. We include all respondents in our model, regardless of whether they had any inpatient stays to avoid complicating the analysis for our

demonstrative purposes.

We provide a step-by-step guide on how to estimate our target parameters with our three selected algorithms using a selection on observables identification strategy. We compare the resulting treatment effect estimates across the alternative methods and draw some conclusions about our impact evaluation problem. Throughout, we specify the functions in `R` that are used to perform the analyses. All tuneable parameters in our fitted models are tuned by cross-validation under the default settings.

Steps:

1. Estimate the nuisance parameters – $m(X_i)$, $\mu_1(X_i)$, $\mu_0(X_i)$ and $e(X_i)$ – using regression-based machine learning methods.

2. Assess the assumption of strong ignorability (that is, unconfoundedness and overlap) to confirm that our causal estimands of interest (that is, the ATE and CATE) can be adequately identified. Assess balance in the covariate distributions of treated and control populations.

3. Perform a simple ordinary least squares (OLS) regression of the outcome (total length of inpatient stay) on treatment (that is, enrolled in subsidised health insurance) to provide a benchmark for the direction of effects.

4. Estimate the ATE, ATT and ATC using methods that are robust to imbalances in the covariate distributions between treated and controls.

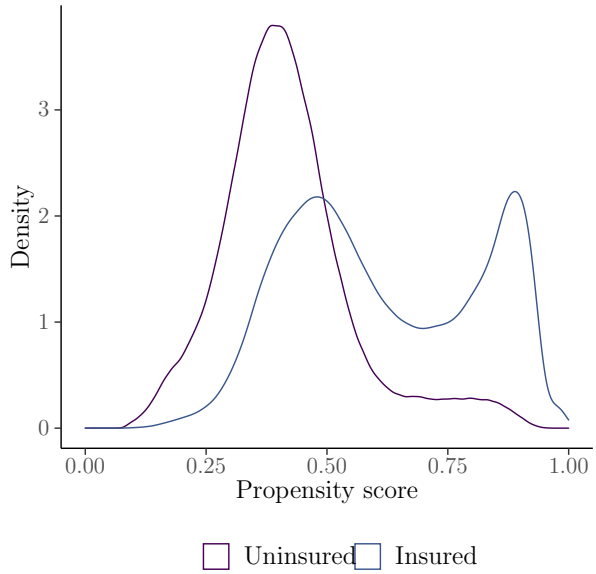
5. Explore individual-level treatment effect heterogeneity by estimating ITEs using our selected machine learning algorithms.

6. Explore subgroup-level treatment effect heterogeneity by estimating GATEs for a selection of effect modifiers.

7. Briefly summarise and discuss findings.

Step 1. We estimate the nuisance parameters using honest random forests (the `regression_forest` function in `grf`). The random forest also produces variance

Figure 1.1: Overlap plot



Note: Density plot showing the distribution of estimated propensity scores for the treated (insured) and the controls (uninsured).

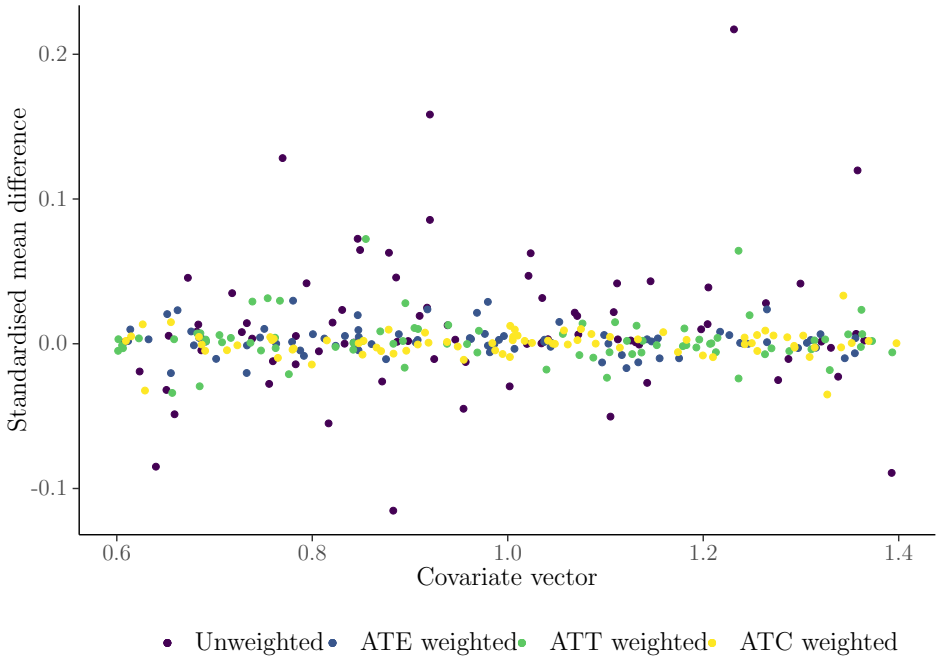
estimates by training small groups of trees and comparing the predictions within and across groups (see Appendix 1.B for details on variance estimation). .

Step 2. We assess our identifying assumption of strong ignorability. Since we use large-scale survey data, which enables a rich set of included controls, we try to minimise unobserved confounding that may explain the process of selection into subsidised insurance. We also review previous impact evaluations of health insurance to guide our selection of confounders and effect modifiers in X, as previously discussed.

We test the overlap assumption by exploring the distribution of predicted propensity scores $\hat{e}(X_i)$. Figure 1.1 presents a density plot of $\hat{e}(X_i)$ for the treated and the control populations. There is a large mass of treated respondents that will almost always receive insurance, which is a potential violation of the overlap assumption. Given this, we change our causal estimands to the ATC and the CATC since we can assume that the overlap condition is sufficiently satisfied for this population.

For each covariate in X, we evaluate the difference in standardised means between the treated and the controls, using a threshold of 0.1. Out of 97 covariates, 5 are

Figure 1.2: Covariate balance between the insured and the uninsured populations



Note: Standardised mean differences are reported for all X between the treated (insured) and the controls (uninsured) before and after inverse probability of treatment weighting for the ATE, ATT and ATC.

not balanced (absolute standardised mean difference (SMD) > 0.1). The covariates with the highest imbalance include a binary indicator for whether the respondent had a social security or family welfare card between August 2016 and March 2017, and a continuous variable for age, suggesting that the probability of being insured is age dependent.

We use inverse probability of treatment weights, constructed using the inverse of the estimated propensity score, that reweight the data to recreate the covariate distributions for the treated and the controls. We generate three sets of weights: ATT weights for the target treated population; ATC weights for the target control population; and ATE weights for the target pooled treated-control population. Although we report all three weights, our focus remains on the ATC due to the overlap issue discussed earlier. Figure 1.2 presents balance statistics on X before and after reweighting the data. Post-weighting, we find that all covariate means are balanced (absolute SMD < 0.1) across the two populations.

Step 3. We perform OLS regressions of the outcome on the treatment, with and

Table 1.1: Ordinary least squares (OLS) regression models

	Model 1	Model 2
Intercept	0.116*	-0.161
	(0.003)	(0.111)
JKN-insured	0.133*	0.138*
	(0.004)	(0.004)
Controls included?	N	Y
Observations (N)	912,810	
R^2	0.001	0.014

Note: OLS regression models with inpatient demand (measured by total length of inpatient stay, in days) as the dependent variable. Standard errors in parentheses. * $p < 0.001$.

without controlling for X_i. Table 1.1 presents the results, showing that being enrolled into subsidised insurance increases the average length of inpatient stay by 0.13 days among the population. This estimate is unchanged when controlling for X_i, suggesting that there is limited confounding bias from the observed covariates. Both estimates are statistically significant at the 0.1% level.

Our OLS regressions are only consistent estimators of the ATE if there is no unmeasured confounding and if the outcome regressions are correctly specified. In the next step, we introduce superior estimators that incorporate additional tools for confounding adjustment when the above criteria is unmet.

Step 4. We propose two alternative estimators for the ATC (for completeness, we also report estimates for the ATE and ATT):

1. Weighted linear outcome regression

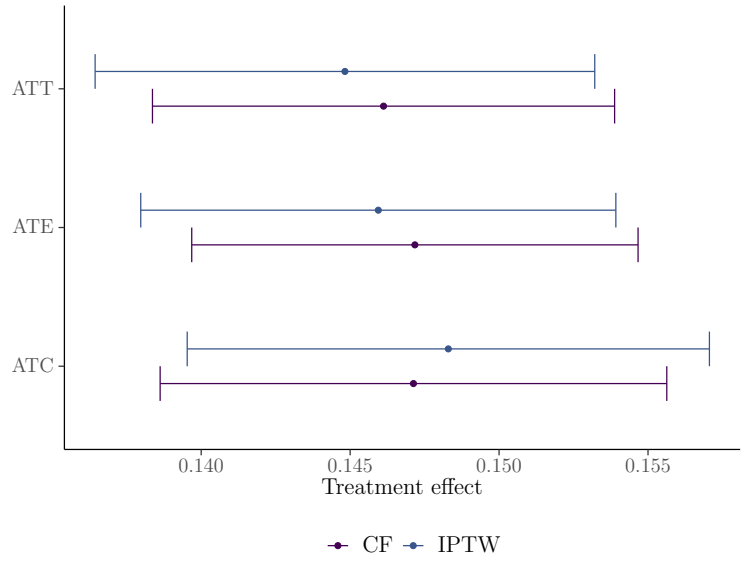
   We take the reweighted data from step 2 and perform a linear regression of Y on X (using `svyglm` from `survey`[6]), which generates an inverse probability of treatment weighted (IPTW) estimate for the ATC.

2. Causal forests

   We use the in-built `average_treatment_effect` function in `grf` to estimate the ATC using the augmented IPTW (AIPTW) estimator (also known as the

[6]https://github.com/cran/survey

Figure 1.3: Estimated ATEs, ATTs and ATCs



Note: Point estimates and 95% confidence intervals (error bars) are reported using the inverse probability of treatment weighting (IPTW) and causal forest (CF) estimators.

doubly robust estimator) (Robins et al., 1994). AIPTW improves upon the previous estimators by combining models for the treatment and the outcome in the same estimator. It produces a consistent estimator for the ATC provided that at least one of the two models is correctly specified. The predicted ITEs from the causal forest (we describe how these are estimated in step 5) are used as inputs into the AIPTW estimator.

Figure 1.3 compares the resulting estimates for the ATC. According to the IPTW-regression estimator, the control group, on average, would increase their length of total inpatient stay by 0.14 days, as a result of being insured. The causal forest (or AIPTW) estimator produces a similar estimate (0.14 days). All estimates are statistically significant at the 5% level, including those for the ATE and ATT.

Step 5. We explore the variation in treatment effects, as a function of the observed covariates, using meta-learners (T-, S- and X-learners) and causal forests (motivated by the R-learner).

Meta-Learners

The implementation of the T-, S- and X-learners varies depending on the choice of base learner. We select honest random forests as the base learner for two reasons: 1) we are interested in obtaining valid confidence intervals, and 2) we have a large dataset with no apparent global structure. The meta-learners are fitted using the nuisance predictions as model inputs. We describe the process of constructing each meta-learner:

1. T-learner

   We construct our ITE predictions $\hat{\tau}_i^T(x)$ by taking the difference between the response functions $\hat{\mu}_1$ and $\hat{\mu}_0$. We also get a non-doubly robust estimate of the ATC by simply taking the mean of of these predictions. To describe treatment effect heterogeneity, we regress $\hat{\tau}_i^T(x)$ on X_i by fitting an honest random forest. We then use the `variable_importance` function in `grf` to rank covariates by how often they were split on at each depth in the forest. Variable importance analysis helps us to detect the most important effect modifiers that drive heterogeneity in treatment effects.

2. S-learner

   We fit an honest random forest that regresses Y_i on X_i and D_i, and evaluate the fitted model at $D_i = 1$ and $D_i = 0$ to generate $\hat{\mu}(x, 1)$ and $\hat{\mu}(x, 0)$ respectively. We take their difference to get $\hat{\tau}_i^S(x)$, and as before, we construct the ATC by averaging the predictions. We then fit another random forest that regresses $\hat{\tau}_i^S(x)$ on X_i, to produce variable importance results.

3. X-learner

   We start by imputing treatment effects for the treated and the controls using the predicted response functions $\hat{\mu}_1(x)$ and $\hat{\mu}_0(x)$. For the treated, we subtract $\hat{\mu}_0$ from the observed treated outcomes $Y(1)$ to give $\tilde{\tau}_i^1$, and for the controls, we subtract the observed control outcomes $Y(0)$ from $\hat{\mu}_0$ to give $\tilde{\tau}_i^0$. Then, we fit separate random forests that regress $\tilde{\tau}_i^1$ and $\tilde{\tau}_i^0$ on X_i. We evaluate the trained forests at the empirical distribution observed in the sample to generate separate CATE estimates for the treated $\hat{\tau}_1(x)$ and the controls $\hat{\tau}_0(x)$. We measure variable importance by identifying the highest-ranked variables

in the separate, treatment-specific forests. The ITE predictions $\hat{\tau}_i^X(x)$ are constructed as a weighted average of the CATE estimates for the treated and the controls, where the weights correspond to the estimated propensity scores $\hat{e}(x)$. The average of the predicted ITEs equates to the ATC.

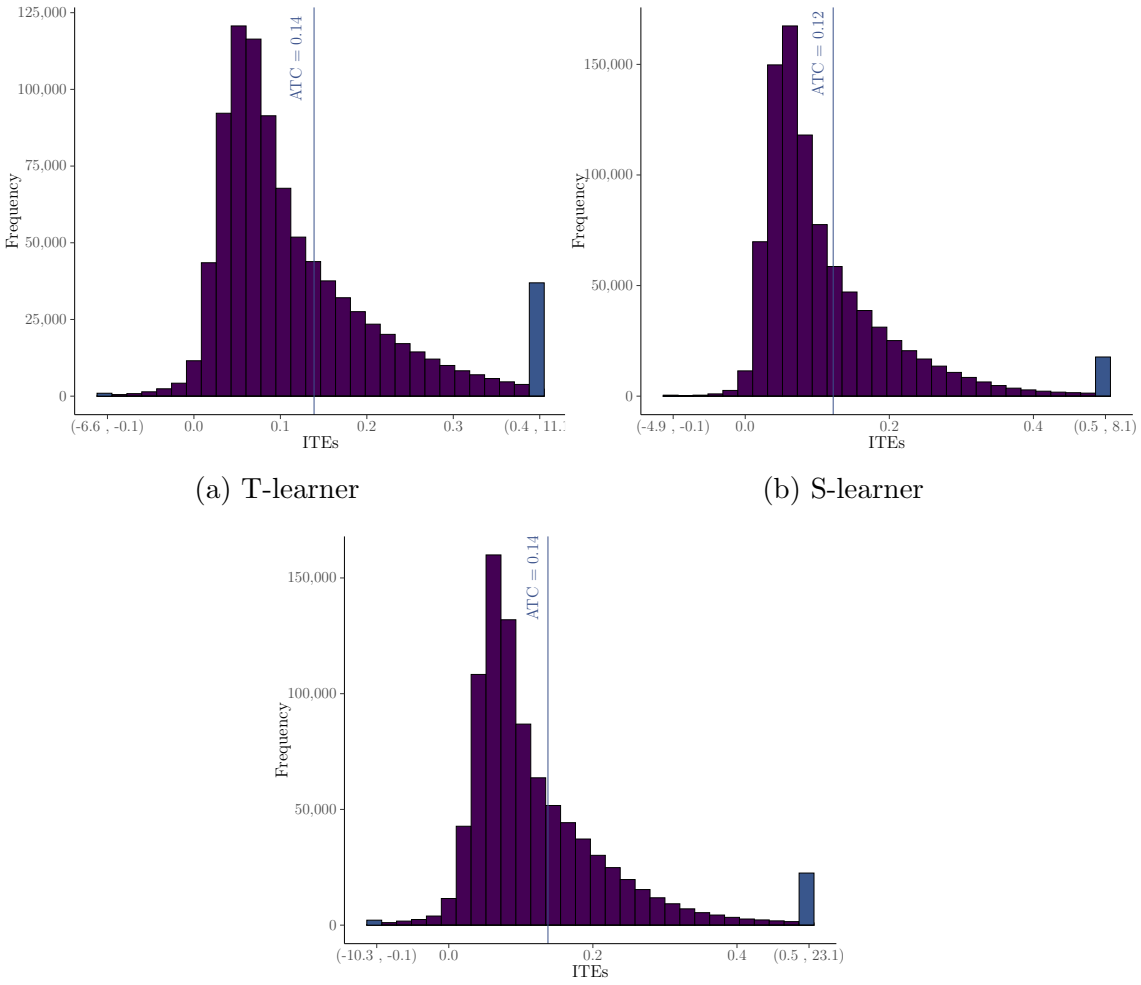
Figure 1.4 presents histograms of the predicted ITEs from the T-, S- and X-learners. We also report the estimated ATC, which for the T- and X-learners is similar at 0.14 days, and for the S-learner is slightly lower at 0.10 days. All ATC estimates are statistically significant at the 5% level. The variation in ITEs is substantial, with the X-learner showing the largest range of estimates (-10 days to 23 days), and the S-learner showing the smallest (-5 days to 8 days). This is expected since the S-learner produces estimates that may be more biased towards zero, as discussed earlier. Despite this variation, the majority of estimates across all meta-learners are concentrated between 0 days and 0.5 days.

Table 1.4 lists the five highest ranked effect modifiers in terms of variable importance, along with their respective percentage of splits, from each of the meta-learners. For the X-learner, there are separate rankings for the treated and the controls since two separate random forests were fitted. The most important effect modifier in the treated group is a binary indicator for whether the respondent is unemployed or retired, which seems plausible since inpatient demand is likely to be associated with age and socioeconomic status. Other important variables include age, marital status and monthly household consumption expenditure per capita. We rely on these variable importance results for our subgroup analyses in Step 6.

The T-, S- and X-learners rely on the flexible outcome models being correctly specified and appropriately adjusting for confounding bias. To improve our treatment effect estimates, we could use additional statistical tools, such as propensity score matching, for example, that performs causal analyses on a subsample of matched treated and control observations with similar characteristics. However, that is beyond the scope of this chapter.

Causal forests/R-learner

Figure 1.4: Estimated ITEs from the T-, S- and X-learners



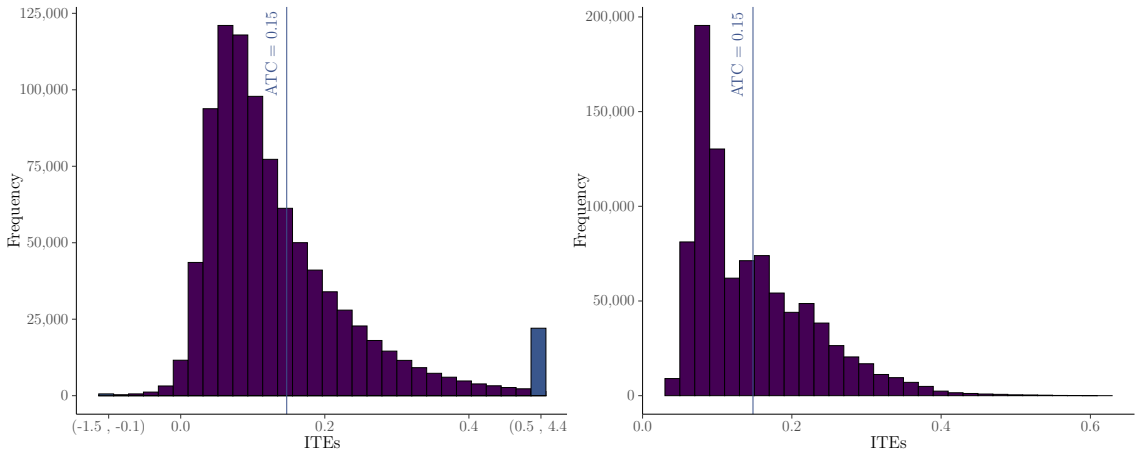
(a) T-learner

(b) S-learner

(c) X-learner

Note: Histograms (with outlier bins) of estimated ITEs from the meta-learners, using honest random forests as the base learner.

Figure 1.5: Estimated ITEs from the causal forest, based on the R-learner



(a) Causal forest trained on all variables in X

(b) Causal forest trained on selected variables in X

Note: Histograms (with outlier bins) of estimated ITEs from the causal forest, based on the R-learner. Estimates are made out-of-bag.

We use the `causal_forest` function in `grf` to train an honest causal forest on the residualised treatment and outcome models. We produce out-of-bag predictions $\hat{\tau}^{-i}(x)$ evaluated at the empirical distribution of the sample. In addition, we use `average_treatment_effect` to generate a doubly robust estimate of the ATC (0.15 days). Figure 1.5a plots the distribution of estimated ITEs, which shows a small range (-2 days to 4 days). Similar to the meta-learners, the majority of estimates are concentrated between 0 days and 0.5 days. We perform variable importance analysis on the trained forest, as before, to rank the effect modifiers. Table 1.4 lists the top five, with age being the most important.

As per the analysis of Athey and Wager (2019), and motivated by the work of Basu et al. (2018), we additionally train a second causal forest, using only those effect modifiers that saw an above-average proportion of splits from the variable importance analysis on the first causal forest. The residualisation used in fitting the first causal forest eliminates any residual confounding effects from including only selected effect modifiers. Figure 1.5b plots the resulting distribution of estimated ITEs from the second causal forest fitted on selected effect modifiers. Compared to the forest fitted on X, the range of ITEs is even smaller (0 days to 0.6 days). The estimated ATC, however, is similar at 0.15 days.

Although the two causal forests display some form of heterogeneity, these results are not sufficient to confirm that the predicted ITEs $\hat{\tau}^{-i}(X_i)$ are a better estimate of $\tau(X_i)$ than the ATE $\hat{\tau}$. Therefore, we perform two tests to determine whether the heterogeneity in $\hat{\tau}^{-i}(X_i)$ and $\tau(X_i)$ is similar:

1. We estimate doubly robust ATCs for two subpopulations by grouping observations based on whether their ITE estimate $\hat{\tau}^{-i}(X_i)$ (in absolute terms) is above or below the median. This analysis provides a qualitative assessment of the strength of heterogeneity, by separating those with "high" and "low" CATCs. Table 1.2a shows that the difference in the ATC (high-low) of 0.14 days is statistically significant at the 5% level, suggesting that some level of heterogeneity exists. We can also compare the average characteristics of observations in the control population that are grouped by high and low CATCs. Table 1.5 presents the results, which shows that respondents with an above-median CATC are more likely to be female, older, in work and married. They are also more likely to be socioeconomically better off according to their household and financial characteristics.

2. Second, we use the `test_calibration` function in `grf` that performs an omnibus test of the calibration of the causal forest. This alternative test for heterogeneity is motivated by the "best linear predictor" method of Chernozhukov et al. (2018a) that aims to find the best linear fit of $\tau(X_i)$ using $\hat{\tau}^{-i}(X_i)$. The residualised outcome $Y_i - \hat{m}^{-i}(X_i)$ is regressed onto two synthetic predictors, $J_i = \bar{\tau}(D_i - \hat{e}^{-i}(X_i))$ and $K_i = (\hat{\tau}^{-i}(X_i) - \bar{\tau})(D_i - \hat{e}^{-i}(X_i))$, where $\bar{\tau}$ is the average of $\hat{\tau}^{-i}(X_i)$. The coefficient on K_i ("differential.forest.prediction") measures the quality of the CATE estimates, with a coefficient of 1 indicating that the estimates are well calibrated. The coefficient on J_i ("mean.forest.prediction") absorbs the ATE, with a coefficient of 1 indicating that the ATE estimate is valid. Table 1.2b presents the results of the calibration test, confirming a valid ATE estimate and the presence of observable heterogeneity in treatment effects.

Step 6. So far, we have explored treatment effect heterogeneity at the individual level. However, we are also interested in how treatment effects vary across different

Table 1.2: Results from tests for heterogeneity

	Estimate	SE
(a) Estimates of high and low ATCs		
High ATC	0.2189*	0.0072
Low ATC	0.0784*	0.0044
Difference in ATC (high-low)	0.1405*	0.0085
(b) Omnibus test		
mean.forest.prediction	0.9992*	0.0277
differential.forest.prediction	1.0587*	0.0690

Note: SE = standard error; *p-value<0.05.

subpopulations, grouped according to certain covariates of interest. These covariates can be selected pre- or post-analysis, depending on the research plan. We select our covariates via both approaches: a pre-analysis, theory-driven variable selection process based on previous subgroup analyses in the existing literature, and a post-analysis, data-driven variable selection process based on our variable importance analyses (see Table 1.4). Our selected effect modifiers consist of continuous variables – age and monthly household expenditure per capita – and binary variables – marital status, employment status and urban/rural location.

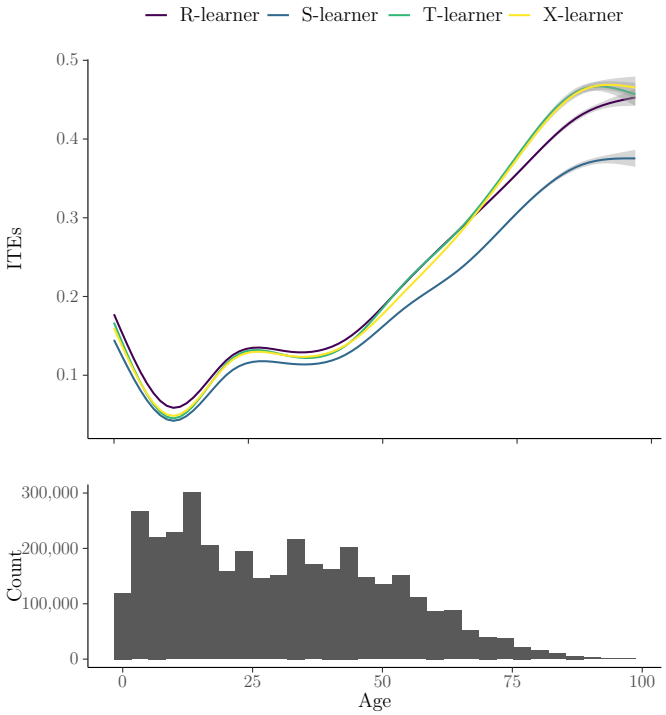
First, we construct smooth plots of the estimated ITEs (with 95% confidence intervals) from the T-, S-, X- and R-learners, as a function of the chosen continuous covariates $X = x$, using the `gam` smoothing function in `ggplot2`[7]. Figure 1.6 presents the results, showing that the policy impact on health care utilisation increases with age and monthly household consumption expenditure. The marginal increase in demand is larger among poorer individuals.

Second, we estimate doubly robust GATCs for our selected effect modifiers using only the R-learner. We use the `average_treatment_effect` function to aggregate the estimated ITEs to the ATC as before, only now we restrict the sample to observations with a given value x of the covariate X. For continuous variables of interest, we group observations into bins or quartiles. Figure 1.7 presents the results, which supports our earlier findings that treatment effects increase with age and household expenditure. Treatment effects are particularly large for the unemployed, however this includes retirees which could reflect an age effect. We additionally find that

[7]https://github.com/tidyverse/ggplot2

Figure 1.6: Estimated ITEs as a function of $X = x$

(a) Age

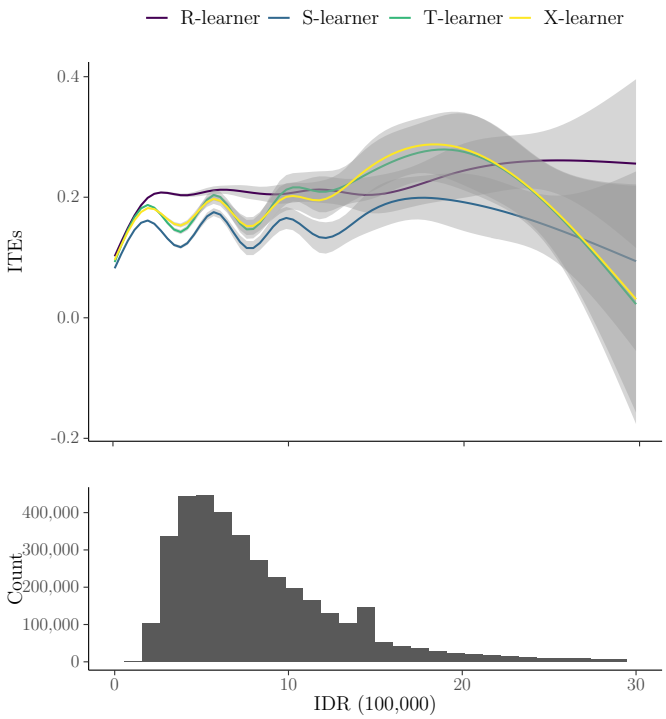


Note: Smooth plots of the estimated ITEs (with 95% confidence intervals) and their associated histogram. Smoothing uses the `gam` method in `ggplot`.

treatment effects are larger for married and widowed respondents (again, a likely age effect), and those that live in urban areas.

Step 7. Our results provide additional insights into the average and heterogeneous impacts of subsidised JKN on inpatient hospital utilisation. We find that overall, there is a positive policy effect on the population, in that subsidised health insurance is associated with an increase in health care utilisation, which is supported by the OLS, IPTW-regression and causal forest (AIPTW) estimators. These findings are in line with previous evaluations of health insurance in Indonesia and other low- and middle-income countries (Erlangga et al., 2019a,b). However, given the demonstrative purpose of this case study, we acknowledge that there are certain limitations in our evaluation approach. In our model, we do not control for respondents' health. Health status, in this particular context, can serve two purposes: it can be a confounder or a mediator or both. If health status is a potential confounder, and we

(b) Monthly household consumption expenditure per capita (excluding outliers > IDR 3,000,000)

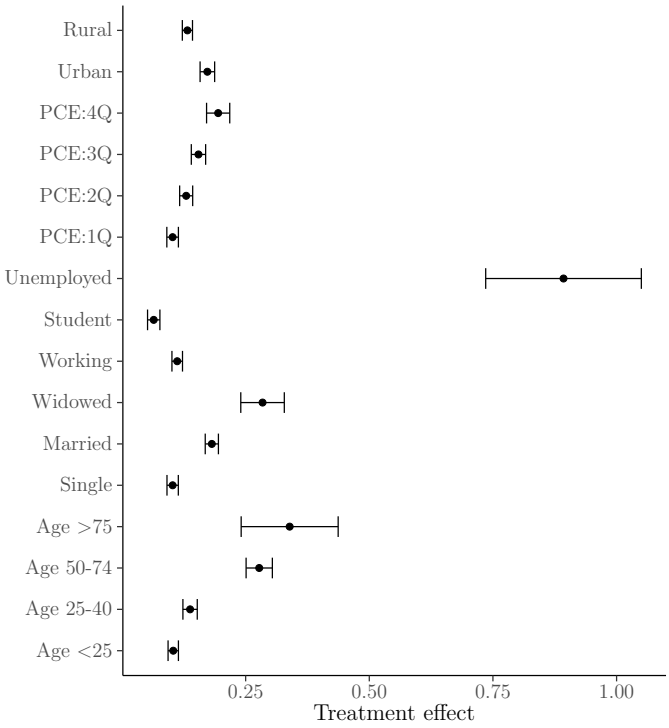


Note: Smooth plots of the estimated ITEs (with 95% confidence intervals) and associated histogram. Smoothing uses the `gam` method in `ggplot`.

do not include it in our model, we are at risk of producing biased treatment effect estimates. A limitation of using cross-sectional data, however, is that we are unable to identify the nature of the health status variable. A more detailed evaluation could also explore the variables in X_i that violate the overlap condition in the treated population, and investigate whether they are true confounders. If so, it may be more appropriate to create a matched sample or to remove observations from the sample that contribute to this violation, prior to any causal analyses. Additionally, in the outcome regression, it may also be worth accounting for the substantial mass point at zero (97% of respondents did not report any inpatient utilisation) with a two-part model. We have highlighted a few of the limitations in our simplified analysis – a more rigorous evaluation is required to obtain robust policy impact estimates.

We find that all of the machine learning algorithms produce similar average impact estimates (the S-learner estimate is more biased towards zero), however the range of ITE estimates varies substantially. Despite this, the majority of estimated ITEs

Figure 1.7: Estimated GATCs for selected effect modifiers



Note: GATCs are estimated using the causal forest. Point estimates and 95% confidence intervals are reported. Q1 = lowest quartile.

across all algorithms are located in the region of 0 days to 0.5 days. We discover that, in addition to our pre-defined variables of interest, additional variables – such as age and employment and marital statuses – are key drivers of heterogeneity. These data-driven subgroups provide a novel contribution to the evidence base for analysing treatment effect modification, since they have not previously been identified as subgroups of interest in this policy context. We find that heterogeneity is driven by age, employment status and socioeconomic factors, in that the elderly population, those who are out of work, and households with a high consumption expenditure respond positively to subsidised health insurance. These findings can provide valuable information for policymakers interested in identifying the main beneficiaries of the health policy.

1.5 Conclusion

In this chapter, we have provided an insight into the current causal machine learning literature on estimating heterogeneity in treatment effects from observed data. We have defined some of the key concepts behind supervised machine learning and have discussed its recent implementation into the causal inference methodology for estimating the ATE and CATE in a selection on observables framework. We have described in detail three causal machine learning algorithms for estimating heterogeneous treatment effects – meta-learners, the R-learner and causal forests – and have explained their relative strengths and limitations. Meta-learners offer the most intuitive method for estimating CATEs. They are user-friendly, transparent and model agnostic. The X-learner, in particular, offers a flexible approach to CATE estimation and performs particularly well under specific scenarios, such as unbalanced study designs, or when the structure of the response functions is known. A limitation of meta-learners, however, is that if the outcome models are misspecified, they cannot fully adjust for confounding bias without implementing additional statistical tools. The R-learner, on the other hand, adjusts for confounding since it is founded on semi-parametric theory and uses orthogonalisation. This makes the R-learner particularly suitable for non-randomised settings. A common application of the R-learner is via causal forests, which combines the R-learner with the adaptive framework of the random forests algorithm. Through a modified splitting criterion and the introduction of honesty, causal forests offer a flexible non-parametric approach to CATE estimation, with the ability to construct confidence intervals.

We demonstrated the application of the described methods to an impact evaluation of Indonesia's national health insurance programme on health care utilisation. The case study resembles that of a typical non-randomised empirical application that relies on large-scale survey data with a binary treatment variable and a rich covariate vector of confounders and effect modifiers. The primary aim of our case study was to show the current scope of causal analyses using the described methods, most notably that CATEs can be estimated for individuals according to their individual covariate profile X_i (ITEs), and for population subgroups that share the same covariate profile

$X = x$ (GATEs). However, we highlight that GATEs can only be estimated for population subgroups that have been selected by the researcher in advance, either prior to CATE estimation through some form of pre-specification plan, or after CATE estimation once variables that drive heterogeneity have been identified by the algorithm. In other words, a current limitation of the described methods is their inability to automatically identify GATEs without some level of researcher input.

The causal machine learning literature continues to grow rapidly. Of particular note are a group of methods that extend the meta-learner framework to allow for doubly robust estimation of $\tau(x)$. The proposed methods are based on a two-stage regression estimator that constructs doubly robust "scores" (based on the doubly robust estimator for the ATE) in the first stage using nuisance predictions, and in the second stage regresses these scores on X_i to obtain $\hat{\tau}(x)$ (Fan and Wu, 2020; Foster and Syrgkanis, 2019; Kennedy, 2020; Lee et al., 2017; Luedtke and van der Laan, 2016b; Semenova and Chernozhukov, 2021; van der Laan, 2006; van Der Laan and Dudoit, 2003; van der Laan and Luedtke, 2015; Zimmert and Lechner, 2019). We briefly highlight a few examples. Kennedy (2020) proposes the DR-learner that incorporates sample splitting (specifically, cross-fitting) into the two-stage estimator to reduce bias and to generate a general result that is "model agnostic" – meaning that any method can be used to fit the models – about the first- and second-stage methods, provided that some conditions are met. Knaus (2022) propose the normalised DR-learner (NDR-learner) that builds on the DR-learner by normalising the inverse probability of treatment weights that are included in the construction of the doubly robust scores, in order to stabilise $\hat{\tau}(x)$ by preventing extreme estimates. However, this approach restricts the choice of machine learning method to those that make predictions using convex combinations of the outcomes (for example, random forests).

Other proposed methods in the literature are particularly suited to high-dimensional settings since they use a concept known as *dimensionality reduction*. The idea is to fit the CATE model on a smaller subset of covariates $V \in X$, while still controlling for the full covariate vector X in the estimation of the nuisance models to sufficiently adjust for confounding. Lee et al. (2017) use the two-step doubly robust estimator

described earlier and assume parametric specifications for the nuisance models to avoid the curse of dimensionality, but allow non-parametric estimators of the CATE function. Abrevaya et al. (2015) consider a similar approach but instead of the doubly robust estimator, they consider the IPTW estimator, which can be sensitive to any misspecification of the propensity score model. Fan et al. (2022) consider the two-step doubly robust estimator but allow flexible machine learning algorithms to estimate the nuisance parameters. For the CATE estimation, they propose a traditional local linear non-parametric regression.

These developments are particularly welcome for those seeking to employ causal machine learning methods for empirical applications, where observational data and confounding are common features of the study design. The data challenge from the 2018 Atlantic Causal Inference Conference provides a number of examples of how machine learning can be applied to estimate treatment effect modification in a non-randomised study (Carvalho et al., 2019). The eight participants used a diverse set of methods based on semi-parametric and ensemble models, and there was also a particular focus on incorporating tools to address confounding. In terms of other applied examples in the broader literature, Appendix 1.C displays a selection of economic papers that use machine learning methods to estimate heterogeneous treatment effects. There is a mixture of randomised and non-randomised studies, with causal forests being the most popular algorithm. There is also a large variation in sample size, which highlights the flexibility of these methods.

It is evident that causal machine learning methods are becoming increasingly popular in study designs that adjust for observed confounding. However, methodological developments for study designs that identify causal effects using a selection on unobservables identification strategy are still in their infancy. In these designs, including observed covariates is not enough to control for confounding due to the presence of unobserved heterogeneity. Instrumental variable analysis is a common tool for estimating causal effects when the unconfoundedness assumption is violated. The idea is to identify causal effects by exploiting an exogenous source of random variation via an instrument. There is a growing literature combining causal machine learning and instrumental variables analysis for estimating average and heterogeneous treatment

effects. For example, Chen et al. (2021) propose a DML approach to instrumental variable quantile regression (IVQR) for estimating low-dimensional causal parameters in high-dimensional data (Chernozhukov and Hansen, 2005). Athey et al. (2019b) extend the GRF algorithm to instrumental variables regression. As before, the honest random forest estimates heterogeneity in causal effects using local moment conditions, but is now identified using the conditional two-stage least squares (2SLS) estimator. Chen and Hsiang (2019) combine the above two methods by incorporating IVQR into the GRF algorithm, which produces quantile treatment effects and variable importance measures for heterogeneity. An alternative method of controlling for unobserved heterogeneity is to utilise panel data methods. The extension of causal machine learning methods from cross-sectional to longitudinal settings is another area of development in the literature. In terms of effect modification, Semenova et al. (2017) produce an estimation and inference method for high-dimensional panel data settings with a large number of heterogeneous treatment effects, using a two-stage method, which they term the *orthogonal lasso*.

Gaining new and important insights into the distributional impacts of treatment can have far-reaching implications for policy. In particular, there is a rich, rapidly evolving literature on optimal policy learning that leverages upon heterogeneous treatment effects to allocation treatment efficiently in the presence of constraints, such as budget and equity (Athey and Wager, 2021; Kallus, 2018, 2020; Kallus and Zhou, 2018a; Kitagawa and Tetenov, 2018). For example, in our impact evaluation case study, policymakers may seek to subsidise health care insurance only for those populations who are likely to benefit from enrolment, in order to maximise some target criterion. As discussed earlier, these population subgroups can be identified according to their observed characteristics. This growing area within policy impact evaluation that effectively "learns" treatment assignment rules through understanding heterogeneity in treatment effects further highlights the scope of machine learning in methods of causal inference, and their combined potential for applied economic research.

Figure 1.8: Conceptual diagram of the causal framework for estimating CATEs

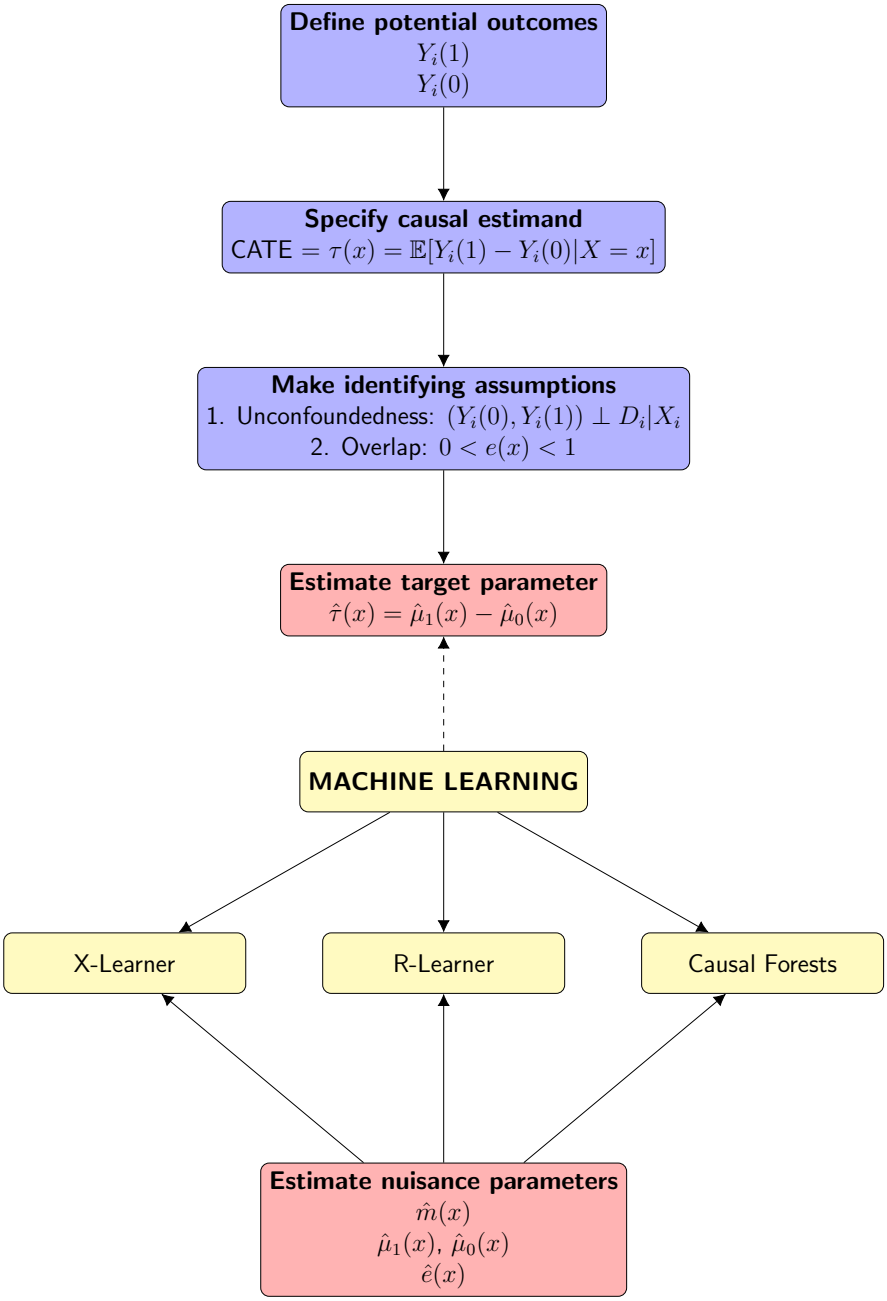


Figure 1.9: X-learner workflow

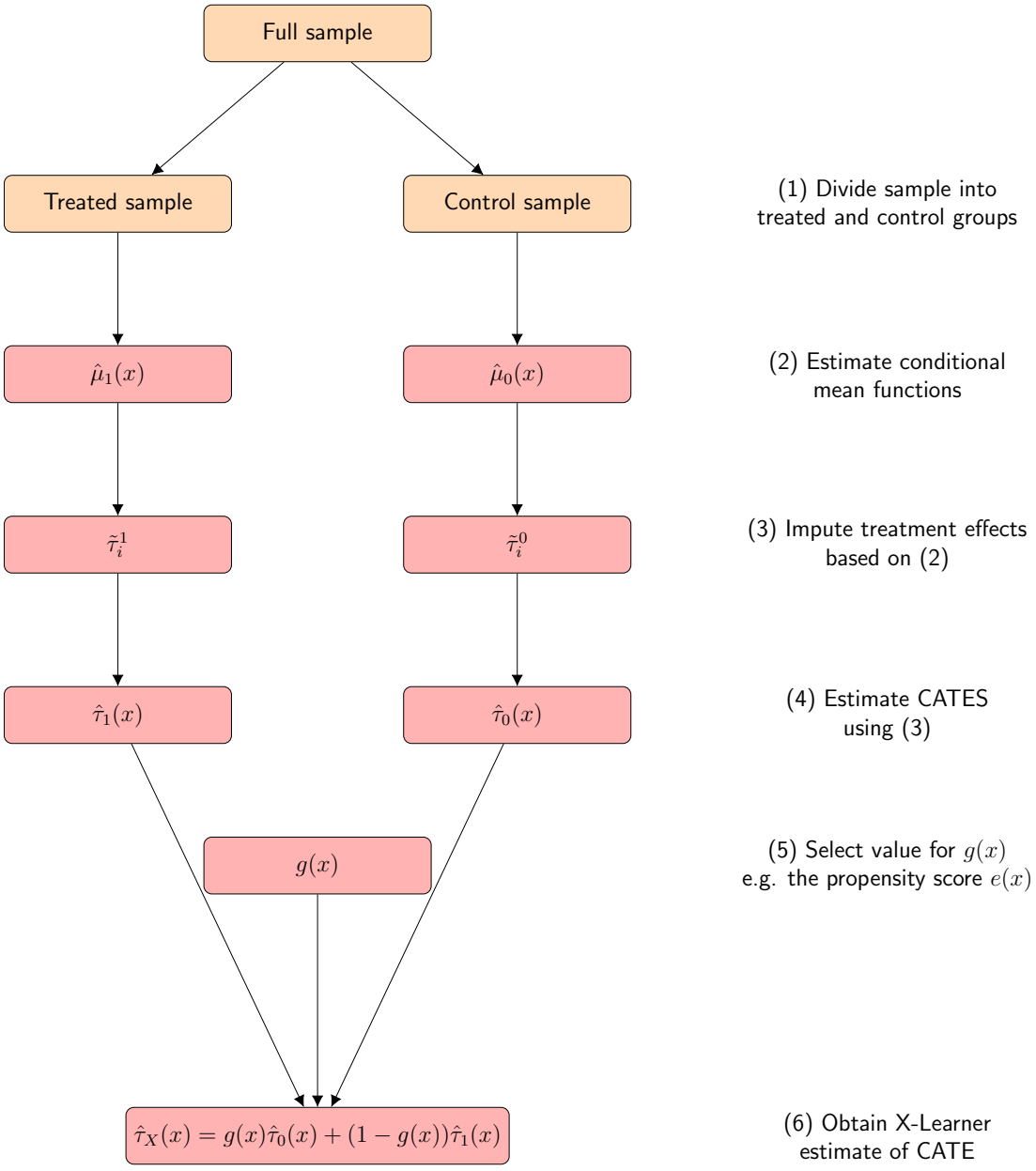


Figure 1.10: Causal forest workflow motivated by the R-learner

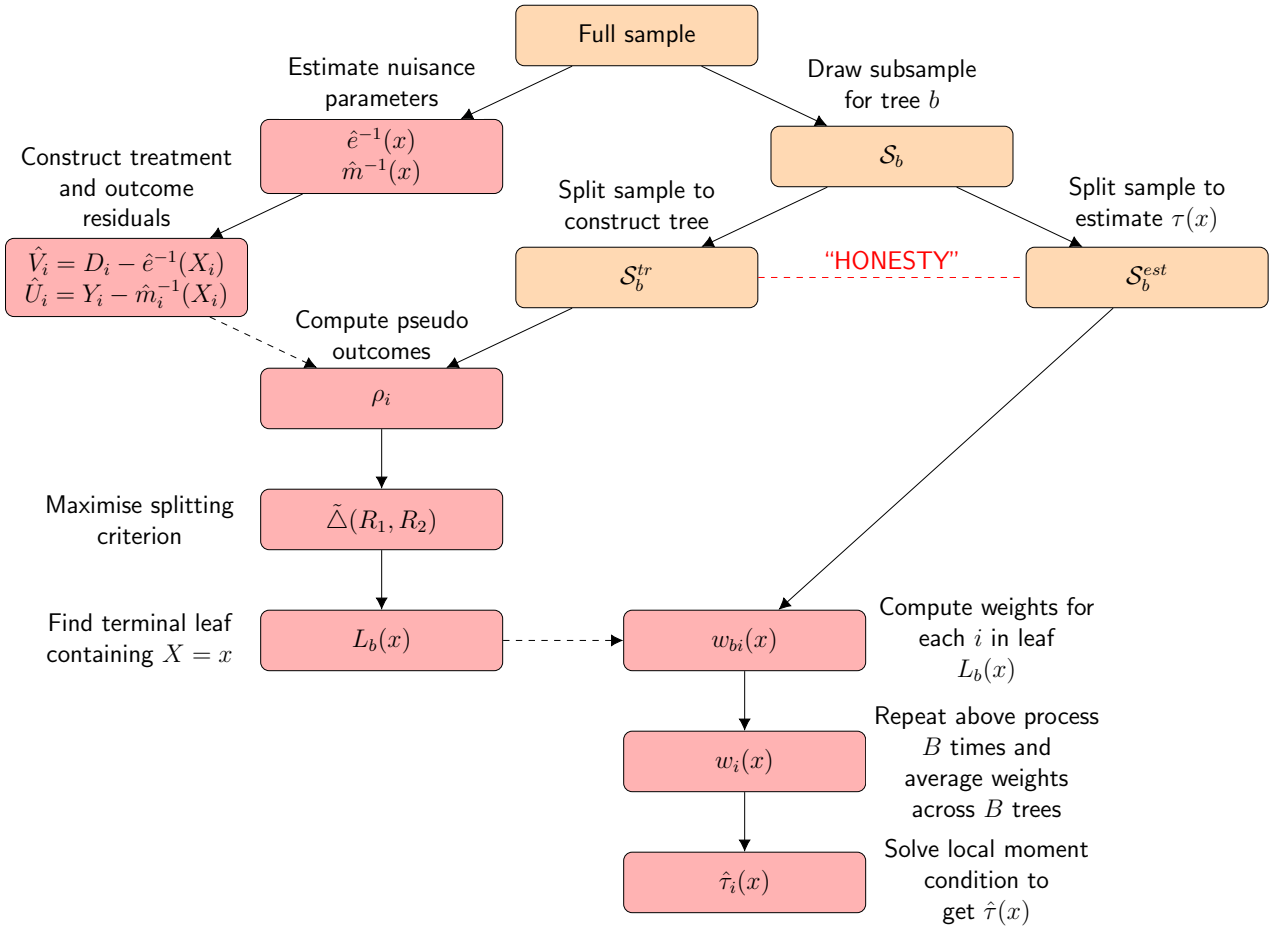


Table 1.3: Descriptive statistics

	Uninsured (n=436,882)		**Insured (n=475,930)**		
	Mean	SD	Mean	SD	SMD
Outcomes					
Total length of inpatient stay (days)	0.116	1.181	0.249	2.132	0.077
Household member characteristics					
Male	0.504	0.500	0.498	0.500	-0.005
Age	28.5	20.5	31.7	20.0	0.158
Education: compulsory	0.561	0.496	0.624	0.484	0.063
Employment status: in employment	0.429	0.495	0.472	0.499	0.043
Employment status: student	0.157	0.363	0.181	0.385	0.025
Marital status: married	0.440	0.496	0.481	0.500	0.042
Used internet in previous 3 months	0.209	0.407	0.197	0.397	-0.013
Household characteristics					
Location: urban	0.381	0.486	0.358	0.479	-0.023
Number of people in household	4.556	1.864	4.678	1.894	0.065
Home occupancy status: owner	0.833	0.373	0.855	0.352	0.022
Toilet: private	0.849	0.358	0.794	0.404	-0.055
Purchases drinking water	0.396	0.489	0.367	0.482	-0.029
Electricity	0.961	0.195	0.933	0.250	-0.028
Natural disaster in previous year	0.132	0.339	0.167	0.373	0.035
Received subsidised rice (Raskin) in past 4 months	0.335	0.472	0.464	0.499	0.128
Had a social security card (KPS)/family welfare card (KKS) between August 2016 - March 2017	0.079	0.270	0.296	0.457	0.217
Not enough food to eat in previous year	0.258	0.438	0.304	0.460	0.046
Savings account	0.437	0.496	0.405	0.491	-0.032
Monthly consumption expenditure per capita (IDR 100,000)	8.810	7.242	8.196	6.489	-0.089
Poverty status: poor	0.128	0.334	0.160	0.366	0.032
Region					
Region: Sumatera	0.282	0.450	0.303	0.460	0.021
Region: Jakarta	0.008	0.089	0.016	0.125	0.008
Region: Java	0.298	0.457	0.253	0.435	-0.045
Region: Bali,NTB,NTT	0.088	0.284	0.084	0.277	-0.005
Region: Kalimantan	0.120	0.325	0.070	0.254	-0.050
Region: Sulawesi	0.131	0.338	0.154	0.361	0.023
Region: Maluku-Papua	0.073	0.260	0.120	0.325	0.047

Note: Sample means and standard deviations (SD) are reported for selected variables in X for the uninsured and insured (enrolled into subsidised JKN) populations. SMD = standardised mean difference.

Table 1.4: Variable importance results from the T-, S-, X- and R-learners

Ranking	Effect modifier	Importance
T-learner		
1	Age	0.27
2	Marital status: single	0.16
3	Marital status: widow(er)	0.10
4	Employment status: student	0.09
5	Employment status: in employment	0.08
S-learner		
1	Age	0.25
2	Marital status: single	0.17
3	Marital status: widow(er)	0.11
4	Employment status: student	0.08
5	Employment sector: primary	0.06
X-learner: treated		
1	Employment status: out of employment	0.18
2	Age	0.14
3	Marital status: single	0.10
4	Monthly consumption expenditure per capita	0.10
5	Employment status: student	0.05
X-learner: controls		
1	Employment status: out of employment	0.24
2	Age	0.18
3	Marital status: single	0.11
4	Marital status: widow(er)	0.09
5	Employment status: student	0.06
R-learner		
1	Age	0.26
2	Marital status: single	0.15
3	Monthly consumption expenditure per capita	0.10
4	Employment status: in employment	0.07
5	Marital status: widow(er)	0.05

Note: Top 5 important effect modifiers are reported based on the variable importance ranking from the trained forests. Importance is measured as the weighted sum of the frequency with which the variable was used to split on at each depth in the forest.

Table 1.5: Descriptive statistics for the control population, grouped into high and low treatment effects

	Low effect		**High effect**	
	Mean	SD	Mean	SD
Outcomes				
Total length of inpatient stay (days)	0.072	0.871	0.161	1.430
Household member characteristics				
Male	0.648	0.478	0.355	0.479
Age	16.7	11.7	40.6	20.5
Education: compulsory	0.537	0.499	0.587	0.492
Employment status: in employment	0.308	0.462	0.552	0.497
Employment status: student	0.298	0.457	0.011	0.105
Marital status: married	0.194	0.396	0.692	0.462
Used internet in previous 3 months	0.283	0.450	0.134	0.340
Household characteristics				
Location: urban	0.387	0.487	0.375	0.484
Number of people in household	4.881	1.813	4.222	1.856
Home occupancy status: owner	0.821	0.384	0.846	0.361
Toilet: private	0.843	0.364	0.856	0.351
Purchases drinking water	0.404	0.491	0.388	0.487
Electricity	0.957	0.203	0.964	0.185
Natural disaster in previous year	0.137	0.344	0.127	0.333
Received subsidised rice Raskin in past 4 months	0.339	0.473	0.332	0.471
Had a social security/family welfare card	0.089	0.285	0.069	0.253
Not enough food to eat in previous year	0.269	0.444	0.247	0.431
Savings account	0.445	0.497	0.429	0.495
Monthly consumption expenditure per capita (IDR 100,000)	8.428	6.750	9.202	7.696
Poverty status: poor	0.142	0.350	0.113	0.317
Region				
Region: Sumatera	0.281	0.450	0.282	0.450
Region: Jakarta	0.008	0.087	0.008	0.090
Region: Java	0.288	0.453	0.308	0.462
Region: Bali,NTB,NTT	0.095	0.294	0.081	0.273
Region: Kalimantan	0.119	0.324	0.121	0.326
Region: Sulawesi	0.122	0.328	0.140	0.347
Region: Maluku-Papua	0.087	0.281	0.059	0.236

Note: Descriptive statistics are reported for the control population only, grouped into "high" or "low" effects according to whether predicted CATEs are below or above the sample median CATE. SD = standard deviation.

Appendix 1.A List of all covariates used for confounder adjustment

Household member-level (binary)

Male

Female

Marital status: single

Marital status: married

Marital status: divorced

Marital status: widow(er)

Has a national identity number

Literacy: Latin letters

Literacy: Arabic letters

Literacy: Other letters

Education: Compulsory

Education: Non-compulsory

Travelled domestically for tourism in 2016

Victim of crime between March 2016-February 2017

Had a cellphone in previous 3 months

Used a computer in previous 3 months

Used internet in previous 3 months

Employment status: working

Employment status: student

Employment status: housekeeper

Employment status: other activities

Employment status: unemployed/retired

Employment sector: primary

Employment sector: secondary

Employment sector: tertiary

Smokes electric cigarettes

Smokes tobacco

Household member-level (continuous)

Age

Number of cigarettes smoked per week

Household-level (binary)

Location: urban

Did not have enough food to eat in previous year

Home occupancy status: owner

Home occupancy status: renter

Home occupancy status: rent-free

Home occupancy status: company-owned

Home occupancy status: other

Has a second home

Roof: concrete

Roof:tile

Roof: asbestos

Roof: zinc

Roof: bamboo

Roof: wood/shingle

Roof: straw/fiber/leaves/metroxylon sagu

Toilet: private

Toilet: shared

Toilet: none

Drinking water: bottled

Drinking water: tap

Drinking water: pump

Drinking water: protected well

Drinking water: unprotected well

Drinking water: protected spring

Drinking water: unprotected spring

Drinking water: river

Drinking water: rain

Drinking water: other

Purchases drinking water

Has electricity

Has experienced a natural diaster in previous year

Has natural tourism in residential area

Has received subsidised rice (Raskin) in past 4 months

Has received Smart Program (PIP) between August 2016 - March 2017

Has had a social security card (KPS)/family welfarecard (KKS) between August 2016 - March 2017

Has received family of hope program (PKH) between August 2016 - March 2017

Has a savings account

Poverty status: poor

Household-level (continuous)

Number of people in household

Number of children in household

Number of infants in household

Number of households in census building/house

Number of families in census building/house

Monthly consumption expenditure per capita (IDR 100,000)

Regional-level (binary)

Region: Sumatera

Region: Jakarta

Region: Jawa

Region: Bali, NTB, NTT

Region: Kalimantan

Region: Sulawesi

Region: Maluka-Papua

Appendix 1.B List of training and tuning parameters in `grf 1.2.0`

Training parameters:

- `sample.fraction` controls the proportion of data used to construct each tree (default is 0.5)
- `num.trees` is the number of trees in the forest (default is 2000)
- `honesty` determines whether honest forests are trained (this is the default)
- `honesty.fraction` controls the proportion of training data used in tree splitting (default is 0.5)
- `honesty.prune.leaves` determines whether empty leaves are pruned away after training to ensure each tree can handle all test points (this is the default)
- `mtry` selects the number of variables considered for each split (default is $\sqrt{p} + 20$)
- `min.node.size` is the minimum number of observations in each tree leaf (default is 5)
- `alpha` controls the maximum imbalance of a split (default is 0.05)
- `imbalance.penalty` controls the penalty imposed on imbalanced splits (default is 0).

Tuning parameters:

- `tune.num.trees` selects the number of trees in each "mini forest" (default is 50)
- `tune.num.reps` selects the number of forests (default is 100)
- `tune.num.draws` is the number of random parameter values considered when choosing the optimal parameter values (default is 1000).

The training and tuning parameters listed above represent those available in the `regression_forest` function. The `causal_forest` function uses the same training and tuning parameters, except that the split balancing parameters are modified for causal splitting. This is because in causal settings, the number of treated vs control observations (and not just the overall number of observations) in each node is important to obtain a good estimate of the treatment effect. In causal splitting, `min.node.size` reflects the minimum number of treated and control observations in each tree leaf (default is 5), and `alpha` and `imbalance.penalty` measure how much information is captured in each node, given by $\sum_{X_i \in P}(W_i - \bar{W})^2$.

All of the above training parameters (apart from `num.trees`) can be tuned via cross-validation using the `tune.parameters="all"` option in `regression_forest`. To use this option, the researcher must only select values for `num.trees` and the tuning parameters (default values can be used). To identify the values of the training parameters selected by cross-validation, the `tunable.params` option can be included in the main function.

Variance estimates can be obtained by providing the `estimate.variance` attribute to the `predict` function. The `ci.group.size` parameter in the `regression_forest` function controls the number of trees in each small group (default is 2). The variance is estimated by training trees in each small group, and comparing predictions within and across groups.

See the `grf` help page for more details on all available parameters and features.

Appendix 1.C Applied economic papers that use machine learning to estimate heterogeneous treatment effects

Article	Sector	Algorithm	Data type	N
"Contemporaneous and Post-Program Impacts of a Public Works Program" (Bertrand et al., 2017)	Labour; Development	Causal forests	Randomised	12,188
"Using Causal Forests to Predict Treatment Heterogeneity: An Application to Summer Jobs" (Davis and Heller, 2017)	Labour	Causal forests	Randomised	6,850; 4,894
"Heterogeneous Employment Effects of Job Search Programs" (Knaus et al., 2022)	Labour	Lasso	Observational	85,198
"Targeting with Machine Learning: An Application to a Tax Rebate Program in Italy (Andini et al., 2018)	Labour	Decision trees; k-Nearest neighbour; Random forests	Observational	3,646
"What Is the Value Added by Using Causal Machine Learning Methods in a Welfare Experiment Evaluation?" (Strittmatter, 2019)	Labour	Random forests	Randomised	33,614
"A Nonparametric Bayesian Analysis of Heterogenous Treatment Effects in Digital Experimentation" (Taddy et al., 2016)	Business	CARTs; Random forests	Randomised	21,000,000

"Estimating Treatment Heterogeneity of International Monetary Fund Programs on Child Poverty with Generalized Random Forest" (Daoud and Johansson, 2019)	Development	Causal forests	Observational	1,940,734
"Causal Tree Estimation of Heterogeneous Household Response to Time-Of-Use Electricity Pricing Schemes" (O'Neill and Weeks, 2018)	Environmental	Causal forests	Randomised	4,225
"Uncovering Sociological Effect Heterogeneity Using Tree-Based Machine Learning" (Brand et al., 2021)	Labour	Causal forests	Observational	4,584

Chapter 2

Estimating Heterogeneous Impacts of Subsidised Health Insurance: A Causal Machine Learning Approach

The impacts of health and social policies can vary according to the observed characteristics of the recipients. Exploring treatment effect heterogeneity is becoming increasingly popular within policy evaluation research. Developments in causal machine learning have rapidly advanced the scope of this work, enabling researchers to estimate effect modification in a flexible, data-adaptive way. In this paper, we combine predictive and causal machine learning to evaluate the impact of the subsidised component of Indonesia's national health insurance programme, the JKN, on the utilisation of inpatient and outpatient health care in 2017. We fit a causal machine learning algorithm, causal forests, to estimate heterogeneous treatment effects, and a predictive algorithm, the super learner, for prediction tasks (that is, estimating the nuisance parameters). To address the abundance of zeros in the utilisation outcomes, we decompose the outcome model into zero and non zero counts using a two-part model. This enables a separate exploration of policy impacts on the decision to seek care and the quantity of care consumed. We summarise and interpret treatment effect heterogeneity using a number of theoretically motivated approaches, including data-driven subgroup analyses and linear projections. Overall, we find positive average impacts of JKN on health care demand and find evidence of treatment effect heterogeneity, for example, some recipients increase demand more than others. For those that decide to access care as a result of being insured, the increase in the quantity of care demanded is particularly pronounced. The policy effect is modified by a set of theoretically motivated covariates (for example, age, household expenditure and health care accessibility), as well as covariates that are identified from our

data-driven approach (for example, employment status and technology usage).

2.1 Introduction

Universal health coverage is a policy priority in many low- and middle-income countries around the world (Banerjee et al., 2021). Over the past decade, there has been a renewed effort to implement national health insurance reforms with the aim of ensuring that everyone can access key health care services without suffering financial hardship (Lagomarsino et al., 2012). Indonesia's Jaminan Kesehatan Nasional (JKN), which was introduced in 2014, is a prime example of these efforts, being the world's largest single payer system and achieving population coverage of over 80% by 2019 (Maulana et al., 2022). Evaluating the impact of large-scale health and social policies, such as the JKN, is essential for understanding whether the policy has proved successful, and for whom. These insights can also provide valuable lessons for other countries considering similar policies. Recent statistical advances have increased the scope of policy evaluation research, most notably the integration of machine learning into the causal inference toolbox. Machine learning offers a flexible, non-parametric and data-driven approach to modelling the relationship between an outcome and a set of covariates, especially in high-dimensional settings where the number of observations and covariates is large (Mullainathan and Spiess, 2017). It has gained traction as a popular alternative to traditional regression models for prediction problems, and has more recently been adapted to estimate causal parameters (Athey, 2017; Athey and Imbens, 2017, 2019; Kleinberg et al., 2015).

One research area in which machine learning is having a profound impact is heterogeneous treatment effect estimation. The majority of recent impact evaluations of health insurance have focused on estimating the overall average treatment effect (ATE) (Erlangga et al., 2019a). Exploring the variation in treatment effects across the population that can be attributed to differences in observable characteristics, can offer important additional insights. In Indonesia, there exist substantial disparities in disease burden and access to adequate, affordable health care, particularly among rural and low-income populations (Agustina et al., 2019). Removing financial bar-

riers to care through universal health coverage may have a disproportionate impact on these population subgroups that can only be identified with an assessment of heterogeneity. Traditionally, heterogeneous treatment effects have been estimated via simple interactions between the treatment variable and potential effect modifiers, however this approach is prone to issues concerning multiple hypothesis testing and potential "cherry-picking" of results (Assmann et al., 2000). It also requires making parametric assumptions, even as data-dimensionality increases, and more complex treatment-covariate interactions may be necessary (Davis and Heller, 2017).

Recently proposed causal machine learning methods address the above limitations by using a more structured approach to subgroup analyses that estimates effect modification in a data-adaptive way (Athey et al., 2019b; Hahn et al., 2020; Künzel et al., 2019a; Nie and Wager, 2021; Powers et al., 2018; Shalit et al., 2017; Tian et al., 2014). The "causal forests" estimator, developed by Athey et al. (2019b), is a prominent example of these methods that combines the flexibility of predictive machine learning with the potential outcomes framework of causal inference to estimate the conditional average treatment effect (CATE) function, which captures heterogeneity in treatment effects. Causal forests predict heterogeneity by searching over possibly the entire covariate space, rather than a few pre-selected subgroups of interest, and maintain strong predictive performance even when there are a large number of true effect modifiers. Estimates of the nuisance parameters – the outcome regression and the propensity score – are required as inputs into the causal forest estimator. Predictive machine learning can be used to choose nuisance models that fit the observed data best, using measures of predictive performance. A popular prediction algorithm is the "super learner", which data-adaptively selects or combines prediction algorithms from a user-defined library of candidates (van der Laan et al., 2007).

In this paper, we apply predictive and causal machine learning to a policy impact evaluation of the JKN on two measures of health care utilisation: outpatient and inpatient demand. Given the excess of zero counts in the outcomes, we propose a two-part model that decomposes the outcome distribution into zero and non-zero counts. This decomposition additionally enables us to separately evaluate the effects

of the policy on the decision to seek care (which we refer to as the "participation" decision) and the quantity of care consumed (the "consumption" decision). We evaluate the average and heterogeneous effects of being enrolled into the component of JKN that subsidises insurance premiums for the poor, in comparison to having no insurance. We use causal forests and the super learner to estimate the CATE and nuisance functions respectively. For the nuisance parameters, we fit different models for the intensive and extensive margins, and then one for the propensity score. We summarise treatment effect heterogeneity using theoretically motivated approaches, including estimating ATEs for population subgroups specified according to values of predicted CATEs or effect modifiers, and finding the best predictors of the CATE through linear projections (Athey and Wager, 2019; Chernozhukov et al., 2018b; Kennedy, 2020; Knaus et al., 2021; Semenova and Chernozhukov, 2021). We expect that any potential change in health care utilisation as a result of being insured is likely to be modified by a range of demographic, geographic and socioeconomic factors. Therefore, we explore how treatment effects vary for subgroups characterised by effect modifiers that have been selected using theory and the related literature, as well as using data-driven approaches, where the effect modifiers are identified from our analysis of causal forest outputs. We conduct our evaluation using data from the 2017 National Socioeconomic Survey (SUSENAS); a large-scale, repeated cross-sectional survey conducted at the household-member level that allows for exploring heterogeneity while also controlling for observed confounding.

We find that, similar to prior evaluations of JKN and other health insurance programmes in low- and middle-income countries, subsidised insurance increases the overall demand for outpatient and inpatient care. For those that do access care, the effects on the quantity of care consumed are particularly pronounced, which has not previously been reported in an evaluation of JKN. There is substantial evidence of heterogeneity in treatment effects. A surprising finding is the negative effect on the decision to seek outpatient care for the population subgroup that comprises the 20% of individuals that are expected to benefit the least from health insurance. Without an assessment of heterogeneity, such adverse effects would be missed. We identify important effect modifiers through our data-adaptive methods, including marital and employment status, as well as proxies for socioeconomic status, such

as technology usage, that are not pre-specified according to theory and the related literature.

Our paper makes several contributions to the empirical policy evaluation literature. We add to the limited evidence base on the impacts of subsidised JKN, with the additional novel focus on heterogeneity. By decomposing our outcome measures of health care utilisation, we assess how health insurance impacts participation and consumption decisions, which is often overlooked. Our study is also relevant to the wider literature on the effectiveness of universal health coverage schemes in low- and middle-income countries, where treatment effect heterogeneity is yet to be fully explored. Methodologically, we contribute to the applied policy evaluation toolkit by demonstrating the potential of combining predictive and causal machine learning in treatment effect estimation. Specifically, we highlight the ability to fit flexible models for health care utilisation and discover potentially new effect modifiers according to the data. We demonstrate that combining theoretically motivated analyses with data-driven methods can improve the evaluation process and generate value for future, evidence-based decision-making.

2.2 The evaluation problem

2.2.1 Background

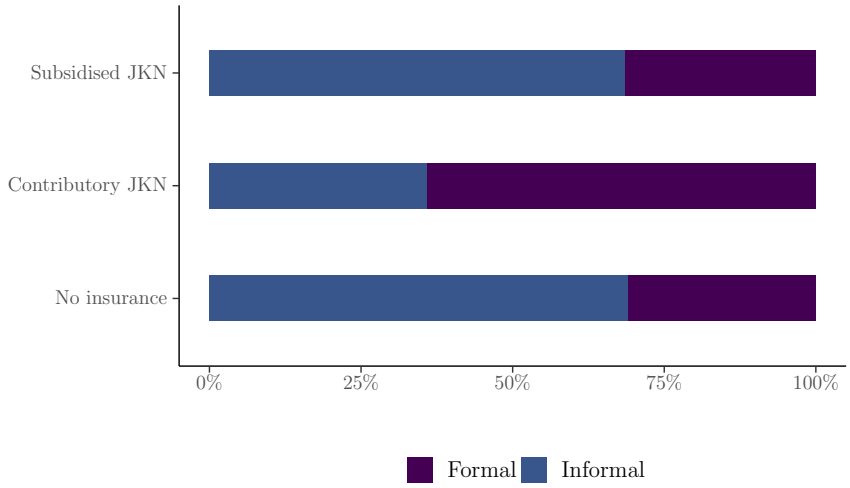
Indonesia is the world's fourth most populous country with 270 million inhabitants in 2019. Amid a backdrop of strong economic growth, the country has recently gained "upper middle-income" status (Mahendradhata et al., 2017). Despite this, Indonesia faces significant health challenges compared to its regional peers. For example, the maternal mortality ratio in 2017 was 177 per 100,000 live births compared to 29 and 27 in Malaysia and Thailand respectively. Life expectancy, despite recent improvements, is also lagging behind its neighbours (71.5 years in 2018, compared to 76 and 77 in Malaysia and Thailand respectively). Geographic and socioeconomic inequalities in access to affordable and quality health care have contributed to Indonesia's health problems (Wiseman et al., 2018). In recent decades, various

health policy reforms, in the form of health care financing schemes, have been introduced in an attempt to improve health-related outcomes across the population. Prior to the implementation of JKN in 2014, a number of social insurance schemes were available for targeted population subgroups: *Askes* for formal public sector workers, *Jamsostek* for formal private sector workers, and *Jamkesmas* for the poor and near-poor. However, workers in the informal sector, which account for 60% of Indonesia's workforce and a large proportion of the country's poorest population, were not eligible for these schemes. One of the motivations behind the development of a nationwide programme was to address this coverage gap by providing the so-called "missing-middle" with a chance to enrol. The existing schemes were unified into a national, single-payer programme, the JKN, as part of larger social welfare policy reforms (Aspinall, 2014; Mahendradhata et al., 2017).

Enrollees of the JKN programme are stratified into two groups: non-contribution beneficiaries (Penerima Bantuan Iural, or PBI) comprising the poor, near-poor and disabled; and contribution beneficiaries (non-PBI) comprising salaried workers, non-salaried workers, non-employees, and their respective families. PBI recipients (that is, the "subsidised" group) are provided with the basic benefits package in exchange for fully government-subsidised JKN premiums. Non-PBI recipients (that is, the "contributory" group) pay insurance premiums either themselves or through their employers, and receive a benefits package according to their membership class. All enrollees have access to at least a comprehensive basic benefits package that covers outpatient and inpatient treatment at all medical facilities, from primary up to secondary care. Auxiliary benefits are either partially or fully covered depending on the membership class.

The objective of JKN was to achieve universal health coverage by 2019 (Mboi, 2015). Although this target has yet to be met, JKN is still the world's largest social health insurance scheme. According to 2017 data, subsidised enrollees were the largest coverage group (comprising 60% of the insured population), followed by formal sector employees and the wealthy. Enrolment among the lower- and middle-income groups has been slower and harder to enforce given the large proportion of informal sector workers in these wealth quintiles (Agustina et al., 2019). It was reported that in

Figure 2.1: JKN coverage and employment sector in 2017



Note: Employment sector is defined as per Badan Pusat Statistik (BPS) Indonesia. Informal sector includes self-employment, employer assisted by temporary workers, freelance employees and unpaid family workers. Formal sector includes employers assisted by permanent workers and employees. Source: SUSENAS 2017.

2016, only 15 million households employed in the informal sector had voluntarily enrolled into JKN (Dartanto et al., 2016). Using data from SUSENAS 2017, Figure 2.1 shows the proportion of formal and informal sector workers that were enrolled into subsidised and contributory JKN, or had no insurance. Among the uninsured sample, the vast majority were employed in the informal sector. Enrolment has been challenging for this population given the voluntary, overly bureaucratic registration process, as compared to the subsidised group and formal sector employees, for whom registration is involuntary (Dartanto et al., 2020b). This is a problem not limited to Indonesia, but in many low- and middle-income countries with social health insurance programmes (Vilcu et al., 2016).

2.2.2 Related literature

There is a large volume of literature analysing the impact of health financing schemes in low- and middle-income countries on key universal health coverage objectives: improved financial protection, health care utilisation, and health outcomes. Previous country-level impact evaluations have generally found positive effects of health in-

surance on various utilisation measures, often as a result of improved affordability and access to care (Escobar et al., 2011; Giedion et al., 2013). Some case-study examples include those from Thailand, Ethiopia, Lao, Colombia, Ghana and China (Alkenbrack and Lindelow, 2015; Limwattananon et al., 2015; Mebratie et al., 2019; Miller et al., 2013; van der Wielen et al., 2018). Systematic reviews of empirical studies by Acharya et al. (2013) and Erlangga et al. (2019b) show more mixed results, with the strength of effects on curative and preventive care varying according to the type of insurance scheme and the target population group. However, the majority of included studies report increases in health care access and utilisation, with only a few finding zero or negative effects. Moreno-Serra and Smith (2015) highlight certain methodological constraints associated with evaluating health insurance policies that can often lead to mixed, and often counter-intuitive, findings. For example, confounding (due to self-selection into insurance) is a particularly concerning limitation that can bias treatment effect estimates if it is not properly addressed within the study design. A simple solution is to randomise the policy assignment, although this is not always feasible, especially for large-scale policies such as the JKN. In non-randomised settings, such as in our case, we try to eliminate confounding by controlling for a large, diverse set of observable characteristics that explain selection into insurance. The availability of rich survey data supports this approach, however we acknowledge the limitations of relying on observed covariates to minimise confounding, and cannot guarantee that other unobserved factors contributing to self-selection may not be controlled for. Alternative study designs that leverage upon panel data or instrumental variables could be used to tackle unobserved confounding, however incorporating these methods within this study were not possible due to data limitations.

Given the relatively recent implementation of JKN, the evidence base exploring its association with universal health coverage objectives is fairly limited. Several studies have evaluated previous Indonesian health reforms (for example, *Askeskin* and the *Health Card* subsidy programme, finding that targeted programmes for the poor had mixed results (Johar, 2009; Sparrow et al., 2013). We highlight some recent studies that have used panel data for a before-after evaluation of the early effects of JKN on health care expenditure and utilisation outcomes. Some additionally perform

subgroup analyses to explore treatment effect heterogeneity. Nugraheni et al. (2020) discover that JKN reduced out-of-pocket delivery costs for expecting mothers, as well as the risk of incurring catastrophic delivery expenditures. Erlangga et al. (2019a) find a positive impact on inpatient and outpatient demand for the contributory group, and a smaller positive effect on just inpatient demand for the subsidised group. They perform traditional subgroup analyses by stratification, finding that policy effects are stronger among wealthier respondents, and in areas with better health infrastructure. Kreif et al. (2022) also show positive average impacts of contributory JKN on maternal health care utilisation and infant mortality, and no significant effects of the subsidised programme. To estimate CATEs, they use causal forests, finding that poorer, lower educated, and rural-based mothers in the contributory group have the greatest increase in health care utilisation. They do not find any significant heterogeneity for the subsidised group. Using a cross-sectional study design, Anindya et al. (2020) find that JKN-insured women are more likely to utilise maternal health care, and that effect sizes are greater among the poor and those living in rural areas.

Exploring treatment effect heterogeneity is becoming an increasingly important component of empirical research, and recent methodological developments have further supported this trend (Carvalho et al., 2019). In particular, there are a growing number of applied economic papers that use causal forests to estimate average and heterogeneous treatment effects using both randomised and non-randomised study designs (Bertrand et al., 2017; Brand et al., 2021; Daoud and Johansson, 2019; Davis and Heller, 2017; Kreif et al., 2022; O'Neill and Weeks, 2018; Strittmatter, 2019). Our paper is one of the first to use causal forests to evaluate a large-scale health policy using observational data.

2.3 Data

SUSENAS is conducted by the central statistics agency, Badan Pusat Statistik, and is the only nationally representative, socioeconomic survey in Indonesia (Johar et al., 2019). It collects annual data on a sample of households, and its members, through

two surveys: a core survey on socioeconomic status, and a consumption survey on household expenditure and income. Each survey samples 300,000 households on average (corresponding to approximately 1.1 million household members), across 34 provinces and 514 districts. A structured two-stage sampling design is used to select households within sample census blocks, which ensures households are representative at the district-level. Frequency weights are additionally provided to reflect the total national population (Johar et al., 2019).

We use cross-sectional data from SUSENAS 2017 to construct a dataset of household members, identified by their respective households. We extract the following measures of health care utilisation: a count variable for the number of outpatient visits in the past one month and a count variable for the total length of inpatient stay (in days) in the past one year. The greater lag period for inpatient care reflects its rarity compared to outpatient care (Bhandari and Wagner, 2006). We consider all outpatient and inpatient treatments that took place at a public or private medical facility.[1] We construct two outcomes for each variable: a binary indicator for whether the respondent had any treatment in the specified time period, which we refer to as the "participation" component; and a count variable for the total intensity of treatment for only those respondents who reported having treatment in the same time period, or the "consumption" component. These separate components are later used to construct the two-part models.

We restrict our sample to respondents who reported having either "no insurance" or "subsidised health insurance" at the time of survey. Subsidised health insurance is defined as being enrolled into the *JKN Penerima Bantuan Iuran (PBI)* or the former *Jamkesda* schemes. Beneficiaries of subsidised insurance who reported having additional health insurance, either through employee contributions (JKN non-PBI) or private plans, are excluded from the analysis as this was not formally allowed.[2]

We extract a large, diverse set of variables to construct the covariate vector X of confounders that affect selection into treatment, while also predicting health care

[1] Medical facility refers to public hospitals, private hospitals, physician/midwife clinics, physician polyclinics, community health centers (puskesmas/pustu), and community based health efforts otherwise known as UKBM (poskesdes, polindes, posyandu, balai pengobatan). Traditional or alternative medical treatments are excluded from the utilisation measures.

[2] Respondents reporting multiple health insurance plans constituted 2.5% of the overall sample.

utilisation. Our variable selection process is guided by theory and related studies that have performed similar evaluations in Indonesia. Following the work of Vidyattama et al. (2014) and Erlangga et al. (2019a), we capture individuals' health care needs by controlling for age, sex, marital status and various socioeconomic measures, including education level, employment status, household expenditure per capita, housing characteristics, asset ownership, and technology usage. Geographical inequalities are controlled for using an urban versus rural distinction and seven regional dummies, given the substantial variation in health care utilisation and economic progress across these administrative boundaries (Johar et al., 2018; Mulyanto et al., 2019c). Lastly, since health care demand relies upon adequate supply-side infrastructure, we control for accessibility to local health care facilities using village-level census data in Potensi Desa data, PODES 2018 (conducted in 2017), which we merge with SUSENAS at the district-level. We gather information on accessibility to the four main types of health care providers: hospitals (private and public), primary care (doctors' clinics, polyclinics, and mobile clinics), community and auxiliary centres (with and without inpatient care), and maternity (hospitals, homes, midwives and mobile clinics). Ease of access is measured using a 4-point scale: very easy, easy, difficult and very difficult. For our supply-side dummies, we collapse the categorical variable into a binary indicator for each provider type, where accessibility to the health care provider is either easy or very easy, and difficult or very difficult is the baseline.

Our final data set comprises 912,812 household members across 297,276 households, of which 475,930 members are in the subsidised "treated" group, and 436,882 are in the uninsured "control" group. Our covariate vector includes a total of 89 variables (see Appendix 2.A for a comprehensive list). We include sample weights to generalise our results to the true, target survey population.

2.4 Methods

2.4.1 The causal framework

We construct an observational dataset (X_i, Y_i, D_i) of household members $i, \ldots, N$, where X_i is the vector of confounders and potential effect modifiers, D_i is the binary treatment (which equals 1 if i is enrolled into subsidised JKN, and 0 if i is uninsured), and Y_i are the continuous outcomes that measure the utilisation of inpatient and outpatient health care. Following the potential outcomes framework of causal inference (Imbens and Rubin, 2015; Rubin, 1974), $Y_i(d)$ denotes the potential outcome that would be observed if household member i was assigned to treatment d. Individual level treatment effects are defined as the difference in the potential outcomes: $\tau_i = Y_i(1) - Y_i(0)$. However, given the fundamental problem of causal inference, τ_i cannot be observed. We can instead take expectations of the difference in the potential outcomes across the population to produce the ATE: $\tau = E[Y_i(1) - Y_i(0)]$.

Since we are interested in exploring the variation in treatment effects across the population, we also define the CATE function, which evaluates the ATE for individuals with the same covariate profile $X_i = x$:

$$\tau(x) = E[Y_i(1) - Y_i(0)|X_i = x], \tag{2.1}$$

This function captures heterogeneity in treatment effects through effect modifiers included in X. The ATE can also be defined as the expectation of the CATE function over a population represented by the distribution of X, $\tau = E_X[\tau(x)]$.

The CATE is our target causal parameter, which, in order be identified using the observed data, requires making the following assumptions on the data generating process. First, the unconfoundedness assumption (also known as selection on observables): $\{Y_i(0), Y_i(1)\} \perp D_i|X_i$, which requires the potential outcomes to be independent of treatment status, conditional on the observed covariates. Second, the overlap assumption: $0 < e(X_i) \equiv P[D_i = 1|X_i = x] < 1$, which requires the probability of being enrolled into subsidised JKN (that is, the propensity score $e(x)$,

which we also refer to as the treatment model) to be bounded away from zero and one. If these two assumptions (jointly referred to as strong ignorability) are satisfied, the conditional expectation of the potential outcomes equals the conditional expectation of the observed outcome. That is, $E[Y_i(1)|X_i = x] = E[Y_i|X_i = x, D_i = 1]$ and $E[Y_i(0)|X_i = x] = E[Y_i|X_i = x, D_i = 0]$. The CATE can therefore be identified as a function of the observed outcomes:

$$\begin{aligned}\tau(x) &= E[Y(1)|X_i = x] - E[Y(0)|X_i = x] \\ &= \mu_1(x) - \mu_0(x),\end{aligned} \tag{2.2}$$

where $\mu_d(x)$ is the counterfactual response surface. We denote $\mu_D(X_i)$ as the conditional expectation function for the observed outcome under the treatment actually received.

Estimating CATEs relies on generating good predictions of $\mu_1(x)$ and $\mu_0(x)$, which makes machine learning ideally suited for this task (Jacob, 2021). Recently proposed machine learning methods for CATE estimation fall into two categories. The first category consists of methods that employ "off-the-shelf" machine learning algorithms (for example, random forests, generalised boosting models, neural networks and the lasso) to estimate the separate components in (2.2). These methods are not specifically designed to estimate CATEs directly and depend on multiple, model agnostic regression tasks (Künzel et al., 2019a; Nie and Wager, 2021). The second category consists of methods that attempt to directly estimate the CATE function using machine learning algorithms that have been adapted for causal tasks (for example, causal forests (Athey and Wager, 2019), causal boosting (Powers et al., 2018) and Bayesian causal forests (Hahn et al., 2020)). We focus on the latter category in this study, in particular causal forests that are designed to generate point and interval estimates of the target CATEs.

To estimate CATEs using causal forests additionally requires estimates of the conditional expectation of the observed outcome marginalised over the treatment: $m(x) = E[Y_i|X_i = x]$, which we refer to as the outcome model. The parameters $e(x)$ and $m(x)$ are jointly known as the nuisance parameters since they are not directly of interest, but are required to estimate the target causal parameter. Since it is not

necessary to understand their underlying structural form, the nuisance models are essentially prediction tasks that can be estimated using machine learning, which allows for a flexible model specification compared to traditional parametric approaches.

2.4.2 Estimating nuisance parameters using the super learner

We estimate $e(x)$ and $m(x)$ using the super learner, which leverages upon various machine learning concepts to generate predictions: "ensembling", which combines a number of heterogeneous algorithms (or "base learners") into a single, optimally-weighted algorithm (a "meta-learner" or "ensemble learner") that aims to improve model accuracy and goodness-of-fit; and K-fold cross-validation, which uses sample splitting to evaluate base learner performance against a user-defined loss function (for example, mean squared error (MSE)) on held-out, validation data.[3] It is a generalisation of the "stacking" algorithm that, under large enough samples, can perform at least as well as the best base learner in the ensemble (Breiman, 1996b; van Der Laan and Dudoit, 2003; van der Laan et al., 2007). There are two versions of the super learner: the "continuous" version that finds the optimal linear weighted ensemble of base learners, and the "discrete" version that selects the best performing learner among the base and ensemble learners. We fit continuous super learners to estimate our nuisance models.

For the treatment model, we select base learners designed for binary outcomes. For the outcome model, we incorporate the two-part model framework to address the large mass points at zero in the distributions of our outcomes. The decision to access care stratifies the population into users and non-users, and for those that decide to access care, the level of care consumed is an entirely different process, depending to a large extent on supply-side factors (Pohlmeier and Ulrich, 1995). Two-part models are popular for modelling health care utilisation since they take into account these separate processes, typically using logit or probit models for the binary "participation" model, and poisson or negative binomial models for the

[3] In K-fold cross-validation, the sample is divided into K folds. Predictions are made on each validation fold k after training the base learners on the remaining $K-k$ folds. Model performance is evaluated by calculating the average loss across all validation folds using predicted and observed outcomes.

truncated-at-zero count "consumption" model (Mihaylova et al., 2011).[4] We propose fitting two super learners to predict the separate components of the two-part model:

$$m(x) = E[Y_i|X_i] = \underbrace{P[Y_i > 0|X_i = x]}_{m_1(x)} \cdot \underbrace{E[Y_i|Y_i > 0, X_i = x]}_{m_2(x)}, \tag{2.3}$$

where $m_1(x)$ and $m_2(x)$ are the respective participation and consumption components.

Table 2.1: Candidate algorithms included in the super learner libraries

Algorithm	**Description**	**Binary**	**Count**
glm-g	Generalised linear model (gaussian)		x
glm-b	Generalised linear model (binomial)	x	
glm-p	Generalised linear model (poisson)		x
glm-nb	Generalised linear model (negative binomial)		x
lasso-g	LASSO (gaussian)		x
lasso-b	LASSO (binomial)	x	
lasso-p	LASSO (poisson)		x
lasso-nb	LASSO (negative binomial)		x
rf	Random forest	x	x
gbm	Generalised boosting model	x	x
nn	Neural network	x	x

Note: Table shows the algorithms included in the respective super learner libraries for binary and count tasks. Binary tasks include the estimation of the treatment model and the participation components of the outcome models. Count tasks include the estimation of the consumption components of the outcome models.

Here, we summarise the super learner procedure for predicting the consumption component of the outcome model $m_2(x)$ (a similar procedure applies for all other prediction tasks, albeit the loss function is adjusted as per the outcome type). We select a diverse set of base learners for each super learner task, including both parametric and non-parametric models (see Table 2.1 for a full list). We include linear models, although by construction, the ensemble learner can give these a small

[4]There is some literature suggesting that the decomposition of the causal ATE into extensive and intensive margins is difficult in two-part models. The decomposition can introduce a form of selection bias since the intensive margin effect is conditioned on the selecting or participation decision, which is a post-treatment variable that may be affected by the treatment itself (Lee, 2017; Staub, 2014). This echoes the earlier debate on the use of sample selection versus two-part models for the problem of limited dependent variables - see Jones (2000) for a summary and overview of the debate.

weight if they do not fit the data well. To compare the performance of the super learner against conventional hurdle models for estimating the two-part model, we include a logistic regression for the participation component, and truncated-at-zero Poisson and negative binomial models for the consumption component. We include the full covariate vector X to adjust for confounding, and select log loss and MSE as the respective loss functions for binary and continuous outcomes. We perform K-fold cross-validation (we set $K = 5$) and generate predictions $\hat{Y}_l$ for each learner $l = 1, \ldots, L$, as well as a measure of model performance in terms of the expected loss. The cross-validated predictions $\hat{Y}_l$ and the observed outcomes Y are used as inputs into a ensemble learner that finds the optimal linear combination of L base learners using stacking. The contribution of each base learner is determined through an ordinary least squares (OLS) regression of Y on $\hat{Y}_l$. The intercept is suppressed and the coefficients that represent the respective weights of the base learners are constrained to be non-negative:

$$E[Y|\hat{Y}_1, \ldots, \hat{Y}_L] = \alpha_1\hat{Y}_1 + \alpha_2\hat{Y}_2 + \ldots + \alpha_L\hat{Y}_L, \tag{2.4}$$

where $\alpha_l \geq 0$. A separate K-fold cross-validation is done on the ensemble learner to evaluate its performance against the individual base learners, and the cross-validated ensemble learner is used to generate nuisance predictions.

See Figure 2.C.1 for a super learner workflow.

2.4.3 Estimating conditional average treatment effects using causal forests

The causal forest approach to heterogeneous treatment effect estimation is founded on semi-parametric statistical theory, requiring a partially linear model for the potential outcome of interest:

$$Y_i = \mu_0(X_i) + D_i\tau(X_i) + \varepsilon_i, \tag{2.5}$$

where the shape of $\mu_0(X_i)$ is unspecified. Given that $E[\varepsilon_i|X_i, D_i] = 0$, where $\varepsilon_i = Y_i - \{\mu_0(X_i) + d\tau(X_i)\}$, (2.5) can be rearranged by rewriting $m(x)$ as $E[Y_i|X_i = x] = \mu_0(X_i) + e(X_i)\tau(X_i)$, and residualising it as follows:

$$Y_i - m(X_i) = (D_i - e(X_i))\tau(X_i) + \varepsilon_i. \tag{2.6}$$

The residuals from the outcome model (the Y-residuals) are regressed onto the residuals from the treatment model (the D-residuals) to debias the estimator (Chernozhukov et al., 2018a; Robinson, 1988). For a given value of the covariate vector $X_i = x$, the CATE estimator $\hat{\tau}(x)$ can be constructed as a weighted linear regression of the Y-residuals on the D-residuals:

$$\hat{\tau}(x) = \frac{\sum_{i=1}^{N} w_i(x)(D_i - \hat{e}(X_i))(Y_i - \hat{m}(X_i))}{\sum_{i=1}^{N} w_i(x)(D_i - \hat{e}(X_i))^2}, \tag{2.7}$$

where $\hat{m}(X_i)$ and $\hat{e}(X_i)$ are the estimated nuisance parameters and $0 \leq w_i(x) \leq 1$ are weights for each observation i based on how frequently it is used to estimate $\hat{\tau}(X_i)$. Observations with a similar covariate profile X_i to x receive a larger weighting in the estimator. The weights w_i are obtained using a neighbourhood weighting function that is solved using the causal forest algorithm.

Causal forests are an extension of the popular random forests algorithm that builds an ensemble of regression trees in order to generate outcome predictions for each observation through recursive partitioning (that is, finding neighbourhoods in the covariate space). The algorithm uses a splitting criterion that partitions the data at the tree root by selecting variables and their respective cut-off thresholds that achieve the greatest reduction in prediction error. This splitting process is repeated at each node until a tree-like structure is formed. The predicted outcome for a new observation is generated by dropping it down the tree until it lands in a terminal node (or "leaf"), and taking the average outcome of all observations that fell into this same leaf during training. Forests repeat this process across an ensemble of trees, constructed using bootstrapped samples of the data, in order to improve prediction accuracy and to limit noise stemming from individual trees. The final forest prediction for each observation is the average prediction across the ensemble.

Causal forests are similar to random forests, except that they average neighbourhood weights across trees, instead of predicted outcomes, and use the resulting weights to solve equation (2.7). There are some important additional modifications. First, the splitting criterion searches for a partition that maximises heterogeneity in treatment effects, rather than minimising prediction error, so that CATEs are similar within leaves and different across leaves. Second, they prevent overfitting by constructing "honest" trees, meaning that each training unit i is either used to construct the tree structure or to estimate the within-leaf treatment effect, but not both. The causal forest predicts CATEs for each observation according to their individual covariate profile, and pointwise confidence intervals are derived using a consistent estimator for the asymptotic variance of $\tau(x)$.

See Appendix 2.B for further details on the causal forest algorithm.

2.4.4 Summarising and interpreting treatment effect heterogeneity

Although the causal forest generates CATE estimates with confidence intervals, using these predictions to draw conclusions about the distribution of $\tau(x)$ is not recommended for various reasons. First, the confidence intervals may, in general, be quite wide due to the complexity of the CATE function. Second, their construction relies on tuning for "undersmoothing" – meaning that the bias of the estimator shrinks faster than its variance – which could result in biased confidence intervals, if tuning is not performed correctly. Cui et al. (2020) highlight some of these coverage issues in their simulation study. Given the challenges associated with interpreting CATEs, the wider literature suggests making inferences on low-dimensional summaries of the CATEs, rather than the predictions themselves. This requires constructing almost unbiased (but noisy) proxies for the predicted CATEs, known as doubly robust scores $\hat{\Gamma}_i$, whose expectation is the augmented inverse probability of treatment weighted (AIPTW) estimator for the ATE (Robins et al., 1994):

$$\hat{\tau} = \frac{1}{N}\sum_{i=1}^{N}\hat{\Gamma}_i(X_i), \quad \hat{\Gamma}_i = \hat{\mu}_1(X_i) - \hat{\mu}_0(X_i) + \frac{D_i - \hat{e}(X_i)}{\hat{e}(X_i)(1-\hat{e}(X_i))}(Y_i - \hat{\mu}_D(X_i)). \quad (2.8)$$

The AIPTW estimator has several desirable properties, including double-robustness, asymptotic normality and efficiency. Also, the adjustment term that weights the Y- and D-residuals by the inverse of the estimated propensity score removes the effects of any regularisation bias that may contaminate the ATE estimate, which is a fundamental concept in the recent double machine learning literature (Chernozhukov et al., 2018a).

Another important double machine learning concept is the role of sample splitting in reducing bias from overfitting, if the nuisance and causal parameters are estimated on the same data. Kennedy (2020) proposes estimating Γ_i in (2.8) using three subsamples, where the first subsample is used to train the outcome model, the second to train the treatment model, and the third to construct the doubly robust scores. To regain full sample efficiency, Chernozhukov et al. (2018a) propose "K-fold cross-fitting", where the role of each subsample is swapped K times so that a doubly robust score is constructed out-of-sample for each $i \in n$. All subsequent analyses (for example, constructing ATEs for population subgroups or finding the best linear predictors of the CATE function) can then be performed on the full data set using the pooled, cross-fitted scores.[5]

Another important role of the AIPTW estimator in (2.8) is to estimate group average treatment effects (GATEs) – defined as ATEs for pre-specified strata of the population – as a means of identifying and exploring treatment effect heterogeneity. Following Chernozhukov et al. (2018b), one form of GATEs, known as "sorted GATES", can be constructed by stratifying $\hat{\tau}(x)$ into quantiles G_q (a popular approach is where $q = 5$), and estimating $\hat{\tau}_q$ for each quantile. The authors suggest that heterogeneity can be identified by testing whether the differences between $\hat{\tau}_q$ for $q = 1, \ldots, Q$ are statistically significant. This requires a difference-in-means estimator that corrects for multiple hypothesis testing. By classifying individuals into quantiles, a separate test for heterogeneity can be performed that compares differences in covariate means across quantiles. This is a type of "classification analysis" that describes the characteristics of populations grouped according to their

[5]Technically, Semenova and Chernozhukov (2021) state that only two subsamples are required to perform subgroup analyses under mild conditions (one for the nuisance models and another to construct the scores), and that three subsamples are only essential when directly evaluating CATEs.

estimated CATEs, and compares them to the average characteristics in the entire sample population. From this analysis, data-driven effect modifiers can be identified by quantifying and ranking the differences in standardised mean characteristics across quantiles using all covariates in X_i. An alternative way of identifying data-driven effect modifiers is to use causal forests' in-built variable importance function that ranks covariates by the frequency with which they were split on when building the forest. However, this approach tends to favour continuous variables given their larger number of splitting points. Theory-driven effect modifiers, in comparison to those identified from data, are selected a priori based on the relevant theory and literature.

The full set of effect modifiers, identified via both theory and data, can be used in a type of univariable subgroup analysis that constructs GATEs using the AIPTW estimator on a restricted sample (Knaus et al., 2021). This analysis demonstrates how treatment effects change when conditioning on a single covariate, however the interpretation of the estimates should be of general trends rather than magnitude of effects, given issues of multiple hypothesis testing and collinearity.

An alternative, mutivariable subgroup analysis proposed by Semenova and Chernozhukov (2021) explores linear summaries of the CATE by estimating projections of the true $\tau(x)$ on simpler hypothesis spaces. The best linear projection (BLP) of $\tau(x)$ is:

$$\{\beta_0^*, \beta^*\} = \text{argmin}\, E\Big[(\tau(X_i) - \beta_0 - A_i\beta)^2\Big], \tag{2.9}$$

where $A_i \in X_i$ is a set of covariates that can be selected either pre-analysis based on existing hypotheses about how $\tau(x)$ varies with certain covariates (for example, age and region), or post-analysis based on the variable importance results from the classification analysis.[6] The BLP parameters β_0^* and β^* in (2.9) are estimated via an OLS regression of $\hat{\Gamma}_i$ on A_i, with standard errors derived in the usual way (Semenova and Chernozhukov, 2021). In theory, the causal forest CATE estimates could replace $\hat{\Gamma}_i$ in the regression but simulation studies have found that estimators based on $\hat{\Gamma}_i$ are more stable given their double robustness property. Since the aim of the projection is to estimate the predictors of the CATE function, and not

[6] A_i could also be the full covariate vector X_i

the CATEs themselves, doubly robust scores are the superior outcome variable for the regression. The estimated regression coefficients have the same *ceteris paribus* interpretation as those from an OLS model, which in comparison to the univariable analysis described in the previous section, can provide even more nuanced insights into the main predictors of treatment effect heterogeneity. However, it is important to state that the projection assumes a linear association between $\hat{\Gamma}_i$ and A_i (the true $\tau(x)$ may not be linear in A_i), so the coefficients should not be interpreted as partial effects.

2.4.5 Our implementation

We apply the described methods to our health insurance policy evaluation problem. Our approach is largely based on methodological contributions by several leaders in the field (Athey et al., 2019b; Athey and Wager, 2019; Chernozhukov et al., 2018a,b; Kennedy, 2020; Knaus et al., 2021; Semenova and Chernozhukov, 2021). We perform all analyses on our two outcome measures – inpatient and outpatient demand – using our three outcome models of interest: the overall two-part model, the participation component and the consumption component.

We explain our implementation in the following steps:

1. We construct the dataset and identify potential theory-driven effect modifiers among the covariate vector X.[7] We refer to previous subgroup analyses from the existing literature to inform our selection of effect modifiers, which include demographic, socioeconomic, geographic and supply-side determinants of health care access, including age, household wealth (measured by monthly household consumption expenditure per capita), the urban versus rural distinction and the local availability of health care facilities (Erlangga et al., 2019a).

2. We perform 3-fold cross-fitting by randomly dividing the data into three equal subsamples, $\mathcal{S} = \{s_1, s_2, s_3\}$, where s_1 is used to train the nuisance models,

[7] For the CATE models, we consider all variables to be effect modifiers throughout.

s_2 to train the causal forest, and s_3 to predict CATEs and construct doubly robust scores. We use three subsamples (as opposed to two) since we have an additional CATE estimation step (using causal forests).

3. We fit super learners (using the `h2o`[8] package in `R`) on s_1 to train the nuisance models – $m(X_i)$ (constructed using the product of $m_1(X_i)$ and $m_2(X_i)$) and $e(X_i)$ – and make predictions on new data, s_2 and s_3. We check that the overlap assumption is satisfied by exploring the distribution of $\hat{e}(X_i)$.

4. We train an honest causal forest (using the `grf`[9] package in `R`) on s_2 using observed and predicted inputs. The forest is trained according to a range of parameters that can be tuned via cross-validation. We select default parameters (that is, 2000 trees for training and 200 trees for tuning) and include sample weights. See Table 2.C.1 for a list of training and tuning parameters used in our forests.

5. We predict $\hat{\tau}(X_i)$ on the testing subsample s_3 by evaluating the CATE estimator for each household member according to their individual covariate profile $X_i = x$.

6. For cross-fitting, we repeat steps 3-5 twice and swap the roles of the subsamples. In the second iteration, we use s_2 for nuisance training, s_3 for causal forest training, and s_1 for predicting $\hat{\tau}(X_i)$. In the third and final iteration, we perform the respective tasks on the remaining ordering of the subsamples $\{s_3, s_1, s_2\}$. This procedure generates out-of-sample predictions – $\hat{e}(X_i)$, $\hat{m}(X_i)$ and $\hat{\tau}(X_i)$ – for each subsample that, when pooled, reconstructs the full data set. See Figure 2.2 for our sample splitting workflow.

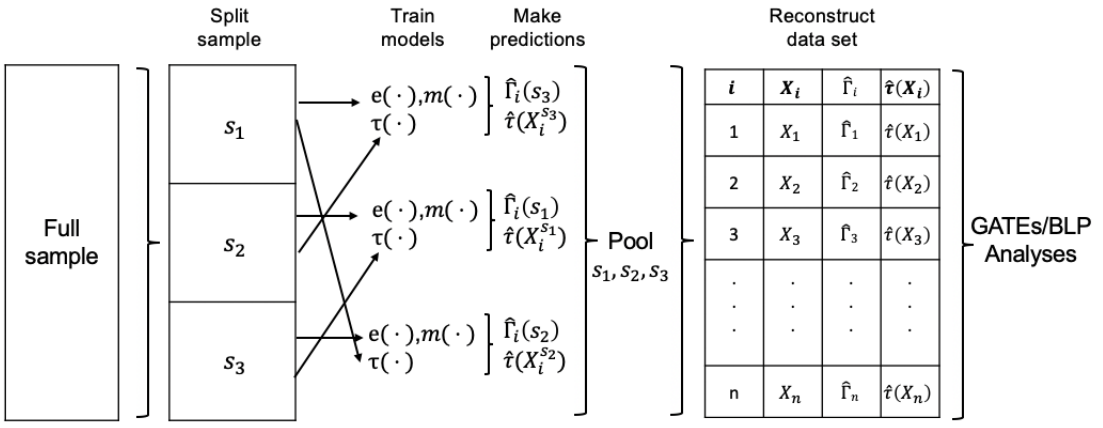
7. We construct doubly robust scores $\hat{\Gamma}_i^{cf}$ using the observed and predicted inputs from the pooled data:

$$\hat{\Gamma}_i^{cf} = \hat{\tau}(X_i) + \frac{D_i - \hat{e}(X_i)}{\hat{e}(X_i)(1 - \hat{e}(X_i))}(Y_i - \hat{m}(X_i) - (D_i - \hat{e}(X_i)\hat{\tau}(X_i)). \quad (2.10)$$

Note that $\hat{\Gamma}_i^{cf}$ in (2.10) is a slightly modified but theoretically identical version

[8]https://github.com/h2oai
[9]https://github.com/grf-labs/grf

Figure 2.2: Sample splitting workflow.



of $\hat{\Gamma}_i$ in (2.8) that incorporates the causal forest CATE estimates $\hat{\tau}(X_i)$ and the conditional mean outcome $\hat{m}(X_i)$ (Athey and Wager, 2021).[10]

8. We divide the pooled data into quintiles (Q1-Q5) by ranking observations according to $\hat{\tau}(X_i)$, and we perform the following subgroup analyses:

   - We construct the ATE and sorted GATEs (for each quintile) using the AIPTW estimator, as in (2.8), but with $\hat{\Gamma}_i^{cf}$.
   - We test whether the differences in sorted GATEs between the bottom quintile (Q1) and higher quintiles (Q2-Q5) are statistically significant using a difference-in-means estimator.[11]
   - We summarise how the joint distribution of covariates in X varies across quintiles by comparing quintile-specific covariate means to overall covariate means using heatmaps. We standardise the covariates in order to compare them on a similar scale. We select our "data-driven" effect modifiers by identifying the top five covariates in X with the largest difference in standardised means between Q1 and Q5. We compare the selection of these effect modifiers to those that are generated from the causal forest's in-built variable importance ranking.
   - We perform univariable subgroup analyses for theory- and data-driven

[10] The construction of $\hat{\Gamma}_i^{cf}$ in (2.10) differs to $\hat{\Gamma}_i$ in (2.8) because $\tau(x)$ is estimated using a causal forest, which fits a partially linear model (as described in section 2.4.3) that relies on $\tau(x)$ being expressed in terms of the conditional mean outcome $m(x)$.

[11] We use the Romano-Wolf correction for multiple hypothesis testing (Clarke et al., 2020).

effect modifiers by constructing GATEs (using the AIPTW estimator) for populations grouped according to the values of these covariates. We dichotomise all continuous variables to aid interpretation.

- We perform multivariable subgroup analyses by finding the BLP of the CATE function on covariates. We regress $\hat{\Gamma}_i^{cf}$ on $A_i \in X_i$, where A_i comprises two sets of covariates: 1) the full covariate vector, and 2) a restricted covariate vector of theory- and data-driven effect modifiers. Again, we dichotomise all continuous variables to aid interpretation. We explore the coefficients on the BLP parameters to identify the covariates that are the main predictors of treatment effect heterogeneity.

2.5 Results

2.5.1 Descriptive statistics

Table 2.2 compares outcomes and selected observed characteristics of the treated and the control populations in terms of sample-weighted means, and unweighted (raw) and inverse propensity score-weighted standardised mean differences (SMD). Although the insured and uninsured populations are comparable for the majority of covariates in X, there exist some small differences between the two groups (that is, covariates with absolute SMD>0.1). The insured population is likely to be older (mean age of 31 compared to 27) and have easier access to primary health care facilities than the uninsured. Moreover, some smaller socioeconomic differences exist between the two groups. The treated tend to be more educated and literate compared to the controls. Their households are likely be larger, have less basic facilities and lower monthly expenditure per capita. Regionally, the insured are more likely than the controls to reside in Sumatera, Jakarta, Sulawesi and Maluku-Papua. In terms of health care utilisation, the treated are more likely to demand health care, as expected.

2.5.2 Selected prediction algorithms (for utilisation and health insurance status)

Table 2.C.2 displays the performance of the super learners in terms of the relative weighting of the individual base learners in each ensemble (using standardised coefficients), and the cross-validated loss of the base and ensemble learners. We find that the super learner is the optimal model in the estimation of the propensity score and the consumption components of the outcome regression. That is, the super learner minimises the cross-validated risk the most relative to the base learners. However, in the estimation of the participation components of the outcome regression, the generalised boosting model slightly outperforms the ensemble, although this results in its larger weighting in the ensemble. We also confirm that the super learner outperforms the separate components of the conventional two-part hurdle model – a logit or probit model for the participation component, and a poisson or negative binomial model for the consumption component.

Figure 2.C.2 plots the distribution of the predicted propensity scores, showing that the overlap assumption is satisfied since there are no observations in either the treated or control populations with extreme scores close to 0 or 1. Figure 2.C.3 displays balance statistics on the full covariate vector X before and after inverse probability of treatment weighting. Post-weighting, all covariate means are balanced (that is, absolute SMD < 0.1).

2.5.3 Group average treatment effects

Figure 2.3 presents point estimates and 95% confidence intervals for ATEs and sorted GATEs (based on quintiles of predicted CATEs) across all three models of interest. On average, subsidised health insurance increases the demand for both types of health care, as expected. However, the variation in sorted GATEs suggests that the impact on health care utilisation as a result of being enrolled into subsidised insurance differs considerably across individuals. On average, subsidised JKN increases the total length of inpatient stay in the previous year by 0.16 days, and the number

of outpatient visits in the previous month by 0.06 visits. These positive effects on inpatient demand are consistent across all three models of interest (that is, the two part, participation and consumption models), whereas this is not the case for the participation component of outpatient care where the sorted GATE is negative for the bottom quintile. In general, the participation effect of insurance is small, only increasing by 0.02 for inpatient care and 0.03 for outpatient care. However, the consumption effect is more substantial. Among respondents who decide to access care, insurance enrolment increases outpatient demand by 0.4 visits and inpatient demand by over 4 days.

Graphically, we also find evidence of treatment effect heterogeneity. Looking at the two-part and participation models, treatment effects on inpatient and outpatient care for the three least affected populations (Q1-Q3) are below the ATE, and only those in the highest quintiles experience larger, above-average effects. The ATE point estimates are therefore being positively skewed by the 20% of individuals who increase health care demand much more than the remaining population. The most worrying finding is the negative impact on the decision to seek outpatient care for the bottom quintile, which could potentially be explained by supply-side and accessibility issues, or greater demand for inpatient rather than outpatient care as a result of insurance.

Table 2.C.3 presents results from our heterogeneity test that explores whether the differences in estimated sorted GATEs between the lowest quintile (Q1) and the higher quintiles (Q2-Q5) are statistically significant (at the 5% level). For inpatient demand, the difference between Q1 and Q5 is consistently significant across the two-part, participation and consumption models. There are also significant differences between Q1 and Q4 for the two-part and participation models. For outpatient demand, we find even greater evidence of heterogeneity, in that the differences between Q1 and all remaining quintiles (Q2-Q5) are consistently significant across all three models of interest.

Figures 2.4 and 2.C.4 present heatmaps showing how the joint distribution of covariates varies for population subgroups ranked according to quintiles of predicted CATEs. Across all inpatient models, we identify substantial variation in covariate

means across subgroups, suggesting that health insurance impacts inpatient demand differently across the population based on individuals' characteristics. The variables with the largest variation across quintiles include those in our list of theory-driven effect modifiers, such as age, sex, accessibility to all four types of health care facilities (maternity, community, secondary and primary), and monthly household per capita expenditure. We find that policy effects overall are greater among respondents who are older (average age of 42 in Q5 compared to 23 in Q1), female, not single and out of work (either through retirement, unemployment or homemaking activities). These respondents are also better off in terms of having basic household facilities (for example, electricity) and higher monthly household expenditure per capita (1,000,000 IDR in Q5 compared to 800,000 IDR in Q1). Geographically, they are more likely to live in urban, rather than rural, areas. A large proportion (17%) of residents in the region of Maluku-Papua are in the least affected category (Q1), which suggests there is geographical heterogeneity in treatment effects. For outpatient demand, age is also the covariate with the largest variation across quintiles, with older respondents consuming more outpatient care (age 42 in Q5 compared to age 25 in Q1). However, aside from sex and monthly household expenditure per capita, we find that new, previously unspecified variables are the main effect modifiers. Treatment effects are greatest among respondents who are female, married, and socioeconomically poor in terms of lower monthly household expenditure per capita (compared to Q1), not having enough food to eat in the previous year, and lower technology ownership and usage. Geographically, there is some small variation across quintiles, in particular for the specific region of Java, suggesting that location is a potential driver of treatment effect heterogeneity for outpatient care. These results are generally supported by those from the participation and consumption models, however we find that the primary effect modifiers differ from the two-part model, suggesting that impact heterogeneity on the decision to access to care and the quantity of care consumed is driven by different types of characteristics. Age is a leading effect modifier in the participation components of both care models, but is less important in the consumption components, with socioeconomic factors being greater drivers of impact heterogeneity. In particular, there is large variation in technology ownership and usage across quintiles.

2.5.4 Subgroup (univariable) analyses

We use the heatmaps from Figure 2.4 to identify additional data-driven effect modifiers that we have not specified a priori, by selecting covariates with the largest differences in means between the highest and lowest quintiles. These include marital and employment status, the availability of basic household amenities (for example, electricity), household size, compulsory education, literacy, and technology usage (for example, internet usage). We also refer to Table 2.C.4 for the variable importance outputs from the causal forest, which in general identifies similar effect modifiers to those from our classification analysis, for example, age, employment, marital status, health care accessibility and household size. We find that the causal forest consistently picks up certain continuous variables, such as age and consumption expenditure, to be the most important drivers of heterogeneity in all data subsamples and across almost all three models of interest, which although plausible, may be due to some level of variable importance bias, as discussed earlier.

Figures 2.5 and 2.6 plot GATEs for population subgroups constructed according to our theory- and data-driven effect modifiers. The results are largely consistent with those found in our earlier analysis that explores the joint distribution of covariates across quintiles, but we also generate some new insights. We already know that older respondents tend to increase their demand for health care more than younger respondents as a result of insurance, however this is not a consistent finding for consumption effects. In fact, the increase in the number of outpatient visits is greatest among those aged 25-49. We also find that, although in general treatment effects increase with household wealth, the participation effect on outpatient demand is greater among poorer households. In addition, policy effects are greater for women compared to men across all models, except for the inpatient consumption model. An unsurprising finding is that insurance increases health care demand in areas where health care accessibility is considered to be easy. However, the disparity in treatment effects between areas with easy and difficult access is clear and highlights the fundamental need to align demand- and supply-side policies. Our data-driven results reaffirm the positive link between socioeconomic status and treatment effects since respondents that are more likely to have basic household amenities, such as

electricity, tend to consume more health. Technology usage is a key effect modifier, but surprisingly those who use internet are, in general, less likely to increase their health care demand compared to those who do not use internet. This is consistently true for all models of interest apart from the participation component for outpatient demand. Married respondents, as well as those out of work through unemployment or retirement, also tend to consume more health care, however this is likely correlated with age which we already know is a key driver of heterogeneity. Being literate and educated is associated with an increase in inpatient care but a decrease in outpatient care, compared to those that that are illiterate and uneducated. Lastly, very small (1-2 members) or very large (11+ members) households also have larger treatment effects. In particular, the larger the household, the greater the quantity of health care consumed.

2.5.5 Predictors of heterogeneity (multivariable analysis)

Figures 2.7 and 2.C.5 plot BLP coefficients from OLS regressions of doubly robust scores on the respective full and restricted covariate vectors. Using the full vector, we generally find similar results to our previous subgroup analyses, confirming age, health care accessibility, geographical region and measures of socioeconomic status as key predictors of heterogeneity. However, in contrast to our earlier findings, we can now summarise heterogeneity using a ceteris paribus interpretation. Further, we find that participation and consumption effects often vary, highlighting the importance of decomposing the overall effect, if possible. When we restrict our covariate vector to a limited set of theory- and data-driven effect modifiers, we find that effect magnitudes tend to be larger and the direction of effects is, for the most part, supportive of the BLP regression on the full covariate vector.

2.6 Discussion

In this paper, we evaluated the average and heterogeneous impacts of Indonesia's subsidised JKN programme on health care utilisation in 2017, three years after its

implementation. Our findings were consistent with those from previous impact evaluations of health insurance programmes in Indonesia that, on average, enrolment into the subsidised scheme eases the financial barrier to accessing health care, leading to an increase in utilisation (Erlangga et al., 2019a; Sparrow et al., 2013). Through our separate evaluations on inpatient and outpatient demand, we find that, similar to Erlangga et al. (2019a), there is a limited participation effect of subsidised health care. Unlike higher-income countries where a demand inducement generally leads to increased consumption, there are additional challenges associated with subsidised care in lower-income countries (for example, time and transport costs, or inadequate health facilities) that may reduce the incentive-based effects of insurance (Johar, 2009). We did, however, find greater consumption effects of insurance – that is, the intensity of health care consumed for those who choose to access it. In our study, we also gathered substantial evidence of heterogeneity in treatment effects for individuals and population subgroups. An unexpected finding was the negative participation effect on outpatient demand (based on the GATE) for 20% of the population, which contradicts the demand inducement described earlier. The existing evidence that reports minimal or adverse health effects of health insurance is currently limited (Chen and Jin, 2012; Dow and Schmeer, 2003; Fink et al., 2013), and even more so for utilisation outcomes (Sheth, 2014). By increasing the affordability of essential care, insurance may create a substitution effect between inpatient and outpatient demand, thus reducing the risk of patients delaying or foregoing treatment, or pursuing traditional therapies. This is supported by the relatively larger consumption effects on inpatient demand. By incorporating the two-part model, and its decomposition, into our analyses, we were able to make these conclusions. Similarly, our heterogeneity results reinforced the potential for CATEs to reveal more nuanced insights that would otherwise be hidden in a population-level analysis. We attributed significant variations in individual-level treatment effects to differences in observed characteristics. In line with previous impact evaluations of JKN, we found stronger policy effects among respondents living in urban areas, where health care is more accessible compared to its rural counterparts (Erlangga et al., 2019a). It is well understood that supply-side constraints limit the effectiveness of health financing policies, and inadequate access to good quality health care

remains a challenge in large parts of Indonesia and other low- and middle-income countries implementing similar schemes (Agustina et al., 2019). Although theory-driven covariates, such as age and household expenditure, were identified as the most important drivers of heterogeneity, our data-driven analysis additionally identified other important variables, such as marital and employment status, and other variables that are representative of socioeconomic status that may not have been considered in a traditional subgroup analysis.

We demonstrated the potential of combining machine learning in prediction and causal tasks. Ensemble machine learning is a promising alternative to parametric regression for nuisance parameter estimation, despite computational costs. Since we are only interested in evaluating our target causal parameters, and not the nuisance parameters, the "black-box" nature of ensemble learning – meaning that the underlying model cannot be easily interpreted – is not considered a limitation in our case study. Ensemble machine learning algorithms, such as the super learner, can be used to improve predictions of target parameters, provided that the researcher has sufficient knowledge to make a priori selections of the covariate vector, library of base learners and tuning parameters (Polley and van der Laan, 2010; Rose, 2013). Causal machine learning similarly has several advantages over conventional CATE estimation approaches (for example, simple interactions between treatment and effect modifiers in regressions), particularly in data sets with a large number of covariates (Strittmatter, 2019). We showed that CATE estimators, such as causal forests, can flexibly make predictions, while automatically controlling for observed confounding, incorporating interaction terms and enabling valid inference. The nuisance parameter and CATE predictions can be used to construct doubly robust proxies for the CATEs, the doubly robust scores, on which we focused our inference using summary measures and projections. This is a generic inferential approach that enables valid inference even if the predicted CATEs are not consistent estimates of the true CATEs. We summarised and interpreted treatment effect heterogeneity using a combination of graphical and statistical approaches that form and describe population subgroups, constructed either using ranked CATEs or theory- and data-driven effect modifiers, which helped us to learn about the types of individuals that benefitted the most and the least from the policy. Throughout, we took advantage

of sample splitting and cross-fitting to avoid overfitting, while achieving full sample efficiency.

Our study has some limitations. First, we use a selection on observables framework that relies on the strong assumption of unconfoundedness. While we try to capture as much observed confounding as possible by including a rich set of covariates, we may not capture all unobserved confounding, which could result in biased treatment effect estimates. Potential confounding could arise from selection bias relating to the insurance enrolment process, particularly as those who are not automatically enrolled may voluntarily register based on their health care needs. Recent methodological advances that allow data-adaptive estimation of CATEs in the presence of unobserved confounding could support future research for this evaluation problem, for example incorporating instrumental variables analysis or panel data methods into the causal forests estimator (Athey et al., 2019a,b; Johnson et al., 2022; Kallus et al., 2019). Second, our study demonstrates the benefits of ensemble machine learning as a prediction tool for nuisance parameter estimation. Although we select a diverse set of parametric and non-parametric base learners in our ensemble, each learner and its model parameters could be optimised further through tuning. However, for simplicity and explainability, we select our base learners according to theory and choose mainly default parameter settings, where appropriate. Since the causal forest employs orthogonalisation to control for confounding, we are not overly concerned with a slight misspecification of the nuisance parameter models.

The findings from this study can form a starting point for thinking about more complex interactions between policy interventions and effect modifiers. In our setting, we limit our subgroup analysis to single binary covariates for interpretability, but in reality, describing the winners and losers of a policy intervention could involve exploring combinations of covariates at higher dimensions. If we are able to harness this information, the scope of personalised decision making could be considerably greater. This idea has sparked a related, rapidly growing literature on methods for designing "optimal policy rules" that map combinations of individual characteristics to a treatment decision (Atan et al., 2018; Athey and Wager, 2021; Kallus, 2018, 2020; Kallus and Zhou, 2018a; Kitagawa and Tetenov, 2018; Luedtke and van der

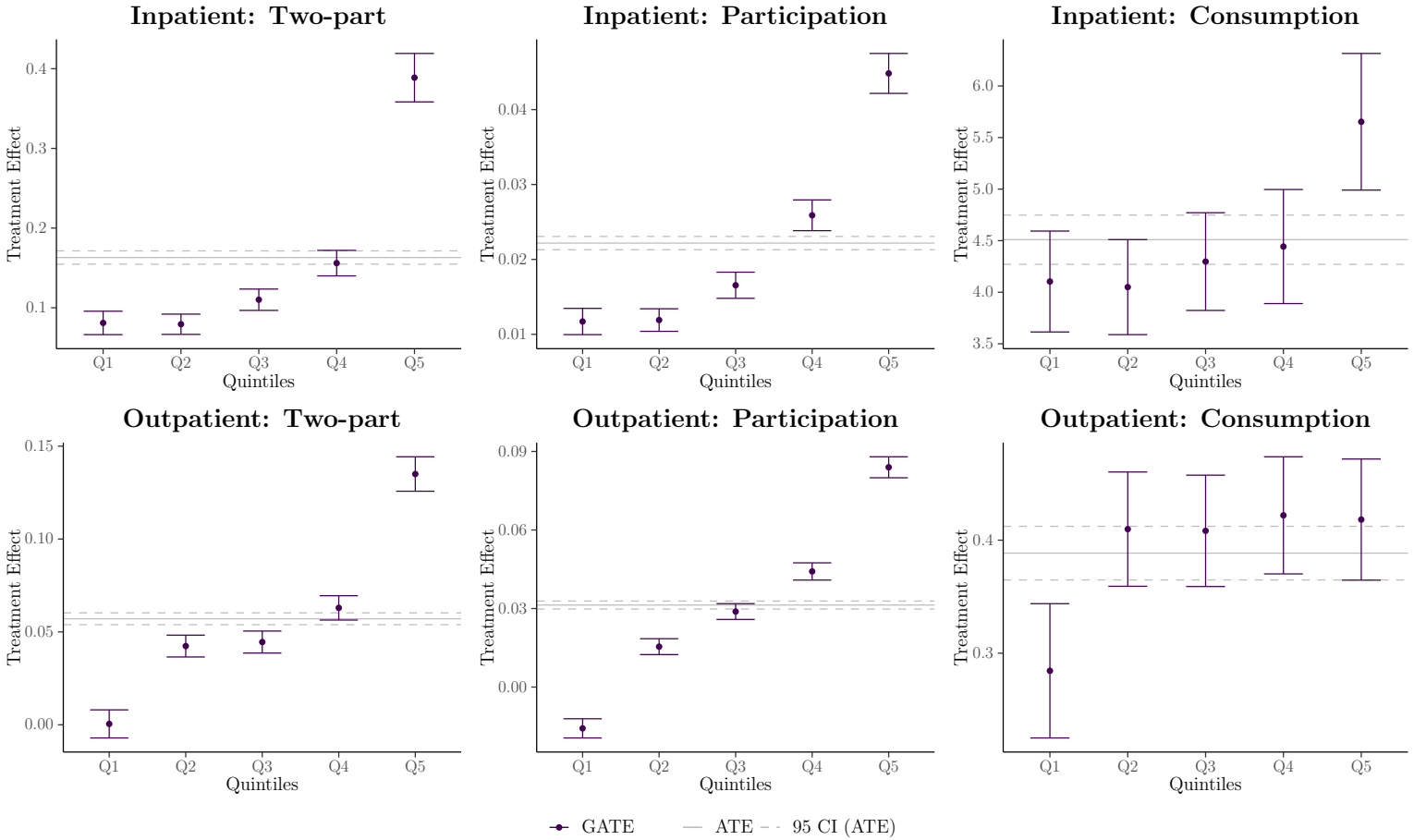
Laan, 2016a,b; Zhou et al., 2022). The objective is to find the optimal policy rule, among the space of all available policies that approximately maximises average outcomes across the population. The policy class can be constrained to satisfy criteria important to decision makers, such as budget and equity constraints, or rules that target only a selection of pre-specified covariates. For large-scale health policies, such as the subsidised JKN programme, an optimal policy rule could be used to prioritise treatment for uninsured individuals that are expected to benefit the most, as per the heterogeneity analysis. The rule could, more broadly, promote equitable implementation based on geographic and socioeconomic factors, for example, while also respecting financial constraints. These are common considerations in universal health coverage policies.

Table 2.2: Descriptive statistics

	Uninsured		Insured		SMD	
	Mean	SD	Mean	SD	Raw	IPTW
Outcomes						
Had inpatient treatment	0.026	0.158	0.043	0.203	0.018	0.023
Total length of inpatient stay in past one year (days)	0.121	1.151	0.247	1.972	0.078	0.095
Had outpatient treatment	0.112	0.316	0.128	0.334	0.015	0.026
Total number of outpatient visits in past one month	0.160	0.563	0.195	0.687	0.056	0.073
Household member characteristics						
Male	0.508	0.500	0.503	0.500	-0.005	-0.003
Age	26.7	19.8	30.5	19.5	0.197	-0.012
Education: compulsory	0.554	0.497	0.629	0.483	0.075	-0.006
Education: non-compulsory	0.236	0.424	0.248	0.432	0.013	0.000
Literate: Latin letters	0.789	0.408	0.871	0.335	0.082	-0.004
Employment status: in employment	0.403	0.490	0.455	0.498	0.053	0.001
Employment status: student	0.153	0.360	0.180	0.384	0.026	-0.002
Marital status: married	0.433	0.495	0.482	0.500	0.050	0.001
Used internet in previous 3 months	0.241	0.428	0.239	0.426	-0.003	0.004
Travelled domestically for tourism in 2016	0.237	0.425	0.218	0.413	-0.019	0.007
Household characteristics						
Location: urban	0.483	0.500	0.459	0.498	-0.024	-0.008
Number of people in household	5.191	2.017	5.282	2.007	0.045	-0.020
Home occupancy status: owner	0.834	0.372	0.850	0.358	0.016	0.006
Has a second home	0.092	0.289	0.115	0.318	0.022	-0.009
Toilet: private	0.876	0.329	0.829	0.376	-0.047	0.010
Purchases drinking water	0.431	0.495	0.411	0.492	-0.020	-0.002
Electricity	0.982	0.133	0.970	0.171	-0.012	0.000
Goods ownership: car	0.102	0.303	0.067	0.250	-0.036	0.004
Natural disaster in previous year	0.112	0.316	0.138	0.345	0.026	-0.006
Not enough food to eat in previous year	0.249	0.433	0.291	0.454	0.042	-0.013
Has a savings account	0.461	0.498	0.432	0.495	-0.029	0.008
Monthly consumption expenditure per capita (IDR 100,000)	8.469	7.194	7.815	6.384	-0.096	0.025
Health care accessibility						
Easy access: secondary care	0.757	0.205	0.741	0.221	-0.076	0.015
Easy access: community care	0.920	0.114	0.913	0.127	-0.059	-0.001
Easy access: primary care	0.819	0.179	0.796	0.213	-0.116	0.030
Easy access: maternal care	0.854	0.176	0.840	0.203	-0.075	-0.001
Region						
Region: Sumatera	0.213	0.409	0.250	0.433	0.038	0.005
Region: Jakarta	0.019	0.137	0.039	0.194	0.020	-0.007
Region: Java	0.536	0.499	0.458	0.498	-0.078	0.017
Region: Bali,NTB,NTT	0.064	0.244	0.062	0.242	-0.001	-0.006
Region: Kalimantan	0.073	0.260	0.046	0.210	-0.026	0.003
Region: Sulawesi	0.070	0.255	0.099	0.299	0.030	-0.011
Region: Maluku-Papua	0.026	0.159	0.044	0.206	0.018	0.000

Note: Sample means and standard deviations (SD) are reported for selected variables in X for the uninsured and insured (enrolled into subsidised JKN) populations. SMD = standardised mean difference. Raw = unweighted. IPTW = inverse probability of treatment weighting for the ATE. All observations are weighted by SUSENAS frequency weights.

Figure 2.3: Estimated sorted GATEs



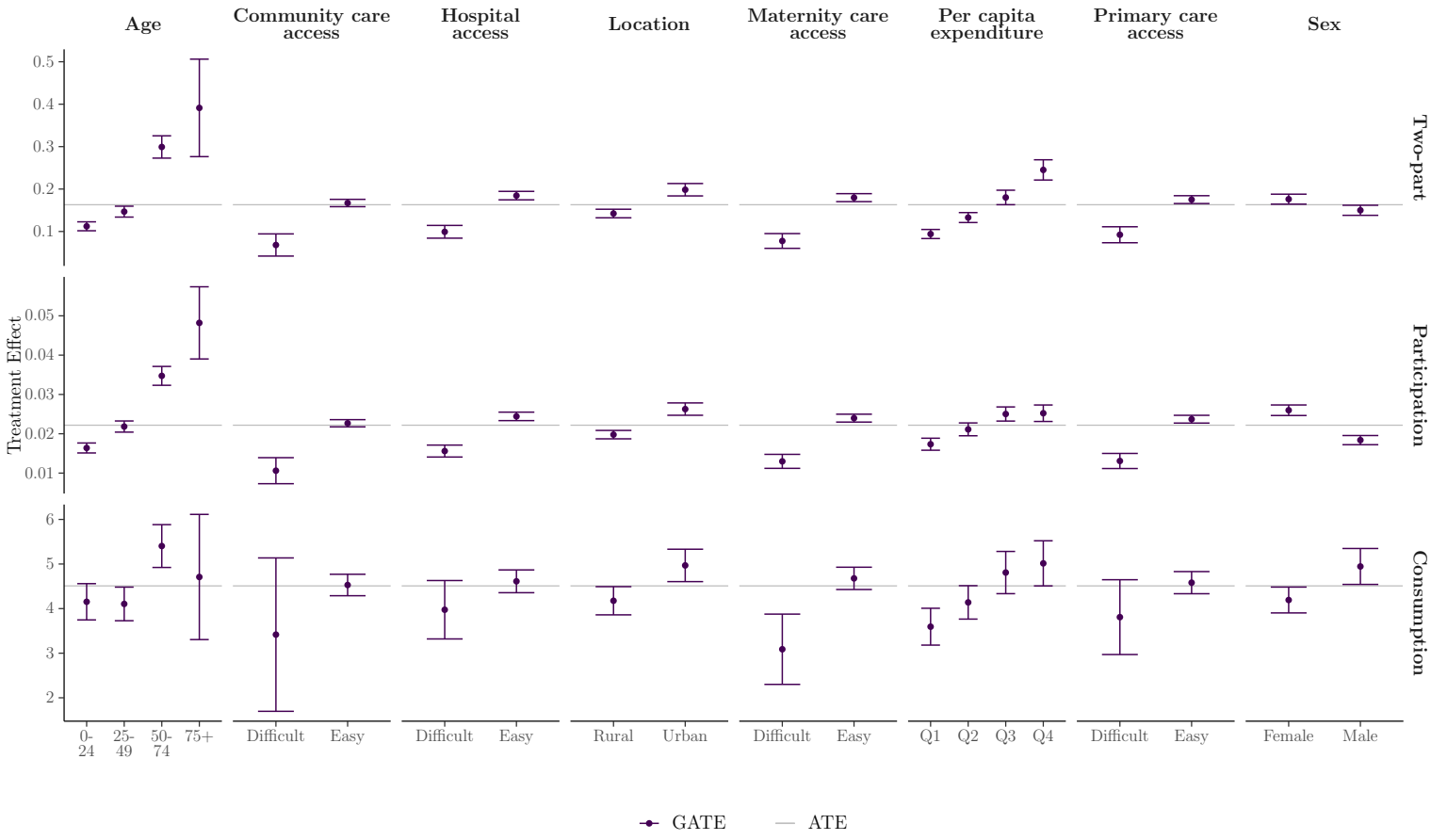
Note: Point estimates and 95% confidence intervals (error bars) are reported. GATEs are the estimated ATEs for each quintile of the population, ranked (in ascending order) by predicted CATEs. ATE point estimates and 95% confidence intervals are also reported. Graphs are separated by type of health care utilisation (inpatient and outpatient) and outcome model (two-part, participation component and consumption component).

Figure 2.4: Heatmaps for classification analysis (two-part model)

Note: Heatmaps show the variation in the joint distribution of covariates for subgroups with different CATEs, estimated using the two-part outcome model $\hat{m}(x)$. The annotated text shows the average value and standard error (in parentheses) of each covariate within each quintile (of predicted CATEs). The colour is a normalised distance of each quintile-specific covariate mean from the overall covariate mean. The darker the colour, the greater the distance. Only the top 20 covariates in X are plotted, in descending order of variation between Q1 and Q5.

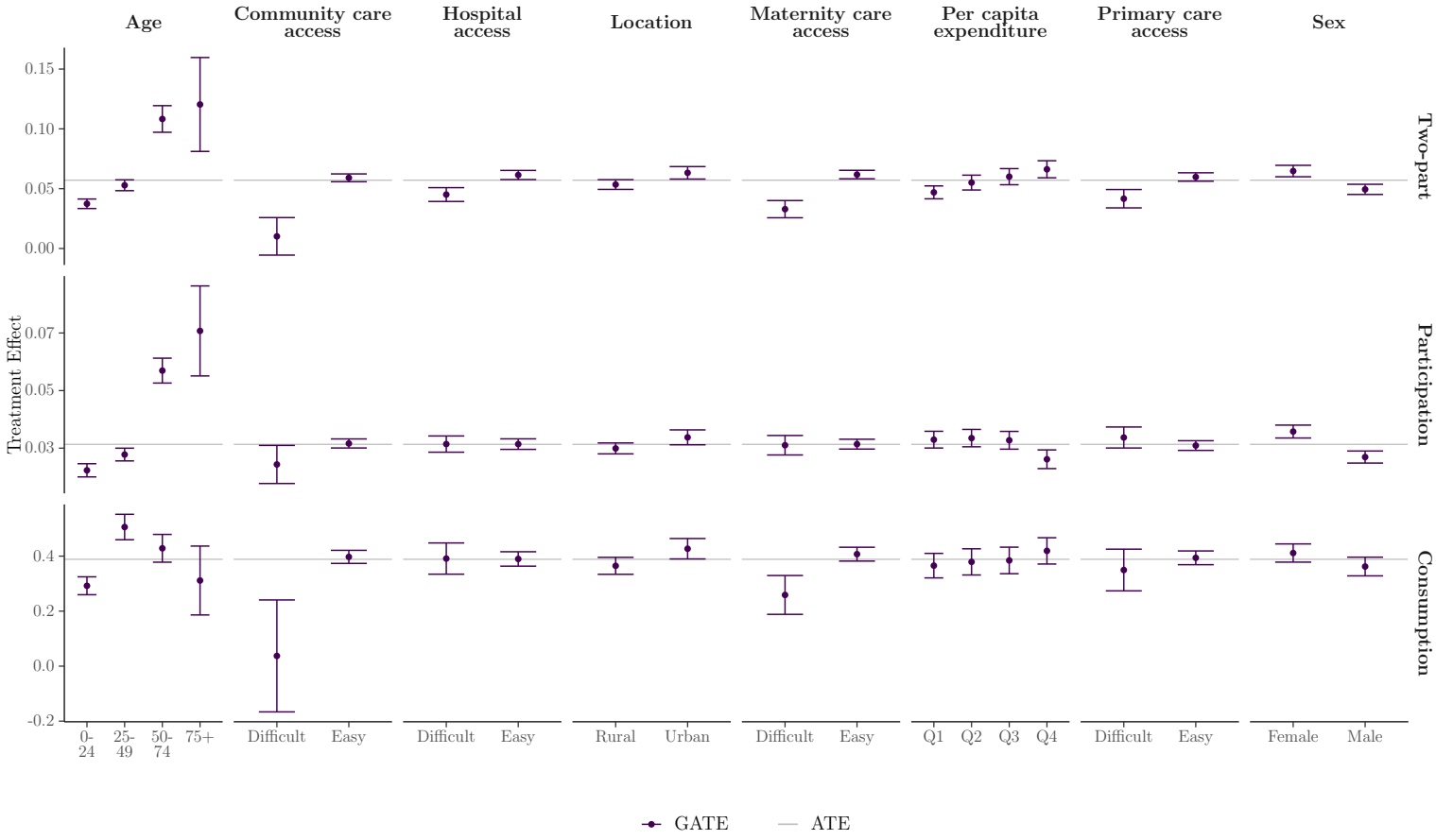
Figure 2.5: Estimated GATEs for theory-driven effect modifiers

(a) Inpatient demand



Note: Point estimates and 95% confidence intervals (error bars) are reported. ATE point estimates are also reported.

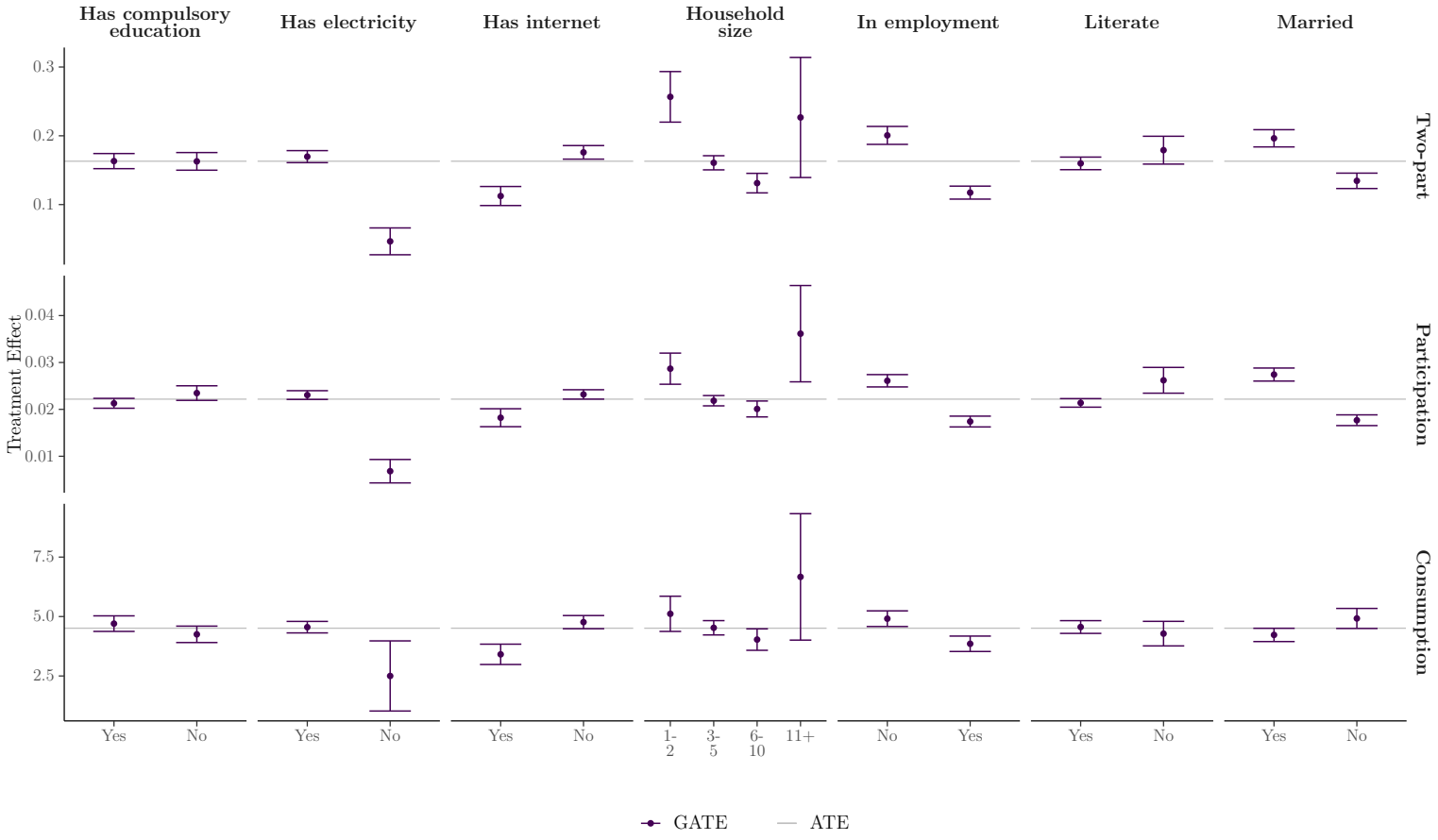
(b) Outpatient demand



Note: Point estimates and 95% confidence intervals (error bars) are reported. ATE point estimates are also reported.

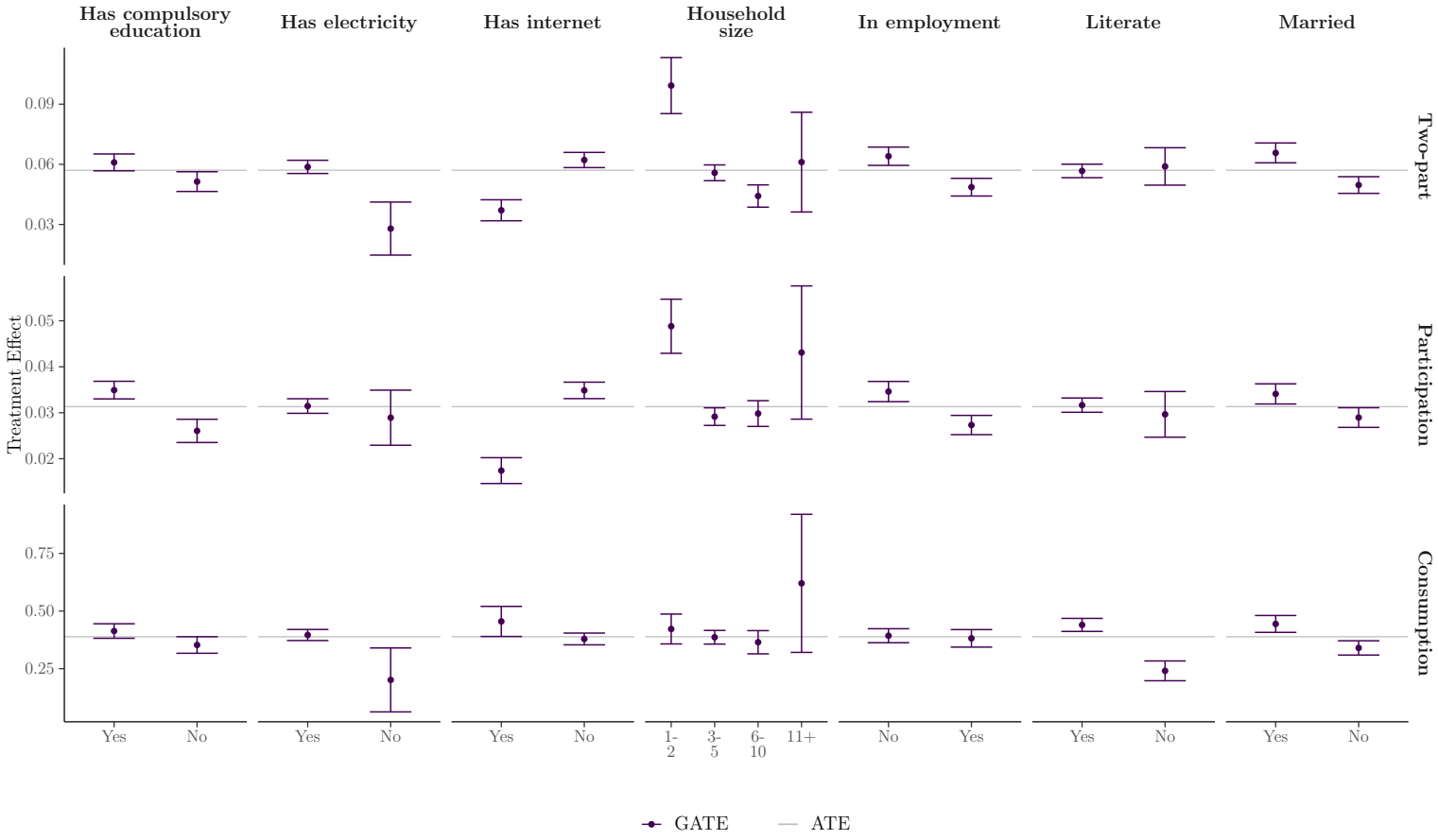
Figure 2.6: Estimated GATEs for data-driven effect modifiers

(a) Inpatient demand



Note: Point estimates and 95% confidence intervals (error bars) are reported. ATE point estimates are also reported.

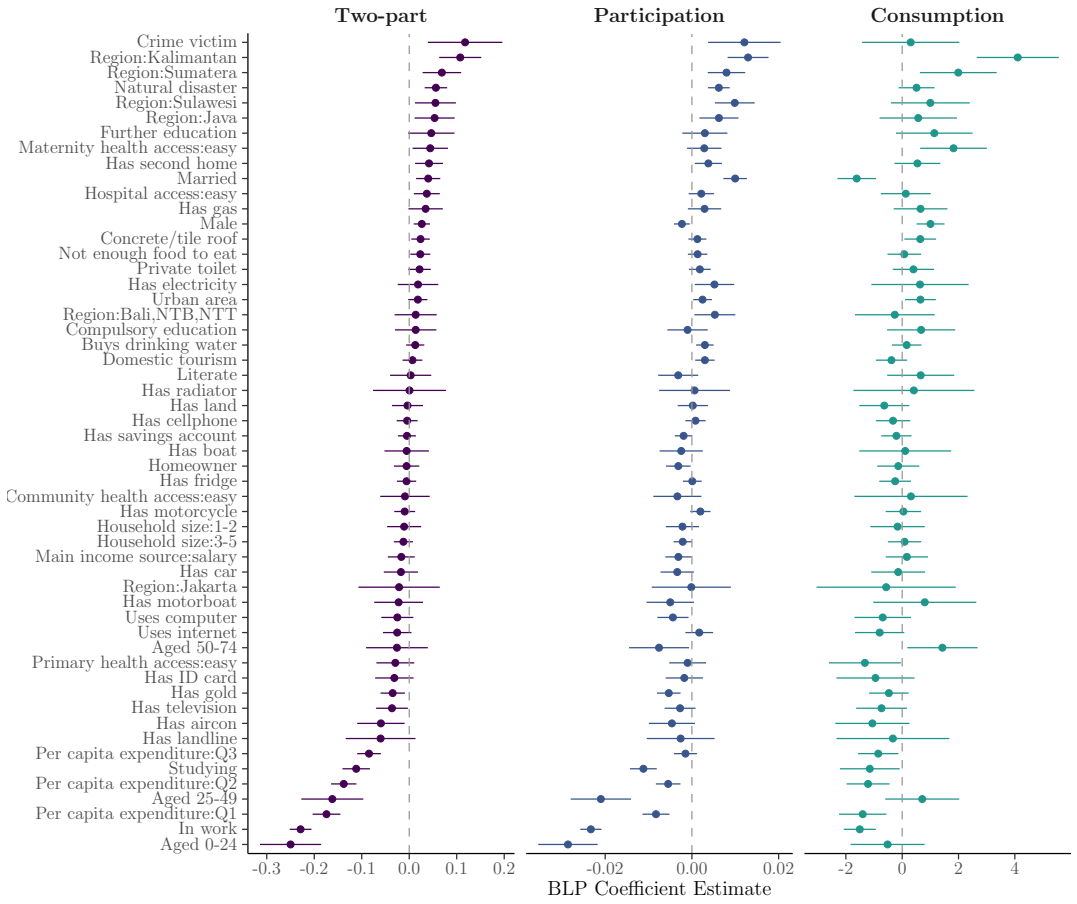
(b) Outpatient demand



Note: Point estimates and 95% confidence intervals (error bars) are reported. ATE point estimates are also reported.

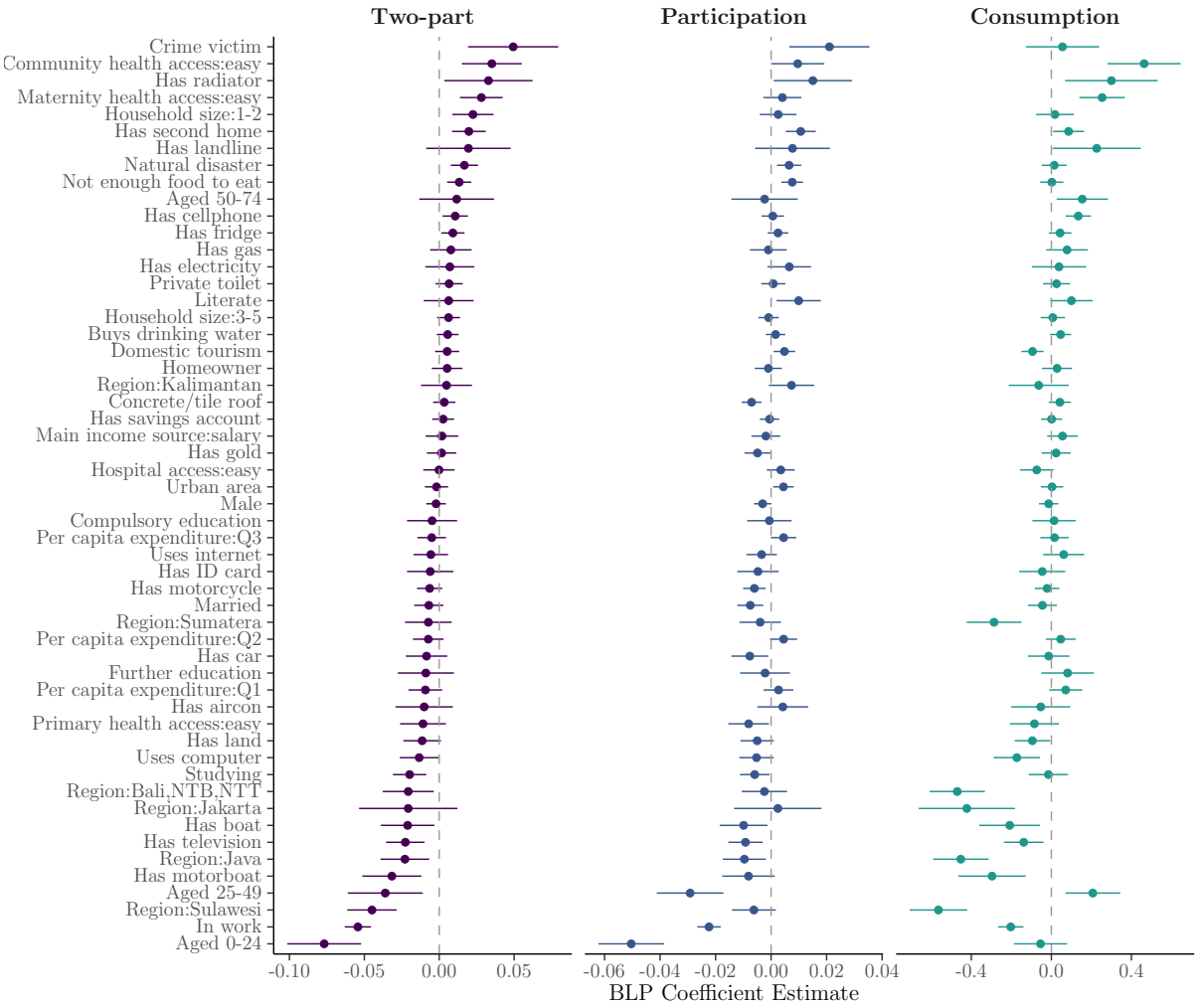
Figure 2.7: Estimated coefficients from the Best Linear Projection (BLP) of Γ_i^{cf} on X_i

(a) Inpatient demand



Note: BLP coefficients $\hat{\beta}$ from an OLS regression of $\hat{\Gamma}_i^{cf}$ on X_i ($\hat{\Gamma}_i^{cf} = \hat{\beta}_0 + \boldsymbol{X}\hat{\beta}$). Continuous variables have been converted to discrete variables. Reference categories include: Age 75+, Per capita expenditure: Q4, Region: Maluku-Papua and Household size:11+. All other variables have a binary interpretation.

(b) Outpatient demand



Note: BLP coefficients $\hat{\beta}$ from an OLS regression of $\hat{\Gamma}_i^{cf}$ on X_i ($\hat{\Gamma}_i^{cf} = \hat{\beta}_0 + \boldsymbol{X}\hat{\beta}$). Continuous variables have been converted to discrete variables. Reference categories include: Age 75+, Per capita expenditure: Q4, Region: Maluku-Papua and Household size:11+. All other variables have a binary interpretation.

Appendix 2.A List of all covariates used for confounder adjustment

Household member-level (binary)

Male

Female

Marital status: single

Marital status: married

Marital status: divorced

Marital status: widow(er)

Has a national identity number

Literacy: Latin letters

Literacy: Arabic letters

Literacy: Other letters

Education: compulsory

Education: non-compulsory

Travelled domestically for tourism in 2016

Victim of crime between March 2016-February 2017

Had a cellphone in previous 3 months

Used a computer in previous 3 months

Used internet in previous 3 months

Employment status: in employment

Employment status: student

Employment status: housekeeper

Employment status: out of employment

Employment status: performs other activities

Household member-level (continuous)

Age

Household-level (binary)

Location: urban

Did not have enough food to eat in previous year

Home occupancy status: owner

Home occupancy status: renter

Home occupancy status: rent-free

Home occupancy status: company-owned

Home occupancy status: other

Has a second home

Roof: concrete

Roof:tile

Roof: asbestos

Roof: zinc

Roof: bamboo

Roof: wood/shingle

Roof: straw/fiber/leaves/metroxylon sagu

Toilet: private

Toilet: shared

Toilet: none

Drinking water: bottled

Drinking water: tap

Drinking water: pump

Drinking water: protected well

Drinking water: unprotected well

Drinking water: protected spring

Drinking water: unprotected spring

Drinking water: river

Drinking water: rain

Drinking water: other

Purchases drinking water

Has electricity

Has experienced a natural diaster in previous year

Has natural tourism in residential area

Has a savings account

Goods ownership: gas

Goods ownership: fridge

Goods ownership: air conditioning

Goods ownership: radiator

Goods ownership: landline

Goods ownership: gold

Goods ownership: motorcycle

Goods ownership: boat

Goods ownership: motorboat

Goods ownership: car

Goods ownership: television

Goods ownership: land

Main income source: salary

Main income source: money transfer

Main income source: investments

Main income source: pension

Household-level (continuous)

Number of people in household

Number of children in household

Number of households in census building/house

Number of families in census building/house

Number of rooms in census building/house

Monthly consumption expenditure per capita (IDR 100,000)

District-level (binary)

Easy access to primary health care

Easy access to community health care

Easy access to maternal health care

Easy access to secondary (hospital) health care

Regional-level (binary)

Region: Sumatera

Region: Jakarta

Region: Java

Region: Bali, NTB, NTT

Region: Kalimantan

Region: Sulawesi

Region: Maluka-Papua

Appendix 2.B Causal forests

Causal forests are an ensembling algorithm comprising B bootstrapped samples, where $b = 1, \ldots, B$ is one sample, or "tree". Similar to its prediction counterpart, the random forest, each tree b partitions the covariate vector X into two equally sized subsets using a recursive splitting process that seeks a splitting variable and a cut-off threshold that optimises a target criterion. For random forests, the criterion is to minimise a loss function (for example, the mean squared error), whereas for causal forests, it is to maximise heterogeneity in treatment effects. The causal forest employs an additional concept, known as "honesty", that divides the bootstrapped sample b into two equal subsamples $\mathcal{S}_1$ and $\mathcal{S}_2$ before splitting, in order to separate the data used to train the tree and make predictions. The subsample S_1 is used for training and begins the splitting process by redefining itself as the splitting (or parent) node $\mathcal{P}$. Since optimising the loss function over all possible splits is too computationally expensive, the algorithm instead captures treatment effects locally through gradient-based optimisation. To determine the split at $\mathcal{P}$, the algorithm calculates pseudo-outcomes ρ_i for $i \in \mathcal{P}$ as follows:

$$\rho_i = A_{\mathcal{P}}^{-1}(D_i - \hat{e}(X_i) - \bar{D}_{\mathcal{P}})(Y_i - \hat{m}(X_i) - \bar{Y}_{\mathcal{P}}) - (D_i - \hat{e}(X_i) - \bar{D}_{\mathcal{P}})\hat{\beta}_{\mathcal{P}}/\mathrm{Var}_{\mathcal{P}}(D_i - \hat{e}(X_i)),$$
$$A_{\mathcal{P}} = \frac{1}{I(i \in \mathcal{P})} \sum_{i \in \mathcal{P}} (D - \bar{D}_{\mathcal{P}})^2,$$

where $\bar{D}_P$ and $\bar{Y}_P$ are the averages of D_i and Y_i over the observations in the parent node P, and $\hat{\beta}_{\mathcal{P}}$ is the solution of the least-squares regression of $Y_i - \hat{m}(X_i)$ on $D_i - \hat{e}(X_i)$ in $\mathcal{P}$. The recursive splitting process is subsequently used on the pseudo outcomes ρ_i, D_i and X_i to partition the parent node $\mathcal{P}$ into two further subsets (or child nodes) C_1 and C_2 by maximising the following criterion:

$$\mathrm{argmax}_{C_1, C_2} \frac{\sum_{i \in C_1} \rho_i^2}{\sum_{i \in \mathcal{P}} I(i \in C_1)} + \frac{\sum_{i \in C_2} \rho_i^2}{\sum_{i \in \mathcal{P}} I(i \in C_2)}.$$

This process of recursive partitioning between the parent and child nodes continues until the terminal nodes (or leaves) contain a minimum number of treated and control observations (this number can be optimised via cross-validation). Ultimately,

the splitting criterion rewards partitions that find strong heterogeneity, such that the within-partition variance of the pseudo-outcomes is maximised. Once the full tree has been grown, the reserved subsample $\mathcal{S}_2$ is used to populate the leaves by pushing each observation down the tree. The proportion of observations in $\mathcal{S}_2$ that fall into the leaves is calculated, and each observation i is assigned a weight w_{bi} equal to this proportion if it falls into the same leaf that contains $X_i = x$, or zero otherwise. This process is repeated across B trees such that:

$$w_i(x) = \frac{1}{B} \sum_{b=1}^{B} w_{bi}(x).$$

The causal forests `grf` package in `R` includes a `variable_importance` function that calculate a measure of 'importance' for each covariate in X_i using the following formula:

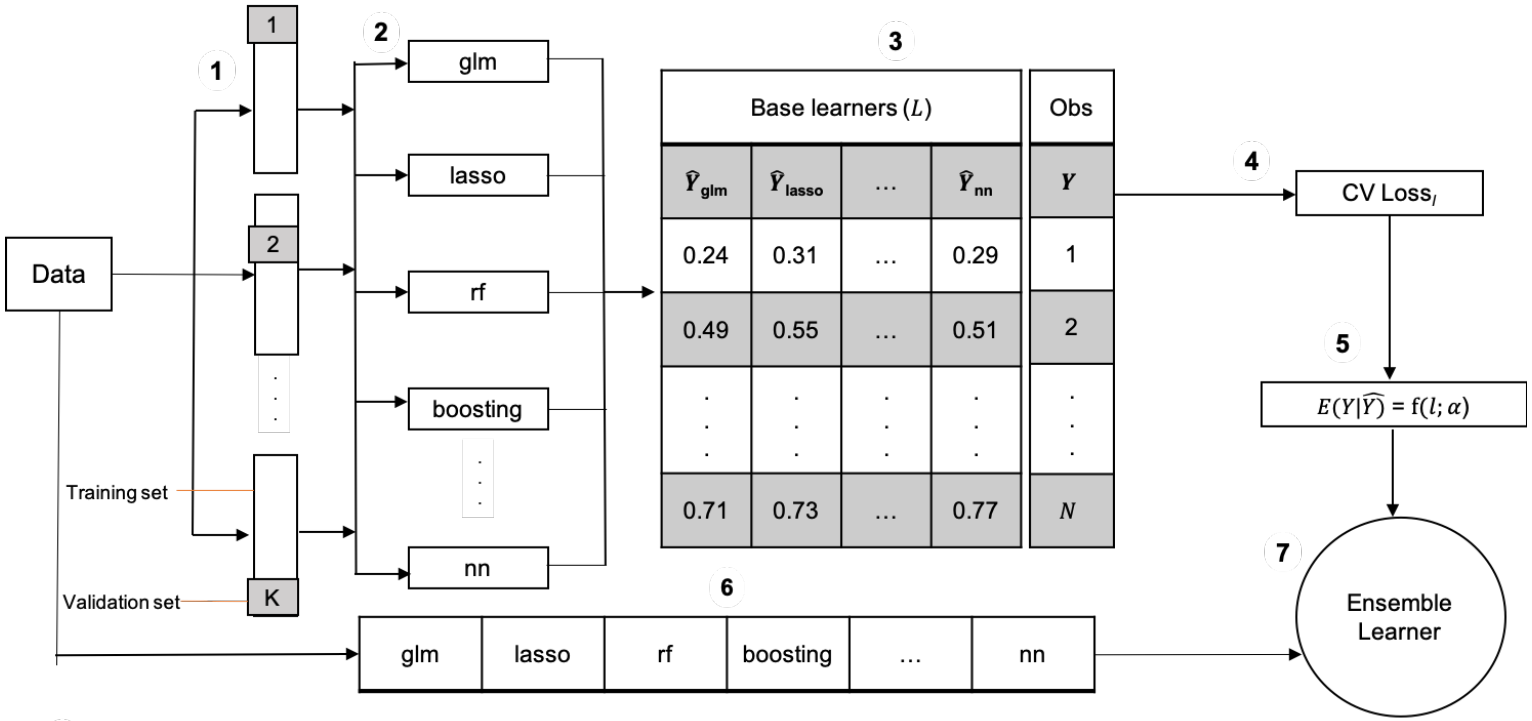
$$\text{Importance}(x_i) = \frac{\sum_{l=1}^{\text{max.depth}} \left(\frac{\sum_{b=1}^{B} \text{number of splits at depth } l \text{ on } x_i \text{ in tree } b}{\sum_{b=1}^{B} \text{total number of splits at depth } l \text{ in tree } b} \right) \cdot l^{-2}}{\sum_{l=1}^{\text{max.depth}} l^{-2}},$$

where the maximum depth of tree B is user-defined. Note that continuous variables are more likely to be considered "important" compared to binary variables as they have more potential splitting points.

Causal forests can also incorporate sample weights into their estimation model, so that treatment effect estimates are representative of the target population. For CATEs, sample weights do not change the causal estimand, but instead prioritise fit on the target population by minimising a weighted version of the loss function constructed using the CATE estimator. For average treatment effects, sample weights do change the causal estimand by estimating the average effect over the sample-weighted population.

Appendix 2.C Additional figures/tables

Figure 2.C.1: Super learner workflow



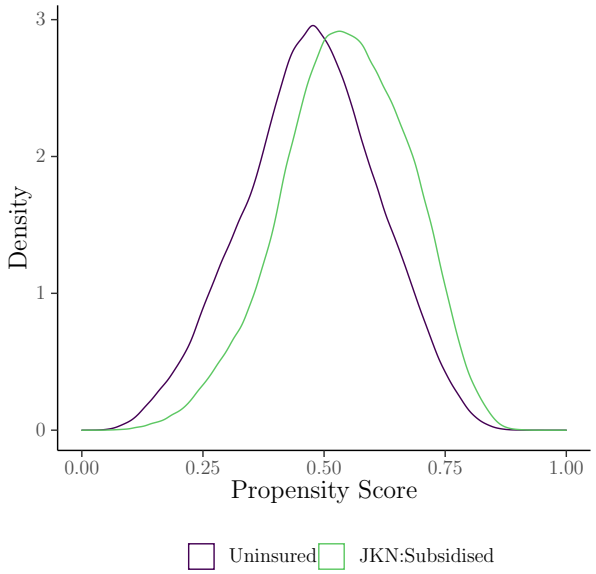
1. Divide data into k=1,…,K folds
2. Train each base learner on the training set for each fold and obtain predictions for the validation set
3. Predict the outcome Y for each base learner l=1,…,L and for each unit i=1,…,N using the validation set in each fold
4. Compute the cross-validated loss for each base learner using the predicted values $\hat{Y}_{l,i}$ from step 3
5. Calculate the optimal weighted combination of base learners by regressing Y on $\hat{Y}$ to estimate the weight vector α
6. Train each of the L base learners on the complete data set
7. Combine the fitted values from step 6 with the estimated weights to form the ensemble learner

Table 2.C.1: Causal forest tuning parameters

		Inpatient									**Outpatient**								
		Two-part			Participation			Consumption			Two-part			Participation			Consumption		
Tuning parameter	Description	s_1	s_2	s_3	s_1	s_2	s_3	s_1	s_2	s_3	s_1	s_2	s_3	s_1	s_2	s_3	s_1	s_2	s_3
sample.fraction	Fraction of the data used to build each tree	0.43	0.41	0.46	0.41	0.49	0.50	0.48	0.50	0.38	0.42	0.39	0.44	0.48	0.49	0.50	0.50	0.34	0.41
mtry	Number of variables tried for each split	22	29	27	23	25	26	27	28	17	24	24	28	23	25	26	30	25	22
min.node.size	Minimum number of observations in each tree leaf	4	2	2	1	1	1	1	1	1	1	1	2	1	1	1	5	2	1
honesty.fraction	Fraction of data used for determining splits	0.72	0.74	0.71	0.75	0.73	0.75	0.63	0.65	0.67	0.78	0.60	0.75	0.67	0.73	0.75	0.50	0.57	0.77
honesty.prune.leaves	Prunes estimation sample tree such that no leaves are empty	0	1	1	1	1	0	1	1	1	0	1	0	1	1	0	1	0	0
alpha	Maximum imbalance of a split	0.01	0.00	0.00	0.03	0.04	0.04	0.07	0.04	0.00	0.04	0.04	0.12	0.05	0.04	0.04	0.05	0.08	0.13
imbalance.penalty	Controls how harshly imbalanced splits are penalised	1.43	0.27	0.37	0.44	0.47	0.06	2.14	0.17	0.29	0.90	0.35	1.38	0.30	0.47	0.06	0.00	0.76	0.28

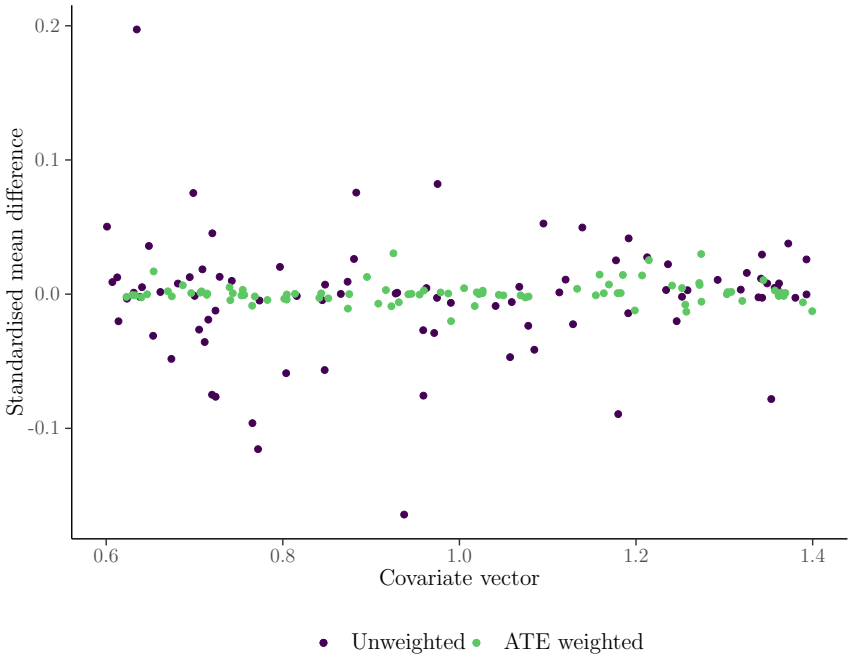
Note: List of all tuning parameters that are tuned by cross-validation in the training of the causal forests. Values are reported for causal forests trained on each subsample $\{s_1, s_2, s_3\}$.

Figure 2.C.2: Overlap plot



Note: Density plot showing the distribution of predicted propensity scores for the treated (subsidised JKN) and the controls (uninsured) in the pooled data.

Figure 2.C.3: Covariate balance between the subsidised JKN and the uninsured populations



Note: Standardised mean differences are reported for all X between the treated (subsidised JKN) and the controls (uninsured) before and after inverse probability of treatment weighting for the ATE.

Table 2.C.2: Estimated cross-validated loss and standardised coefficients from fitted super learners

	Propensity score		**Outcome regression**							
			Inpatient				Outpatient			
			Participation		Consumption		Participation		Consumption	
Algorithm	Loss	Coef	Loss	Coef	Loss	Coef	Loss	Coef	Loss	Coef
ensemble	0.550	-	0.142	-	2.545	-	0.342	-	0.376	-
gbm	0.552	1.246	0.140	0.237	2.588	0.004	0.339	0.447	0.376	0.110
glm-b	0.655	0.000	0.143	0.045	-	-	0.347	0.039	-	-
glm-g	-	-	-	-	2.552	0.092	-	-	0.381	0.000
glm-nb	-	-	-	-	3.243	0.000	-	-	1.070	0.000
glm-p	-	-	-	-	2.596	0.000	-	-	0.382	0.019
lasso-b	0.656	0.000	0.143	0.081	-	-	0.347	0.053	-	-
lasso-g	-	-	-	-	2.566	0.000	-	-	0.388	0.000
lasso-nb	-	-	-	-	3.243	0.000	-	-	1.070	0.000
lasso-p	-	-	-	-	2.576	0.003	-	-	0.382	0.000
nn	0.656	0.000	0.161	0.000	2.578	0.022	0.361	0.027	0.388	0.013
rf	0.673	0.000	0.145	0.167	2.549	0.090	0.355	0.068	0.382	0.009

Note: All estimates are averaged across the three subsamples $\{s_1, s_2, s_3\}$. Cross-validated loss estimates ("Loss") are reported for the base learners and the ensemble learner. Loss functions for the binary-response and count-response models are log loss and mean squared error respectively. Standardised regression coefficients ("Coef") are reported from the OLS regression of the observed outcomes on the cross-validated predicted outcomes. They can be interpreted as the "importance" of each base learner in making a prediction in the ensemble.

Table 2.C.3: Results from heterogeneity test using difference-in-means estimator

	Inpatient				**Outpatient**			
	Est	Std.Err	Unadj p-val	Adj p-val	Est	Std.Err	Unadj p-val	Adj p-val
Two-part								
Q2-Q1	-0.002	0.013	0.900	0.899	0.042	0.005	0.000	0.000
Q3-Q1	0.029	0.013	0.027	0.046	0.044	0.005	0.000	0.000
Q4-Q1	0.075	0.013	0.000	0.000	0.062	0.005	0.000	0.000
Q5-Q1	0.308	0.013	0.000	0.000	0.135	0.005	0.000	0.000
Participation								
Q2-Q1	0.000	0.001	0.892	0.893	0.031	0.002	0.000	0.000
Q3-Q1	0.005	0.001	0.001	0.001	0.045	0.002	0.000	0.000
Q4-Q1	0.014	0.001	0.000	0.000	0.060	0.002	0.000	0.000
Q5-Q1	0.033	0.001	0.000	0.000	0.100	0.002	0.000	0.000
Consumption								
Q2-Q1	-0.054	0.377	0.885	0.886	0.125	0.038	0.001	0.001
Q3-Q1	0.193	0.377	0.608	0.820	0.124	0.038	0.001	0.001
Q4-Q1	0.339	0.377	0.368	0.695	0.138	0.038	0.000	0.001
Q5-Q1	1.549	0.377	0.000	0.000	0.134	0.038	0.000	0.001

Note: Table reports estimates and standard errors of the differences in sorted GATEs (for quintiles of predicted CATEs) between the lowest quintile (Q1) and higher quintiles (Q2-Q5). Unadj p-val does not correct for multiple hypothesis testing. Adj p-val uses the Romano-Wolf procedure to correct for multiple hypothesis testing.

Figure 2.C.4: Heatmaps for classification analysis (participation and consumption models)

(a) Participation model

Note: Heatmaps show the variation in the joint distribution of covariates for subgroups with different CATEs, estimated using the participation component of the outcome model $\hat{m}_1(x)$. The annotated text shows the average value and standard error (in parentheses) of each covariate within each quintile (of predicted CATEs). The colour is a normalised distance of each quintile-specific covariate mean from the overall covariate mean. The darker the colour, the greater the distance. Only the top 20 covariates in X are plotted, in descending order of variation between Q1 and Q5.

(b) Consumption model

Note: Heatmaps show the variation in the joint distribution of covariates for subgroups with different CATEs, estimated using the consumption component of the outcome model $\hat{m}_2(x)$. The annotated text shows the average value and standard error (in parentheses) of each covariate within each quintile (of predicted CATEs). The colour is a normalised distance of each quintile-specific covariate mean from the overall covariate mean. The darker the colour, the greater the distance. Only the top 20 covariates in X are plotted, in descending order of variation between Q1 and Q5.

Table 2.C.4: Variable importance results from estimated causal forests

(a) Inpatient demand

	s_1		s_2		s_3	
Ranking	Effect modifier	Importance	Effect modifier	Importance	Effect modifier	Importance
Two-part						
1	Unemployed/retired	0.182	Unemployed/retired	0.226	Age	0.188
2	Age	0.174	Age	0.162	Unemployed/retired	0.168
3	Consumption expenditure	0.091	Consumption expenditure	0.119	Consumption expenditure	0.104
4	Widow(er)	0.078	Number of rooms	0.079	Household size	0.080
5	Working	0.073	Single	0.041	Company-owned house	0.057
Participation						
1	Age	0.199	Single	0.160	Single	0.187
2	Working	0.092	Age	0.139	Age	0.171
3	Widow	0.057	Married	0.075	Married	0.123
4	Single	0.054	Female	0.072	Student	0.061
5	Maternity care access	0.052	Male	0.061	Hospital access	0.042
Consumption						
1	Hospital care access	0.135	Consumption expenditure	0.153	Consumption expenditure	0.117
2	Consumption expendiutre	0.112	Unemployed/retired	0.077	Household size	0.068
3	Primary care access	0.088	Age	0.055	Age	0.066
4	Age	0.074	Rents house	0.054	Community care access	0.061
5	Maternity care access	0.053	Hospital access	0.053	Hospital access	0.058

Note: Top 5 important effect modifiers are reported based on the variable importance ranking from the trained causal forests fitted on each subsample $\{s_1, s_2, s_3\}$. Importance is measured as the weighted sum of the frequency with which the variable was used to split on at each depth in the forest.

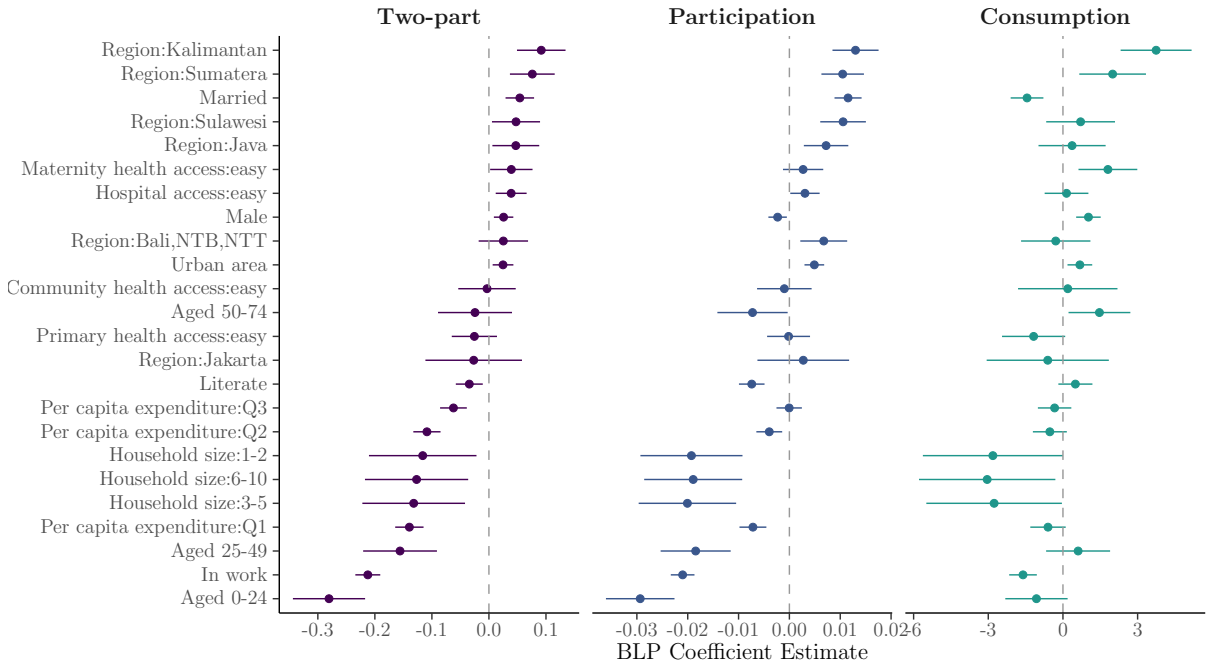
(b) Outpatient demand

	s_1		s_2		s_3	
Ranking	Effect modifier	Importance	Effect modifier	Importance	Effect modifier	Importance
Two-part						
1	Age	0.162	Age	0.224	Age	0.217
2	Primary care access	0.083	Widow(er)	0.090	Hospital access	0.087
3	Hospital access	0.079	Hospital access	0.072	Consumption expenditure	0.068
4	Widow(er)	0.073	Consumption expenditure	0.061	Single	0.059
5	Number of households	0.073	Maternity care access	0.059	Maternity care access	0.054
Participation						
1	Age	0.178	Age	0.212	Age	0.230
2	Hospital access	0.131	Widow	0.110	Single	0.068
3	Primary care access	0.066	Consumption expenditure	0.078	Widow(er)	0.065
4	Maternity care access	0.060	Hospital access	0.063	Hospital access	0.051
5	Consumption expenditure	0.059	Internet	0.059	Maternity care access	0.041
Consumption						
1	Consumption expenditure	0.084	Consumption expenditure	0.130	Consumption expenditure	0.107
2	Community care access	0.074	Maternity care access	0.095	Hospital access	0.107
3	Age	0.071	Age	0.088	Primary care access	0.096
4	Primary care access	0.068	Hospital access	0.067	Maternity care access	0.084
5	Maternity care access	0.066	Primary care access	0.059	Age	0.070

Note: Top 5 important effect modifiers are reported based on the variable importance ranking from the trained causal forests fitted on each subsample $\{s_1, s_2, s_3\}$. Importance is measured as the weighted sum of the frequency with which the variable was used to split on at each depth in the forest.

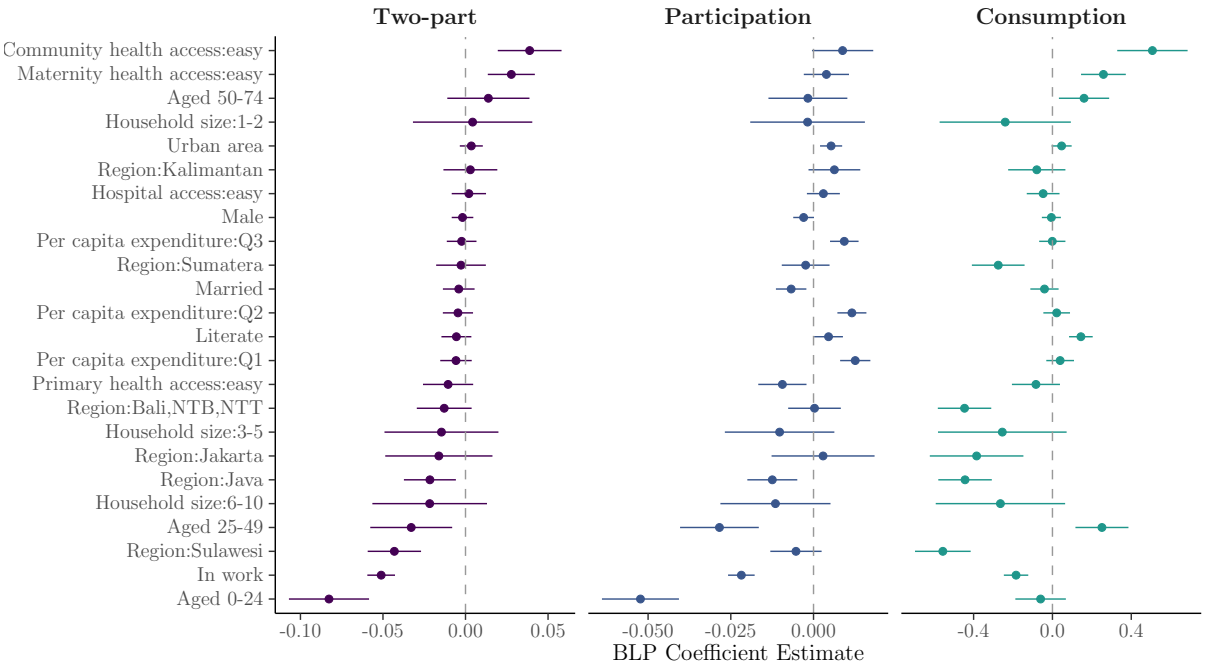
Figure 2.C.5: Estimated coefficients from the Best Linear Projection (BLP) of Γ_i^{cf} on $A_i \in X_i$

(a) Inpatient demand



Note: BLP coefficients $\hat{\beta}$ from an OLS regression of $\hat{\Gamma}_i^{cf}$ on $A_i \in X_i$ ($\hat{\Gamma}_i^{cf} = \hat{\beta}_0 + \boldsymbol{A}\hat{\beta}$). Continuous variables have been converted to discrete variables. Reference categories include: Age 75+, Per capita expenditure: Q4, Region: Maluku-Papua and Household size:11+. All other variables have a binary interpretation.

(a) Outpatient demand



Note: BLP coefficients $\hat{\beta}$ from an OLS regression of $\hat{\Gamma}_i^{cf}$ on $A_i \in X_i$ ($\hat{\Gamma}_i^{cf} = \hat{\beta}_0 + \boldsymbol{A}\hat{\beta}$). Continuous variables have been converted to discrete variables. Reference categories include: Age 75+, Per capita expenditure: Q4, Region: Maluku-Papua and Household size:11+. All other variables have a binary interpretation.

Chapter 3

Using Policy Learning to Reduce Catastrophic Health Expenditures: An Evaluation of Indonesia's JKN Insurance Schemes

Policy learning uses observational data to learn optimal policy rules that maximise a policy's expected outcomes, by mapping an individual's covariate profile to a policy decision. Popular approaches involve estimating heterogeneous policy impacts via the conditional average treatment effect (CATE) function, and assigning the policy to units with a beneficial impact. Alternatives include simpler rules that may allow for easier interpretation (for example, tree-based rules). We construct a rule that assigns households to Indonesia's two subsidised health insurance programmes, using an objective function that corresponds to the total reduction in the risk of incurring "catastrophic health expenditure"; a measure of the financial protectiveness of insurance. We use an ensemble algorithm, the super learner, that finds the best weighted combination of candidate estimators of the rule, including different regression specifications of CATE-based rules (for example, linear models and data-adaptive models) and simpler rules (for example, fixed-depth decision trees). We learn a second rule that constrains the proportion of households that can receive the policy, to reflect budget restrictions. We evaluate the rules using doubly robust estimators of the counterfactual mean outcome. The optimal unconstrained rule achieves better expected outcomes (that is, a reduction in the risk of incurring catastrophic expenditures) than the actual assignment and static rules, that assign the same programme to all. The optimal resource-constrained rule also outperforms the actual assignment. Geography, particularly the urban-rural distinction, is the

main differentiating factor between the assignment under the rules and the actual assignment, which is known to be associated with the availability of health services and a determinant of health spending.

3.1 Introduction

Policy impact evaluation that leverages upon existing data on current or past programmes has become an important tool for identifying whether a policy works, and for whom (Athey and Imbens, 2017). Estimates of heterogeneous treatment effects that capture varying policy impacts owing to differences in observable population characteristics, can be used to "learn" rules that assign the policy to those with the largest net benefits (Dehejia, 2005; Kitagawa and Tetenov, 2018; Manski, 2004). Essentially, policy learning takes an individual's observed covariate profile as an input, and outputs a policy decision (Montoya et al., 2022). The idea is to construct an optimal policy rule that achieves larger welfare gains than an assignment strategy that is either arbitrary (that is, not motivated by any evidence) or static (that is, assigning the policy to everyone or to no one) (Athey and Wager, 2021). Personalised policy rules can play an essential role in the allocation of social programmes. Many governments subsidise access to key services, such as health care, but are typically bound by resource constraints. When only a certain fraction of the population can be enrolled into a programme, an assignment strategy that prioritises eligibility to those that are expected to benefit the most can have an important effect on overall programme performance (Bhattacharya and Dupas, 2012).

There is a fast growing literature on learning optimal policy rules using observational data, where propensity scores are unknown and treatment assignment may be endogenous (Athey and Wager, 2021; Bertsimas and Kallus, 2020; Kallus, 2017, 2018; Kallus and Zhou, 2018b; Kitagawa and Tetenov, 2018; Zhou et al., 2022). Popular approaches use an assignment rule that relies on estimating the conditional average treatment effect (CATE) function that captures heterogeneity in treatment effects. Methodological advances in CATE estimation using machine learning, combined with the increasing availability of rich data to control for confounding, have

rapidly increased the scope of this field (Athey et al., 2019b; Kennedy, 2020; Künzel et al., 2019a; Nie and Wager, 2021; Shalit et al., 2017). CATE-based policy learning methods generally fall into two categories. The first category comprises rules that assign policy according to the sign of the CATEs, which can be estimated using traditional regression methods or more complex, data-adaptive methods (Luedtke and van der Laan, 2016b; van der Laan and Luedtke, 2015). We refer to this category as "threshold-based rules". The second category comprises simpler rules based on the CATE function that may allow for easier interpretation (Athey and Wager, 2021; Kitagawa and Tetenov, 2018). Decision trees are a popular representation for constrained policies since they produce an interpretable solution (Amram et al., 2022; Bertsimas et al., 2019; Sverdrup et al., 2020). The trees use the CATE estimates to directly select an optimal rule among different candidate rules that maximises expected outcomes. The CATE function itself is model agnostic, but is typically estimated using causal forests; a non-parametric causal machine learning algorithm (Athey et al., 2019b). We refer to this category as "tree-based rules". Other popular methods for policy learning that do not rely on estimating the CATE function, include other direct estimation approaches (e.g. outcome weighted classification), regression-based approaches (e.g. Q-learning), and optimal structural nested models (Murphy, 2003; Qian and Murphy, 2011; Robins, 2004; Zhao et al., 2012). Given the variety and scope of methods available for policy learning, it is often unclear which approach to use, and how to specify the estimation task. A popular prediction tool, the super learner (also known as stacked regression), addresses this problem, using ensembling and K-fold cross-validation to learn the optimal policy rule from a user-specified selection of candidate estimators of the rule. (Luedtke and van der Laan, 2016b; Montoya et al., 2021, 2022; van der Laan and Luedtke, 2015).

In this paper, we use the super learner to estimate optimal policy rules in the context of Indonesia's national health insurance scheme, the Jaminan Kesehatan Nasional (JKN). We restrict our attention to the subsidised version of the JKN (known as *Penerima Bantuan Iuran*, or PBI), which targets lower-income and vulnerable households by subsidising insurance premiums either at the national (PBI-APBN) or sub-national (PBI-APBD, formerly *Jamkesda*) level. Sub-national health care financing policies were initially created to target the large informal working sector

that are eligible for subsidised insurance (Sparrow et al., 2013). Although JKN has been successful in enrolling the majority of the population, there remains a relatively large, PBI-eligible, uninsured population, and so a targeting strategy to support enrolment has been proposed. We therefore remove the uninsured from our sample since the overall aim of JKN is to achieve full coverage and to leave no one uninsured. An evaluation of APBD compared to APBN could provide insights into the types of people that would benefit from being enrolled into sub-nationally- versus nationally-funded health care. Sparrow et al. (2013) suggest that an important motivation for decentralised health spending is that local governments are better placed to close the coverage gap since they are good at identifying and targeting local health care needs. At the same time, however, they are less likely to be able to address supply-side constraints (for example, health care availability and accessibility) compared to the national government, which could compromise the effectiveness of the policy.

Our aim is to develop an optimal policy rule from the perspective of the Indonesian state government that efficiently assigns each household to "treatment" (APBD) or to "control" (APBN), using an objective function where net benefits are measured in terms of financial protection; a key dimension of universal health coverage. We use the official indicator for financial protection among the Sustainable Development Goals (SDG indicator 3.8.2); catastrophic health expenditure. Health care expenditure is defined as being catastrophic if out-of-pocket payments on health exceed either 10% or 25% of total household income or consumption (Boerma et al., 2014; Saksena et al., 2014; Wagstaff, 2008; Wagstaff et al., 2018). We use the broader 10% measure in our study.[1]

We model our optimal policy rule under two scenarios; with and without resource constraints, learning a separate rule for each assignment problem. For the constrained policy, we limit the proportion of households that can be enrolled into APBD to 10%, to reflect the proportion of the state government's health care budget that is distributed to local governments. Our library of candidate estimators

[1]Some researchers adjust total household consumption by subtracting essential spending on items such as food, clothing and shelter. However, the process of deciding which goods should or should not be included in the subtraction is ambiguous (Wagstaff et al., 2018) Therefore, we use total consumption in our measure of financial protection, as per the official SDG indicator.

for the unconstrained policy include threshold-based rules (implied by parametric and non-parametric estimates of the CATE function), tree-based rules (we include shallow decision trees of varying depths) and static rules that either assign APBD or APBN to all households. Given the construction of constrained policies, only threshold-based rules can be included in the candidate library. We evaluate the performance of our learned rules using efficient, doubly robust estimators of the counterfactual mean outcome. We characterise households that are counterfactually assigned to APBD and APBN under the estimated optimal policy rules by comparing differences in covariate means of the learned subgroups to those under the actual assignment. For our study, we use large-scale data from the 2017 National Socioeconomic Survey (SUSENAS) and the 2018 Village Potential Statistics Census (PODES).

We find that the financial impact of being enrolled into APBD over APBN varies with households' observed covariates, which provides some justification for developing CATE-based policy rules. The most important predictors of treatment effect heterogeneity relate to households' demographic composition, dwelling characteristics and socioeconomic status. The assignment strategies under both the unconstrained and constrained rules have a lower expected risk of catastrophic expenditure than the actual assignment, according to the doubly robust estimator of the counterfactual mean outcome. Geography is one of the main differentiating features between the assignment under the optimal rules and the status quo. Households that are counterfactually assigned to APBD are more likely to live in urban areas, with better household facilities and more accessible health care, whereas more rural-based households are assigned to APBN, where health care is less accessible. This distinction is even more pronounced under the constrained rule. Some of the best performing candidate algorithms in the super learner library include simple linear models and shallow decision trees, which can, to some extent, be interpreted.

Our study contributes to the applied statistics and economics literature on how estimating treatment effect heterogeneity can be used to maximise expected benefits from social policies (Bhattacharya and Dupas, 2012; Imai and Strauss, 2011; Kube et al., 2019). To our knowledge, this is one of the first studies to apply policy learn-

ing methods to the evaluation of a system-level health policy. We also contribute to the health economics literature on the impact of subsidised health insurance on financial protection, which is currently of interest given the renewed commitment to universal health coverage in many developing countries (Banerjee et al., 2021). We conduct an impact evaluation of Indonesia's JKN that aims to uncover any differential effects on catastrophic health expenditure of being enrolled into the alternative subsidised schemes. We also take into account the restrictions on the Indonesian state budget by estimating a constrained rule, which is a realistic but often ignored constraint on the policymaker's objective function in the applied literature (Bhattacharya and Dupas, 2012). Our results provide some guidance to decision makers that are considering implementing tailored policies based on expected welfare gains. In terms of methodology, we demonstrate the value in using the super learner for optimal policy learning, and highlight the potential trade off between interpretable and complex rules. We extend the functionality of the super learner to incorporate tree-based rules, in order to include more interpretable estimators of the rule within the candidate library.

3.2 The policy setting

As of mid-2019, Indonesia's ambitious JKN programme has enrolled 83% of the population since its inception in 2014 (Prabhakaran et al., 2019). The largest proportion of JKN enrollees receive PBI, whose insurance premiums are fully subsidised by the government. However, despite efforts undertaken to expand coverage, the effectiveness of the programme in protecting households across the nation from catastrophic health expenditures and impoverishment remains to be seen, as is the case for other low- and middle-countries (LMICs) implementing universal health coverage policies (El-Sayed et al., 2018; Pratiwi et al., 2021). Table 3.1 presents the number of PBI-insured household members that used insurance to fund their outpatient and inpatient health care, using data from SUSENAS 2017. Among APBD enrollees, 38% and 48% did not use insurance to pay for outpatient and inpatient care respectively. Among APBN enrollees, the respective proportions are 11% and 34%. This suggests that the benefits packages for PBI-enrollees are either insufficient to

meet health care needs, or that enrollees are knowingly or unknowingly not utilising their insurance, resulting in out-of-pocket costs for households. An analysis of a previous subsidised insurance programme (that is, Jamkesmas) reported that, although in general enrollees were positive about the scheme, some were unaware of the benefits, while others preferred to pay out-of-pocket for health care as they feared discrimination and longer waiting times (Harimurti et al., 2013). A 2019 review of the existing evidence base on the financial impacts of health insurance in low- and middle-income countries presented mixed findings (Erlangga et al., 2019b). Out of 14 studies reporting the impact of health insurance on catastrophic health expenditure, 9 found positive effects (that is, a reduction in the risk of incurring catastrophic expenditure), 3 found no statistically significant effect, and 2 found negative effects. Among the studies that explored the financial impacts of subsidised schemes, either no significant effects were found, or negative effects were reported among the insured (Bernal et al., 2017; Sparrow et al., 2013). Although increases in out-of-pocket costs are an undesirable consequence of insurance, they can also represent a greater willingness to access health care that may not have been consumed in its absence. If the newly insured value the associated health improvements, they could end up demanding better quality care that may be expensive or in short supply (Wagstaff and Lindelow, 2008). In their systematic review, Erlangga et al. (2019b) did indeed find that the effects of insurance are modified by supply-side factors captured by proximity to health facilities. Comparable evidence was presented in a recent evaluation of Indonesia's JKN, which found positive overall effects on out-of-pocket health spending for subpopulations grouped according to wealth, location and access to health care facilities, which they interacted with the treatment variable (Maulana et al., 2022).

Similar to other health financing schemes in low- and middle-income countries, the JKN has struggled to recruit and retain the large informal working sector, who previously had limited access to publicly-funded health insurance (Vilcu et al., 2016). It was expected that informal workers would self-enrol into the contributory version of JKN that is funded by income contributions, but uptake has been slow and largely dependent on health care needs (Dartanto et al., 2020a). Of those that did enrol, many are either unable or unwilling to pay, which has put at risk the financial

protection it had intended to offer beneficiaries (Muttaqien et al., 2021). When JKN was formed, the proposition was to have just one subsidised programm, APBN, that covers the entire eligible population, with enrollees identified using a specific targeting strategy (which we later describe). However, prior to JKN, many district governments designed local health financing schemes, including APBD, to address coverage gaps in the non-poor informal sector. Local governments largely depend on the central government for funding revenue but have a certain degree of autonomy in public health decision-making (Kruse et al., 2012; Sparrow et al., 2017; World Bank, 2020). Since 2021, the central government has contributed towards 10% of APBD recipients' premium contributions through the state budget, with the remainder being funded through other sources (Dutta et al., 2020). Local governments have, in some instances, been more successful than the central government in enrolling the uninsured informal sector who are eligible for PBI (Sparrow et al., 2017; World Bank, 2020). Although JKN intended to unify all previously implemented schemes, integrating Jamkesda (now APBD) into a single APBN scheme is complicated due to financial and organisational constraints (Agustina et al., 2019; National Team for the Acceleration of Poverty Reduction (Indonesia), 2015). Given APBD's success in enrolling parts of the population that are missed by APBN, it has been proposed that local governments become more involved in expanding PBI subsidies to the uninsured and informal sector through a targeted JKN enrolment strategy (Dutta et al., 2020; World Bank, 2020).

Table 3.1: Demand for inpatient and outpatient health care in 2017

	N	**Inpatient demand (N)**			**Outpatient demand (N)**		
		Accessed care	Used APBD	Used APBN	Accessed care	Used APBD	Used APBN
PBI-APBD	120,346	3,096	1,926 (62%)	-	13,271	6,857 (52%)	-
PBI-APBN	212,334	8,744	-	7,786 (89%)	29,970	-	19,841 (66%)

Note: Table reports the number of household members that used inpatient and outpatient health care, by PBI insurance scheme, and whether they used insurance to pay. Source: SUSENAS 2017.

Our study aims to address some of these challenges by demonstrating how we can develop optimal policy rules that can be used by decision makers to efficiently as-

sign different versions of PBI among the eligible population. Given the overarching aim of JKN is to insure the entire population, the decision question in this policy learning problem is whether to assign a given household to APBD versus APBN, and not whether they should be insured or not. A secondary justification for restricting the sample to PBI recipients (and excluding the uninsured) is to mitigate the effects of selection bias in our estimation of treatment effects and policy rules. We acknowledge that our study relies on observational data, and although we include a rich set of observable covariates to control for confounding, there may still be unobserved confounders that affect selection into health insurance. This issue is particularly pronounced when estimating causal effects on the uninsured versus the insured, since self-enrolment is often linked to health care need, irrespective of eligibility. By focusing on the two insured groups, the risk of selection bias is reduced since enrollees are unlikely to have self-selected their insurance type.

The optimal policy rules we aim to learn are a function of a low-dimensional vector of covariates that can represent the decision maker's selection criteria. We refer to the targeting strategy that was used to enrol the eligible population into PBI, to inform our covariate selection process. All social protection programmes in Indonesia use proxy means testing (PMT) models to identify and directly target the poorest households (Alatas et al., 2012; Vilcu et al., 2016). Social, economic and demographic information on the poorest 24.5 million households (corresponding to 96 million people) was collected in 2011 to form a unified registry, which has since been updated using 2015 census data. The PMT models use stepwise regressions to determine which assets and demographic characteristics are highly associated with household consumption, as a means to identify eligible households for PBI (National Team for the Acceleration of Poverty Reduction (Indonesia), 2015). Table 3.2 lists the targeting variables included in our policy learning model that are motivated by the selected covariates used in the PMT model specification.[2]

[2]Note that we consider a broader set of variables for confounding adjustment.

Table 3.2: Targeting variables included in our optimal policy learning model

Household member characteristics and demography
Household head marital status (1 if married, 0 otherwise)
Number of household members
Number of productive household members (aged 15-64)
Number of children at school (aged 5-14)
Socioeconomic status
Number of household members in work
Household head education level (1 if completed compulsory education, 0 otherwise)
Number of household members that completed compulsory education
Dwelling characteristics
Homeowner status (1 if homeowner, 0 otherwise)
Wall type (1 if concrete, 0 otherwise)
Roof type (1 if concrete/roof tile, 0 otherwise)
Lighting source (1 if electricity, 0 otherwise)
Drinking water source (1 if protected, 0 otherwise)
Toilet facility (1 if private/shared, 0 otherwise)
Cooking fuel (1 if electricity/gas, 0 otherwise)
Asset ownership
Refrigerator (1 if owns, 0 otherwise)
Gas canister >5.5kg (1 if owns, 0 otherwise)
Number of cellular phones in household
Car/motorcycle (1 if owns, 0 otherwise)
Gold/jewellery >10g (1 if owns, 0 otherwise)

Note: Targeting variables are based on the original variables included in the PMT model specification to determine the PBI selection criteria.

3.3 Data

SUSENAS is an annual national household socioeconomic survey that collects data on approximately 300,000 households, and its members, across 514 districts. A two-stage sampling design is used to ensure that households are representative at the district-level, and frequency weights are provided to reflect the population. The survey comprises two modules that gather information on basic household characteristics, such as occupation, education, health, housing, asset ownership and consumption expenditure. We use cross-sectional data from the 2017 survey to construct a household-level dataset, and restrict the sample to households where all members reported having the same insurance type at the time of survey (that is, either PBI-APBD or PBI-APBN). Households where at least one member reported having alternative health insurance cover, either through their employers or

private plans, are excluded from the analysis. We generate a binary indicator for our catastrophic expenditure outcome, which equals 1 if health care expenditure, as a proportion of non-food household expenditure, exceeds 10%, and 0 otherwise (Wagstaff and Lindelow, 2008; Wagstaff et al., 2007; Wagstaff and van Doorslaer, 2003; Xu et al., 2003).

We refer to the related literature on health insurance and catastrophic expenditure, and control for a rich set of variables that may influence both the selection into treatment and the outcome. Using SUSENAS, we control for head-of-household characteristics since they are representative of household well-being (Bookwalter et al., 2006). These include age, sex, marital status and socioeconomic status (which encompasses literacy, level of education, occupation and use of technology). We additionally capture broader household characteristics by controlling for the number of household members that are educated, in work, of a productive working age (between 15 and 64), or currently at school. To further capture households' socioeeconomic status, we include a selection of variables that reflect asset ownership and housing characteristics, including household amenities (e.g. types of water supply, toilet facilities and electricity), small households assets (e.g. fridge, television, computer), and other household characteristics (e.g. number of rooms, home ownership).[3]

Finally, we control for certain village-level characteristics using Potensi Desa data, PODES 2018 (conducted in 2017), which collects information on villages' availability of natural and human resources, and can be merged with SUSENAS at the district-level. We include information on whether access to all types of health care facilities (that is, primary and secondary care, community health care, maternity care and pharmacies) is considered to be easy.

See Appendix 3.A for a full list of variables included in our model for confounding adjustment.

[3]We do not control for health care need since we are unable to identify whether health status is a confounder or a mediator.

3.4 The optimal policy framework

We define our causal parameters of interest using the potential outcomes framework (Holland, 1986; Imbens and Rubin, 2015; Rubin, 1974). We construct an observational dataset of households $i = 1, \ldots, N$ that are characterised by the tuple (X_i, Y_i, D_i), where X_i is a vector of confounders and effect modifiers; $Y_i \in \{0, 1\}$ is a binary outcome variable, which equals 1 if the household's health expenditure as a proportion of total expenditure exceeds 10%; and $D_i \in \{0, 1\}$ is a binary treatment variable, which equals 0 if all household members are enrolled into the nationally funded PBI scheme, PBI-APBN (the "control"), and 1 if all household members are enrolled into the locally funded PBI scheme, PBI-APBD (the "treatment"). Individual level treatment effects are defined as $\tau_i = Y_i(1) - Y_i(0)$, where $Y_i(1)$ and $Y_i(0)$ are potential outcomes under the treatment and the control. The average treatment effect (ATE), defined as the difference in the expected potential outcomes across all households, is denoted as $\tau = E[Y_i(1) - Y_i(0)]$.

To learn a policy rule π, we first define our targeting criteria $V \in X$, where V could include all variables in X, or a subset. The policy rule maps V into a binary decision, $V \to \pi(v) \in \{0, 1\}$. We specify a policy class Π that encodes any constraints on the policymaker's objective function (for example, resource or functional form restrictions) that the policy rule must satisfy, $\pi \in \Pi$. We encode a resource constraint κ so that at most a $\kappa \in [0, 1]$ proportion of households can receive treatment (APBD). We consider two scenarios: an unconstrained environment where $\kappa = 1$, and a constrained environment where $\kappa = 0.1$, to reflect the proportion of the state health care budget that is distributed to local governments.[4] The counterfactual outcome for a household if treatment is assigned using a policy rule is denoted as $Y_i(\pi(V_i))$. Since smaller values of Y correspond to better outcomes (we want to minimise catastrophic expenditure), the optimal policy rule π^* can be defined as a minimiser of the expected counterfactual mean outcomes over all candidate rules

[4]The imposed resource constraint is mainly for demonstrative purposes to model the Minister of Finance's Regulation No. 78/PMK.02/2020, which stipulates that 10% of APBD recipients' contributions will funded by the central government through the state budget. Our model does not take into account that the remaining 90% will be funded through other sources.

$\pi \in \Pi$:

$$\pi^* \in \arg\min_{\pi \in \Pi} E[Y_i(\pi(V_i))] \text{ s.t. } E[\pi(V_i)] \leq \kappa, \tag{3.1}$$

where $E[Y_i(\pi(V_i))]$ is the policy value, our target causal parameter. The optimal policy rule can be found using the conditional average treatment effect (CATE) function, which is defined as the expected difference in the potential outcomes, as a function of the households' observed covariate profile $V_i \in X_i$:

$$\tau(v) = E[Y_i(1) - Y_i(0)|V_i = v]. \tag{3.2}$$

The CATE function can be used to determine whether there is any added value from implementing a policy rule based on treatment effect heterogeneity, as opposed to treating everyone or no one (which is a function of the sign of the ATE). One way to explore this is to construct a priority scoring function $S(\cdot)$ that ranks observations by their CATEs, and plots the Targeting Operating Characteristic (TOC) curve that compares the ATE of treating a certain fraction of observations q (prioritised by $S(X_i)$) to the overall ATE (Yadlowsky et al., 2021):

$$TOC(q : S) = E[Y_i(1) - Y_i(0)|F_s(S(X_i)) \geq 1 - q] - E[Y_i(1) - Y_i(0)], \tag{3.3}$$

where F_S is the distribution function of $S(X_i)$. If $q = 1$, the first term is the ATE.

The utility of implementing an estimated optimal CATE-based policy rule $\hat{\pi}^*$ can be measured by estimating its policy value relative to a given reference rule $\pi^r(V_i)$:

$$U(\hat{\pi}^*) = E[Y_i(\hat{\pi}^*(V_i)) - Y_i(\pi^r(V_i))]. \tag{3.4}$$

Reference rules could include static rules that assign either APBD or APBN to all households. The corresponding utilitarian regret, defined as the difference between the expected utility from the learned policy versus the best policy in the class Π (this is the oracle policy, denoted π'), is given by:

$$R(\pi) = \max_\pi\{U(\pi') : \pi' \in \Pi\} - U(\hat{\pi}^*), \tag{3.5}$$

which we aim to minimise. The regret tends to be positive since policymakers are

unable to implement the oracle policy π' given finite samples and policy constraints.

Identifying our target causal parameter requires the following assumptions on the data generating process:

(a) Unconfoundedness: $Y_i(d) \perp D_i | X_i = x, \forall d \in D, \forall x \in \mathcal{X}$

(b) Overlap: $0 < e(x) \equiv P[D_i = 1 | X_i = x] < 1, \forall x \in (X)$

(c) Stable Unit Treatment Value Assumption (SUTVA): $Y_i = Y_i(D_i)$

We refer to $e(x)$ as the propensity score. If the unconfoundedness assumption is satisfied, the conditional expectation of the potential outcomes corresponds with the conditional expectation of the observed outcomes: $E[Y(d)|X_i = x] = E[Y_i|X_i = x, D_i = d]$. Therefore, $\tau(x)$ can be identified as a function of the observed outcomes:

$$\tau(x) = E[Y_i|X_i = x, D_i = 1] - E[Y_i|X_i = x, D_i = 0] \tag{3.6}$$

$$= \mu_1(x) - \mu_0(x), \tag{3.7}$$

where $\mu_d(x)$ is the counterfactual response surface.

3.5 Learning optimal policy rules

Learning optimal policies requires making some important decisions about the restrictions on the policy class Π. Policymakers may want to impose restrictions on the functional form (decision rules may need to be simple enough to implement and evaluate), budget (there may be a finite set of resources), and fairness (vulnerable populations may require special targeting measures (Kitagawa and Tetenov, 2018; Zhou et al., 2022). In many cases, Π may only include policies that are interpretable and easy to implement, or meet some other normative criteria, such as targeting a set of pre-defined characteristics. Learning $\pi^* \in \Pi$ involves three components: (i) assigning "scores" to each observation according to their predicted CATEs; (ii) finding the empirical counterpart of (3.1) that estimates the policy value of a given rule;

and (iii) selecting an optimiser that searches for the rule, among all candidate rules, that minimises the objective function.

A proposed solution is based on the the well-known augmented inverse probability of treatment weighted (AIPTW) estimator (also known as the doubly robust estimator) for the ATE, $\hat{\tau} = \frac{1}{N}\sum_{i=1}^{N}\hat{\Gamma}_i(X_i)$, of which the main component is the estimated doubly robust scores, $\hat{\Gamma}_i = \hat{\Gamma}_i(1) - \hat{\Gamma}_i(0)$ (Robins et al., 1994):

$$\hat{\Gamma}_i = \underbrace{\hat{\mu}_1(X_i) + \frac{D_i}{\hat{e}(X_i)}(Y_i - \hat{\mu}_1(X_i))}_{\hat{\Gamma}_i(1)} - \underbrace{\hat{\mu}_0(X_i) + \frac{1-D_i}{1-\hat{e}(X_i)}(Y_i - \hat{\mu}_0(X_i))}_{\hat{\Gamma}_i(0)} \tag{3.8}$$

$$= \hat{\mu}_1(X_i) - \hat{\mu}_0(X_i) + \frac{D_i - \hat{e}(X_i)}{\hat{e}(X_i)(1-\hat{e}(X_i))}(Y_i - \hat{\mu}_D(X_i)), \tag{3.9}$$

where $\hat{\mu}_D(X_i)$ is the conditional expectation function for the observed outcome under the treatment actually received (that is, the outcome regression). The construction of the doubly robust scores in (3.9) requires estimates of $e(x)$, $\mu_1(x)$, $\mu_0(x)$ and $\mu_D(x)$, which are jointly referred to as the nuisance parameters.

If the doubly robust scores are constructed in a certain way (that is, the nuisance parameters are estimated using machine learning methods and K-fold cross-fitting[5]), they can be used to generate estimates of the target causal parameter that have better finite-sample performance (than non-doubly robust estimators), are robust to bias and overfitting, and have good asymptotic properties (Chernozhukov et al., 2018a; Newey and Robins, 2018; Zhou et al., 2022). These doubly robust scores can be used (beyond using them to estimate the ATE) to evaluate the value of a given policy rule, and to select the optimal policy rule given a policy class. The empirical solution to (3.1) minimises the counterfactual mean outcome under the rule using the doubly robust scores:

$$\hat{\pi}^* \in \arg\min_{\pi\in\Pi} \frac{1}{n}\sum_{i=1}^{n}\hat{\Gamma}_i(\pi(V_i)) \text{ s.t. } E[\pi(V_i)] < \kappa, \quad \hat{\Gamma}_i(\pi(V_i)) = \pi(V_i)\hat{\Gamma}_i(1) + (1-\pi(V_i))\hat{\Gamma}_i(0). \tag{3.10}$$

[5]Cross-fitting is a form of efficient data-splitting used to reduce overfitting, where the training data is divided into K folds and the prediction for a data point in fold k is made using the other $K-k$ folds. The roles of each fold are swapped K times, and the resulting estimates for each fold k are averaged across the K folds.

3.5.1 Threshold-based rules

Unrestricted policies could be learned easily if we had access to the true CATE function $\tau(v)$. In an unconstrained setting, policy assignment for a given unit could be based on a simple sign rule:

$$\pi^* = \mathbb{I}\{\tau(v) < 0\}. \tag{3.11}$$

Put simply, the sign rule indicates that treatment should only be assigned when it is effective (in our policy learning problem, that means a negative CATE). In recent years, a number of flexible estimators for $\tau(v)$ have been proposed, particularly using non-parametric models and machine learning tools (examples include Athey et al. (2019b); Hahn et al. (2020); Kennedy (2020); Künzel et al. (2019a); Nie and Wager (2021); Shalit et al. (2017)). In particular, van der Laan and Luedtke (2014) and Luedtke and van der Laan (2016b) propose the two-stage doubly robust estimator, where the first stage requires constructing the doubly robust scores $\hat{\Gamma}_i$ as in (3.9), and the second stage involves regressing $\hat{\Gamma}_i$ on $V_i \in X_i$ using any regression-based approach. The first stage takes care of confounding adjustment by controlling for the full covariate vector X_i in the outcome and treatment models, so that the second stage CATE estimation can focus on the targeting criteria $V_i \in X_i$. The same authors additionally propose the super learner to select the best estimator among candidate CATE estimators using cross-validation. For example, candidate estimators could include simple linear models with no interactions, or more complex models that do not enforce linear relationships (for example, regression splines and random forests). More recently, Kennedy (2020) extends this work by incorporating sample splitting into the two-stage procedure, which firstly allows for CATE-specific error bounds that rely only on stability conditions for the second stage estimator, and secondly reduces bias that can arise from repeated use of the same data for different estimation tasks. Other versions of two-stage doubly robust estimators have been proposed in the literature, but the second stage estimators tend not to be model agnostic (Fan and Wu, 2020; Semenova and Chernozhukov, 2021; Zimmert and Lechner, 2019).

An advantage of threshold-based rules is their ability to incorporate resource constraints, where there is a maximum proportion of the population that can be assigned to treatment (Luedtke and van der Laan, 2016a). This constrained minimisation problem can be solved by sorting the population according to their estimated CATEs $\hat{\tau}(V_i)$ in increasing order, and assigning treatment to those with the lowest estimates until the constraint is met. See Appendix 3.C for further details of this constrained optimisation problem.

3.5.2 Tree-based rules

Athey and Wager (2021) and Zhou et al. (2022) propose a policy class of fixed-depth decision trees that directly optimise net policy benefits according to the expected individual treatment effects. The splits of the tree use the covariate values to guide units into a specific leaf, which is associated with a policy decision. Essentially, the optimiser takes $\hat{\Gamma}_i$ and V_i as inputs, and searches through the space of all candidate trees to identify the one that solves the minimisation problem in (3.10). The outputs can be represented visually using a tree-like structure that shows the splitting covariates and their respective cut-off thresholds, which are used to classify households into treatment and control. Athey and Wager (2021) advise fitting shallow trees (that is, an interaction depth of 2 or 3) to prevent overfitting and for computational efficiency, although cross-validation can also be used to choose the optimal depth. Further details on the tree-search algorithm can be found in Sverdrup et al. (2020) and Zhou et al. (2022).

To construct $\hat{\Gamma}_i$, Athey and Wager (2021) propose a modified version of (3.9) that replaces the first component (that is, the difference between the response functions) with a causal forest estimate of the CATE function $\hat{\tau}^{cf}(v)$:

$$\hat{\Gamma}_i^{cf} = \hat{\tau}^{cf}(V_i) + \frac{D_i - \hat{e}(X_i)}{\hat{e}(X_i)(1-\hat{e}(X_i))}(Y_i - \hat{m}(X_i) - (D_i - \hat{e}(X_i)\hat{\tau}^{cf}(V_i)), \tag{3.12}$$

where $\hat{m}(X_i) = E[Y_i|X_i = x]$.[6] In this construction, the nuisance estimates adjust

[6]Note that $\hat{m}(X_i)$ is different to the outcome regression $\hat{\mu}_D(X_i)$ in (3.9) as it is marginalised over the treatment D_i.

for confounding by controlling for X_i, while the CATE estimates only consider V_i. Causal forests are a causal adaptation of the random forest prediction algorithm. They find neighbourhoods of observations with similar CATEs by regressing the Y-residual $Y_i - m(X_i)$ on the D-residual $D_i - e(X_i)$, and recursively partitioning the data into leaves to maximise the within-leaf heterogeneity in treatment effects, thus forming a causal tree. Each observation is dropped down the tree and assigned a weight based on how frequently it is used to estimate $\hat{\tau}(V_i)$ at $V_i = v$. This process is repeated across many causal trees using bootstrapped samples, to improve prediction accuracy and to limit noise stemming from individual trees. Each tree in the forest is "honest", meaning that the bootstrapped samples are split into two, to separate the data used to construct the tree and to make predictions, thus preventing overfitting. The CATE estimator is then constructed as:

$$\hat{\tau}^{cf}(v) = \frac{\sum_{i=1}^{N} w_i(v)(D_i - \hat{e}(X_i))(Y_i - \hat{m}(X_i))}{\sum_{i=1}^{N} w_i(v)(D_i - \hat{e}(X_i))^2}, \tag{3.13}$$

where $w_i(v)$ are the weights derived from the forest splitting on the vector-valued gradient of $\tau^{cf}(v)$.[7] In theory, the doubly robust scores $\hat{\Gamma}_i^{cf}$ in (3.12) could be regressed onto V_i to estimate CATEs as per the method described in the previous section. However, since the causal forest directly produces CATE estimates $\tau^{cf}(V_i)$, the additional regression step is not required.

A limitation of the causal forest algorithm is its black-box nature, which means that the estimated CATE function is non-parametric, with no straightforward interpretation. Moreover, the estimator uses all V_i that is specified by the researcher, while it is possible that an even smaller set of covariates contribute to heterogeneity, and are therefore useful for building the policy rule. The super learner, introduced in the next section, can consider models of varying complexity in a single framework, and enable the selection of the best performing rule in a data-adaptive way.

[7]The CATEs are estimated "out-of-bag", meaning that for each observation dropped down the tree, a prediction is made using trees that did not use this observation during the training process. Out-of-bag prediction produce CATE estimates $\hat{\tau}^{cf}(V_i)$ without the need for explicit data splitting techniques (Athey et al., 2019b).

3.5.3 Super learning the optimal policy rule

The super learner was originally proposed by van der Laan et al. (2007) as a data-driven method for selecting or combining prediction algorithms, among a library of candidate algorithms, using K-fold cross-validation. The implementation of the super learner requires the researcher to define the following: the candidate library, the method for combining candidate estimates (known as the "meta-learner") and the loss function to evaluate candidate performance. The super learner has become a popular prediction tool and more recently has been adapted for policy learning (Luedtke and van der Laan, 2016b). One of its favourable properties is that, in large enough samples, it performs at least as well as the best performing candidate in the library (van der Laan et al., 2007).

In the policy learning setting, the objective of the super learner is to improve the estimation of the policy rule by constructing a meta-learner that, from a library of candidate estimators, either chooses the best performing one (that is, the "discrete" super learner) or chooses the optimal weighted convex combination of candidates (that is, the "continuous super learner). We consider the scenario where we select the continuous super learner. Similar to the prediction task, the implementation requires the user to choose the candidate library, the meta-learner and the loss function. We explain each of these choices in more detail.

The candidate library could include threshold-based rules (that are implied by CATEs, estimated using any parametric or non-parametric method), tree-based rules (that implement decision trees based on doubly robust CATEs, typically estimated using causal forests), and static rules that either treat all or treat none, irrespective of the observed covariate profile.[8]

The meta-learner is used to find weighted convex combinations of candidate estimators of the CATEs or candidate estimators of the policy rules.[9] These weighted convex combinations are estimators of $\tau_\alpha(V_i)$ and $\pi_\alpha(V_i)$ respectively (where α is the

[8]Other policy learning methods that we do not consider (for example, outcome weighted classification, regression-based approaches) could also be included in the candidate library.

[9]Convex combinations are found using the simplex method, which represents a linear programming problem as a system of linear equations, and defines an algorithm for finding the solution to this system of linear equations.

weight vector associated with a given convex combination), denoted by:

$$\tau_\alpha(V_i) = \sum_j \alpha_j \tau(V_i), \qquad \alpha_j \geq 0 \, \forall \, j, \sum_j \alpha_j = 1 \tag{3.14}$$

$$\pi_\alpha(V_i) = \mathbb{I}[\sum_j \alpha_j \pi_j(V_i) > 0.5], \qquad \alpha_j \geq 0 \, \forall \, j, \sum_j \alpha_j = 1, \tag{3.15}$$

where the weights $\alpha_1, \ldots, \alpha_J$ are non-negative and sum to one.[10] Note from (3.15) that a given convex combination of candidate policy rules is made using a weighted majority vote, meaning that if the weighted average of candidate rules is greater than 0.5, $\pi_\alpha(V_i)$ is equal to one; and zero otherwise. For candidate libraries that only include threshold-based rules, either type of meta-learner can be selected. For candidate libraries that include rules other than threshold-based rules (for example, tree-based or static), only the type of meta-learner that optimally combines candidate policy rules (as in (3.15)) can be selected. When learning constrained rules that restrict the proportion of units that can be treated, only the type of meta-learner that optimally combines candidate CATEs (as in (3.14)) can be selected.[11]

The loss function L evaluates the performance of the convex combination of candidate CATEs or policy rules, using, for example, the counterfactual mean outcome under the rule $L_{E[Y_i(\pi(V_i))]}$. This quantity can be estimated using a suitable estimator ψ, such as the AIPTW estimator ψ^{AIPTW} as in (3.10), or the targeted minimum loss-based estimator (TMLE) ψ^{TMLE}, which aims to "target" or debias the target parameter (for example, ψ^{AIPTW}) by updating the initial estimate of the outcome regression predictions under the rule $\mu_{\pi(V_i)}(X_i)$ (van der Laan and Luedtke, 2015).[12] In the meta-learning step of the super learner procedure, the algorithm finds the weighted convex combination of candidate estimators of the CATEs or the policy rules that minimises L.

[10] The discrete super learner that selects only one candidate algorithm also uses weighted convex combinations of the algorithms, although the candidate-specific weight α_j can only be 0 or 1.

[11] For constrained rules, the candidate library can only include threshold-based rules because the units must be ranked according to their estimated CATEs.

[12] TMLE is a doubly robust estimator that updates an initial estimator of the target parameter in a targeted way to optimise the bias-variance trade off. In the context of policy learning, this involves estimating a "clever" covariate $H_i = \frac{\mathcal{I}[D_i = \pi(V_i)]}{e(X_i)}$, and updating $\hat{\mu}_{\pi(V_i)}(X_i)$ by fitting a generalised linear model that regresses Y_i on an offset term $\text{logit}\{\hat{\mu}_{\pi(V_i)}(X_i)\}$ and H_i, with the intercept estimated using maximum likelihood estimation. See van der Laan and Rose (2018) for more details on TMLE.

3.6 Implementation of methods

In this section, we explain our procedure for estimating and evaluating optimal policy rules by applying the methods described in section 3.5 to our PBI assignment problem. Our procedure is largely based on the work of van der Laan and Luedtke (2015) and Luedtke and van der Laan (2016b), who develop the theory behind the super learner framework for policy learning, and Montoya et al. (2021) and Montoya et al. (2022), who provide the implementation and interpretation. We refer to Phillips et al. (2022) for guidance on how to specify the super learners, including defining the K-fold cross-validation scheme and selecting the candidate library of algorithms. For the majority of our analyses, we rely on the following software packages in `R`, and modify code where necessary to suit our implementation: `SL.ODTR`[13,14], `SuperLearner`[15], `grf`[16] and `policytree`[17].

3.6.1 The workflow

We describe our workflow in the following steps (see Figure 3.1 for an accompanying visual diagram):

1. We select $V \in X$ as the list of targeting variables included in Table 3.2. As previously discussed, these variables are representative of decision makers' selection criteria for assigning subsidised health insurance (PBI), and we assume that these are also relevant for our policy learning problem. We denote this vector of covariates as $V1$. We also consider the fact that the optimal policy rule may belong to a simpler policy class, where the true CATE function depends on an even more restricted subset of covariates. Screening methods, such as regularisation, are a useful pre-processing step for variable selection

[13]https://github.com/lmmontoya/SL.ODTR

[14]We extend the functionality of `SL.ODTR` to incorporate tree-based rules into the candidate library and to enable minimisation problems. We also modify the cross-validated loss function from ψ^{TMLE} to ψ^{AIPTW}.

[15]https://github.com/ecpolley/SuperLearner

[16]https://github.com/grf-labs/grf

[17]https://github.com/grf-labs/policytree

(Wang and Barbu, 2019). We construct a second covariate vector $V2 \in V1$ that corresponds to the covariates with non-zero coefficients from a doubly robust adaptive LASSO regression (Bahamyirou et al., 2022).[18] We divide the full data into two parts with a 30:70 split, using the smaller partition (≈28,000 households) to learn $V2$, and the larger partition (≈65,000 households) to learn policy rules.

We implement the adaptive LASSO on the smaller 30% data partition using the following steps:

- We fit super learners to estimate the nuisance parameters – $e(X_i)$, $\mu_1(X_i)$, $\mu_0(X_i)$ and $\mu_D(X_i)$ – and construct $\hat{\Gamma}_i$ by plugging in the nuisance predictions (see step 2 for the candidate algorithms included in the super learners).
- We regress $\hat{\Gamma}_i$ on $V1_i$ to obtain $\tilde{\beta}_l$, the estimated coefficient of $V1_{(l)}$ for $l = 1, \ldots, p$.
- We construct coefficient-specific weights $\hat{w}_l = \frac{1}{|\beta_l|^\gamma}$ (we set $\gamma = 1$) so that the regularisation penalises more those coefficients with lower estimates in the initial linear regression.
- We fit a LASSO regression of $\hat{\Gamma}_i$ on $V1_i$ (again on the full 30% data partition) using $\hat{w}_l$ as the penalty factor associated with each coefficient [19]. The tuning parameter λ is selected using cross-validation. The final selection of covariates in $V2$ are those with non-zero coefficients.

We do not use this 30% data partition again and from here on, we refer to the remaining 70% data partition as the full data.

2. We specify J candidate algorithms for estimating the optimal policy rule $\pi^*(V_i)$ for $j = 1, \ldots, J$. Our first set of candidate algorithms are threshold-based rules that rely on various regression specifications of $\hat{\Gamma}_i$ on V_i to generate estimated

[18] The adaptive LASSO is a regularisation method based on the traditional LASSO. The algorithm has the oracle property, in that it consistently selects the right subset of variables, and has an optimal estimation rate. It also uses coefficient-specific weights in the regularisation so that true-zero coefficients are less likely to be selected than in the traditional LASSO. See Zou (2006) for further details.

[19] We use the same 30% data partition to perform both regressions in the adaptive LASSO, but in theory, the data could be further partitioned to separate the data used for both tasks

CATEs for each household $\hat{\tau}(V_i)$. Candidate regressions include generalised linear models, generalised boosted models (with an interaction depth of 2), multivariate adaptive polynomial spline regression, neural networks and support vector machines. Including a diverse range of candidate regressions in our library, from simple parametric models to more flexible, data-adaptive algorithms, hedges against the possibility that the optimal policy rule could be best modelled by a simple (rather than a more complex) estimator. Simpler models can reduce the risk of overfitting in finite samples. We also include causal forest CATE estimates $\hat{\tau}^{cf}(V_i)$ in (3.13) as a candidate estimator in our threshold-based library. Our second set of candidate algorithms are tree-based rules that rely on $\hat{\tau}^{cf}(V_i)$. We include shallow decision trees (depths 1, 2 and 3). Our third set of candidate algorithms are static rules that assign either APBD or APBN to all households, irrespective of their covariate profile. See Table 3.3 for a full list of candidate algorithms included in our library. Note that all candidate tree-based and threshold-based estimators are separately fitted on $V1$ and $V2$.

Table 3.3: Candidate estimators included in the super learner library

Estimator	Description	Inputs
Threshold-based rules		
GLM	Generalised linear model	$\hat{\Gamma}_i, V_i$
GLMi	Generalised linear model with interactions	$\hat{\Gamma}_i, V_i$
GBM	Generalised boosted model (depth 2)	$\hat{\Gamma}_i, V_i$
PM	Multivariate adaptive polynomial spline regression	$\hat{\Gamma}_i, V_i$
NN	Neural network	$\hat{\Gamma}_i, V_i$
SVM	Support vector machines	$\hat{\Gamma}_i, V_i$
CF	Causal forest	$\hat{\tau}^{cf}(V_i)$
Tree-based rules		
PT1	Policy tree (depth 1)	$\hat{\Gamma}_i^{cf}, V_i$
PT2	Policy tree (depth 2)	$\hat{\Gamma}_i^{cf}, V_i$
PT3	Policy tree (depth 3)	$\hat{\Gamma}_i^{cf}, V_i$
Static rules		
Treat all	Assign PBI-APBD to all households	-
Treat none	Assign PBI-APBN to all households	-

Note: Inputs refer to the parameters required to learn the policy rule. All threshold- and tree-based rules are separately fitted on $V1$ and $V2$. The super learner library for the constrained policy rule only includes threshold-based rules.

3. To ensure that there is sufficient overlap between covariate distributions in the APBD and APBN populations, we follow the methodology that proposes to trim the sample by removing households with extreme propensity scores near 1 or 0 – that is, almost always assigned to APBD or never assigned to APBD, irrespective of observed characteristics (Stürmer et al., 2021). Extreme scores can affect inverse probability of treatment weights, thus creating bias and excessive variance in the treatment effect estimators (Li et al., 2019). Trimming the sample effectively changes the causal estimand to the expected policy value for households with sufficient overlap.[20] We fit a super learner on the full data, and following Crump et al. (2009), we remove approximately 4,000 households with $\hat{e}(X_i)$ outside of the range [0.1,0.9]. We refer to the remaining data as the trimmed data.

4. We divide the trimmed data randomly into two equal subsamples, denoted s_1 and s_2; s_1 is used to train the candidate models for estimating the policy rule, and s_2 is used to test the models and make predictions on new, unseen data. Sample splitting prevents the same data being used for training and testing, which limits the risk of overfitting. To account for a potential loss in efficiency from subsetting the data, we use 2-fold cross-fitting that swaps the roles of s_1 and s_2 and recreates the trimmed dataset by pooling the test predictions of the nuisance models, CATEs and policy rules. In the following steps, we describe our procedure on the iteration where s_1 and s_2 are the respective training and testing data.

5. We perform K-fold cross-validation by splitting s_1 into K folds (we choose $K = 2$ based on the effective sample size, as per Phillips et al. (2022)), where each fold k serves as the validation data and the remaining $K - k$ folds (in this case, just 1) serve as the training data, on which we fit each of our J candidate estimators. For each fold k, we perform the following on the training data:

    - We estimate the nuisance parameters – $e(X_i)$, $\mu_1(X_i)$, $\mu_0(X_i)$, $\mu_D(X_i)$ and

[20] The households to whom this condition applies will be more similar to each other in terms of observed characteristics, but less similar to households for whom treatment assignment is more certain.

$m(X_i)$[21] – using the super learner and construct doubly robust scores.[22,23] Our candidate library for the nuisance models include a generalised linear model, LASSO regression and a generalised boosted model (with an interaction depth of 2).

- We train each of our J candidate estimators using the respective inputs detailed in Table 3.3.

Using the validation fold k, and for each candidate algorithm j, we use the trained models to predict CATEs $\tau_j^k(V_i)$ and policy rules $\pi_j^k(V_i) = \mathbb{I}\{\tau_j^k(V_i) < 0\}$ for each household $i \in n$.

6. We implement the continuous super learner, which uses weighted convex combinations of the predicted candidate-specific CATEs (as in (3.14)) or the predicted candidate-specific policy rules (as in (3.15)) from the previous step. For the unconstrained policy rule, where $\kappa = 1$, we use the second approach (since our candidate library includes non-threshold-based estimators), and for the constrained policy rule, where $\kappa = 0.1$, we use the first approach (since our candidate library can only include threshold-based estimators).

   We select the AIPTW estimator ψ^{AIPTW} as the loss function for estimating the counterfactual mean outcome under the rule $L_{E[Y(\pi_j(V_i))]}$. The super learner obtains estimates of the loss for each candidate convex combination of algorithms within each validation fold k, and averages them across all folds K to produce a single estimated loss for each candidate convex combination of algorithms. The candidate convex combination with the lowest cross-validated loss is selected as the optimal candidate weighting α^*. This is the meta-learning step.

7. We fit each candidate estimator of the CATE $\tau_j(V_i)$ and the optimal policy rule $\pi_j(V_i)$ on the full training data s_1, and generate predictions $\hat{\tau}_j(V_i)$ and $\hat{\pi}_j(V_i)$ on the testing data s_2. Using the optimal weights α^* from the previous step, we combine the candidate predictions to yield the super learner estimate

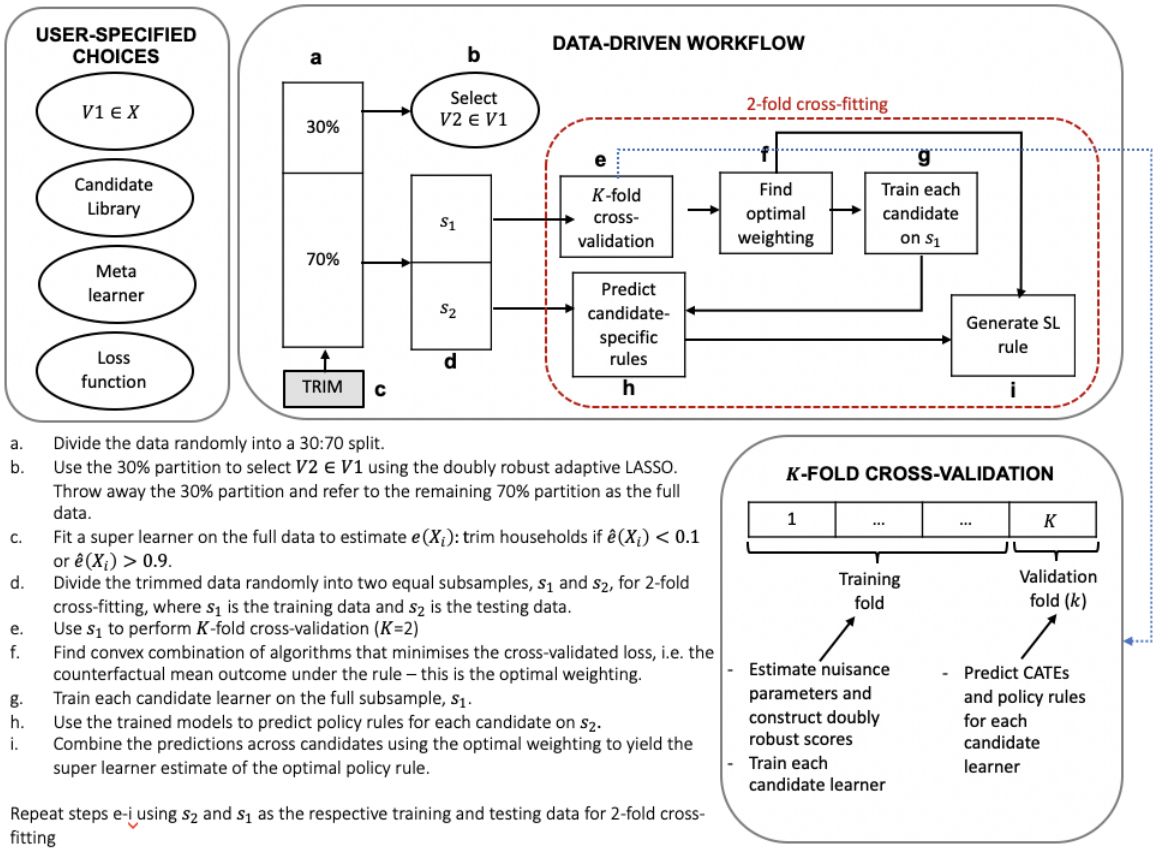
[21] We require predictions of $m(X_i)$ to estimate $\hat{\tau}^{cf}$ and to construct $\hat{\Gamma}_i^{cf}$.

[22] The doubly robust scores include both $\hat{\Gamma}_i$ and $\hat{\Gamma}_i^{cf}$

[23] Fitting and predicting both the nuisance parameters and the doubly robust scores on the same training data is a form of "nested" cross-validation (Coyle, 2017).

of the optimal policy rule, where $\hat{\pi}^*_{\tau(V_i)}(V_i) = \mathbb{I}[\tau_{\alpha^*}(V_i) > 0]$ (that is, the rule based on the sign of the optimal weighted convex combination of candidate CATEs) and $\hat{\pi}^*_{\pi(V_i)}(V_i) = \hat{\pi}_{\alpha^*}(V_i)$ (that is, the rule that corresponds to the optimal weighted convex combination of candidate rules).

Figure 3.1: Our policy learning workflow, focusing on the data-driven components.



3.6.2 Evaluating policy rules

We first analyse the estimated CATEs from our threshold-based candidates, to determine whether there is any value in learning rules that exploit treatment effect heterogeneity in assigning APBD over APBN. We plot TOC curves, as in (3.3), where $S_j(X_i)$ are candidate-specific prioritisation scores that are sorted (in decreasing order) on $\hat{\tau}_j(X_i)$, and q is the fraction of the population that is assigned to APBD

(ranging from 0.1 to 1 in 0.1 intervals) (Yadlowsky et al., 2021). We also separately plot sorted Group Average Treatment Effects (GATEs) that stratify $\hat{\tau}_j(X_i)$ into quintiles, and estimate $\hat{\tau}_j$ for each quintile. Significant differences in the estimates between the bottom quintile and the higher quintiles provide additional evidence of heterogeneity (Chernozhukov et al., 2018b).

If we are able to justify that there is some benefit to learning rules that go beyond the ATE and leverage upon CATEs, we can then evaluate the super learner and candidate estimates of the policy rule. First, we compare the counterfactual mean outcomes of the estimated optimal policy rules to the static rules and the actual policy assignment, to determine whether personalised policy rules outperform the status quo. Then, we investigate the performance of each candidate estimator and its relative contribution to the super learner, to identify which types of underlying models the policy rules are based on. Since we use the AIPTW estimator ψ^{AIPTW} as the loss function in the meta-learning step (see step 6 in the workflow), we make it our primary estimator of the counterfactual mean outcome. For robustness, we also use the TMLE ψ^{TMLE}. Second, we identify the natural subgroups formed from the super learner estimate of the policy rule by characterising households that are counterfactually assigned to APBD and APBN, and comparing them to households under the actual policy assignment. Third, we plot visual representations of the tree-based rules that are made up of splitting nodes (that is, the covariates to split on and their associated splitting values) and leaf nodes (that is, the policy decisions), for additional subgroup analyses.

3.7 Results

3.7.1 Descriptive statistics

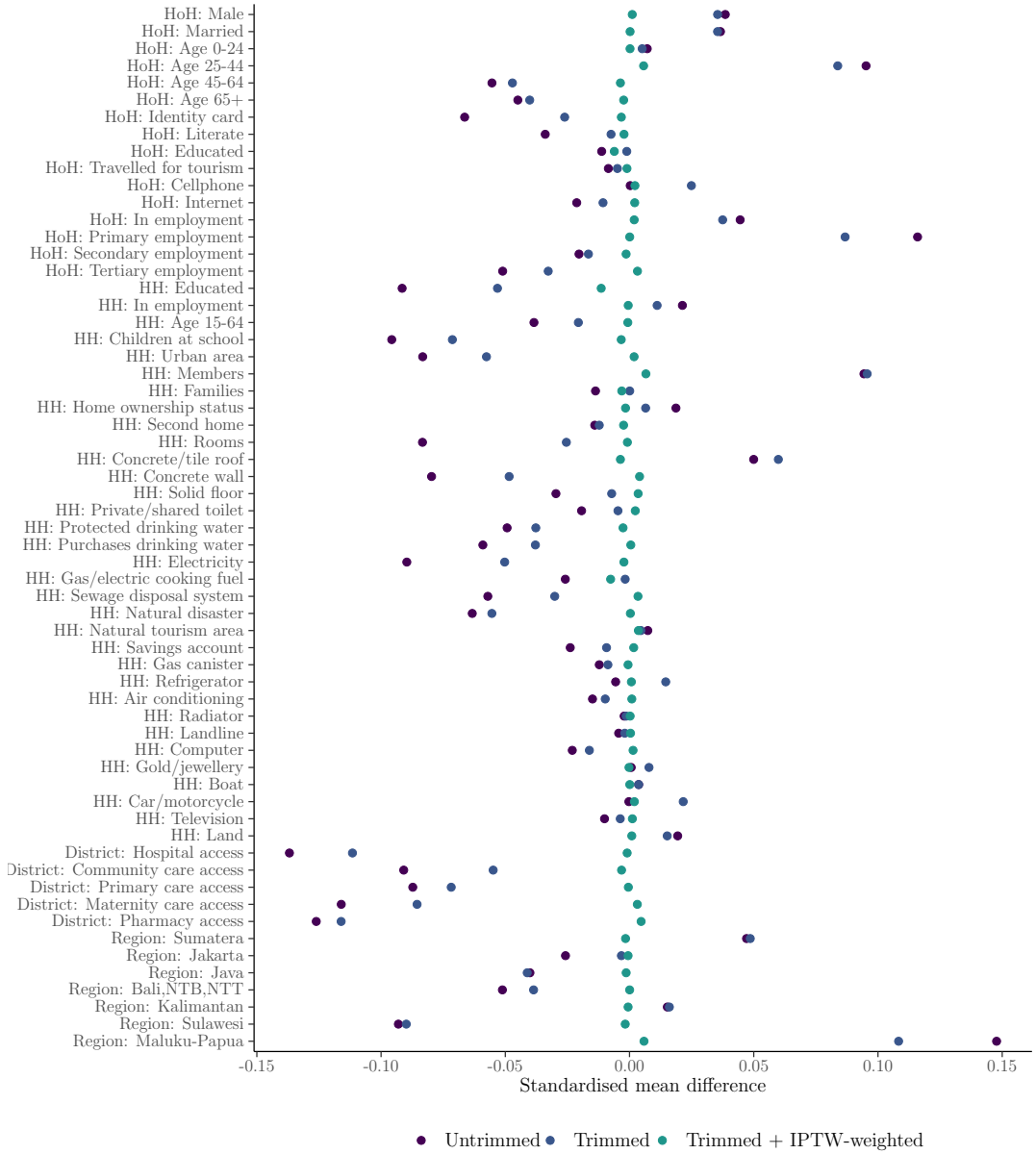
Table 3.4 compares the average characteristics of households that are assigned to PBI-APBD and PBI-APBN under the actual policy assignment D_i for a selection of observed covariates from X_i. Although the two samples are largely comparable, we find some small differences. Compared to their APBN counterparts, the heads

of APBD households are more likely to be in the 25-44 age bracket, married and in employment, particularly in the primary sector. On average, their households tend to be larger in size (average of 3.7 members compared to 3.5 for APBN), although with fewer members that are educated. There are also fewer members that use the internet or have cellular phones, and the household is slightly less likely to have basic household amenities such as electricity and drinking water. Both samples report having easy access to all types of health care facilities, although accessibility is significantly better for APBN households, especially for hospitals. On average, community care centres are the most accessible health facility across all households, and hospitals are the least accessible. Geographically, the average household enrolled into PBI is rurally based. This is especially true for APBD households, who, according to the data, are more likely than APBN households to live in Sumatera, Kalimantan and Maluku-Papua, which represent a few of the regions in the country with high percentages of rural populations (Mardiansjah et al., 2021). Figure 3.2 displays balance statistics on the full covariate vector X, showing that after reweighting the trimmed sample using the inverse of the propensity score, all covariate means are balanced (absolute SMD < 0.1) across the two subpopulations. For any small remaining imbalances, we address these through our doubly robust approach.

3.7.2 Exploring treatment effect heterogeneity

Figure 3.3 plots TOC curves showing estimated improvements in the ATE (compared to the overall ATE) from assigning APBD to increasing fractions of the population, ranked according to their candidate-specific CATE estimates. The area above the TOC curve provides some evidence of treatment effect heterogeneity since the ATEs for highly ranked subpopulations outperform the overall ATEs. We find that, in general, the CATE models that consider $V1$ display more heterogeneity than those that consider $V2$. In both instances, generalised boosted models find very little variation in treatment effects, whereas neural networks and support vector machines find more heterogeneity. The notable differences in effect estimates across the candidates algorithms highlight the potential limitations of relying on a single

Figure 3.2: Covariate balance between households assigned to PBI-APBD and PBI-APBN.

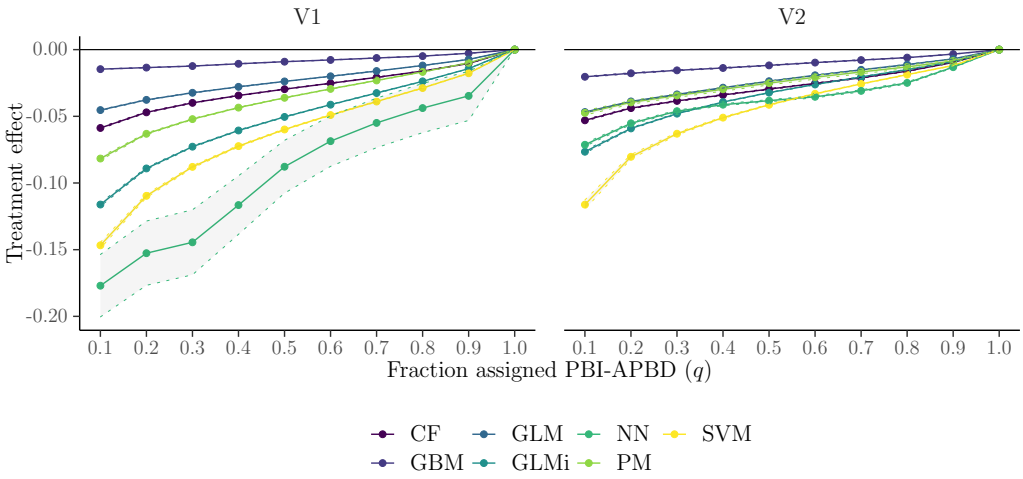


Note: Standardised mean differences between the "treated" (PBI-APBD) and the "controls" (PBI-APBN) are reported in the unweighted untrimmed sample, unweighted trimmed sample, and weighted trimmed sample (using inverse probability of treatment weights for the ATE). HoH = head of household. HH = household. Covariate labels have been abbreviated - see Appendix 3.A for further details on the included covariates.

CATE-based candidate estimator of the policy rule, as opposed to an ensemble.

Figure 3.4 plots sorted GATEs for quintiles of CATEs estimated by each threshold-based candidate estimator. Differences in sorted GATEs between the lowest quintile (Q1) and the higher quintiles (Q2-Q5) are reported in Table 3.B.1. Our findings

Figure 3.3: Targeting Operator Characteristic (TOC) curve



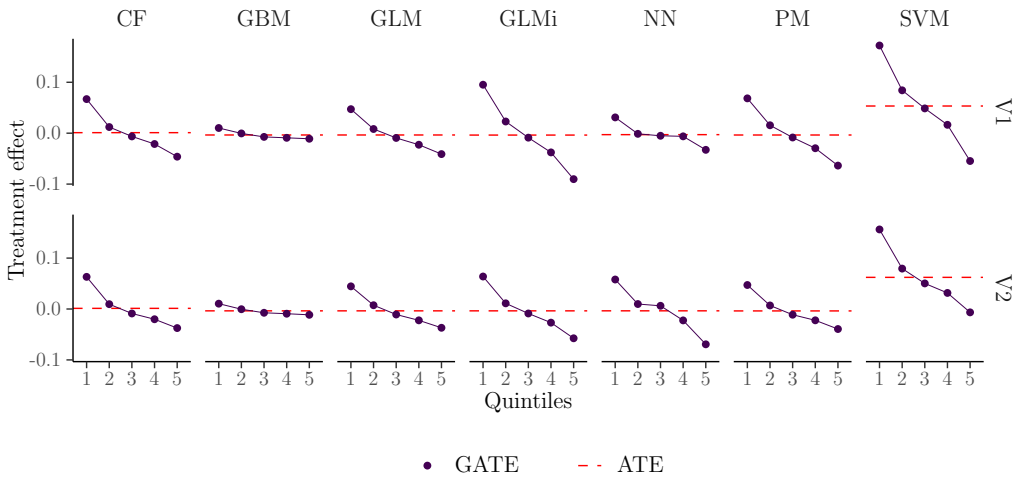
Note: TOC curve plots the cumulative estimated ATEs on catastrophic expenditure from assigning PBI-APBD (compared to PBI-APBN) to increasing fractions of the population, ranked by prioritisation scores $S(X_i)$. The scores represent the estimated CATEs from threshold-based candidate estimators of the policy rule. Separate TOC curves are displayed for CATE estimators that consider $V1$ and $V2$.

support those from Figure 3.3, in that there is evidence of heterogeneity in treatment effects. An evaluation that focuses solely on overall ATEs would conclude that there are limited impacts of assigning APBD over APBN, and the evident variation in impacts would be missed. The sorted GATEs tell us that households in the higher quintiles (predominantly Q4 and Q5) have a lower risk of suffering catastrophic health expenditures from receiving APBD over APBN, while those in the lower quintiles (Q1 and Q2) are at a higher risk. In summary, we think that our heterogeneity analysis provides justification for learning CATE-based policy rules.

3.7.3 Variable selection

Figure 3.B.2 plots the coefficients for $V1$, as a function of the $\log(\lambda)$ values used in the adaptive LASSO model. We select the covariates with non-zero coefficients that are associated with the value of $\log(\lambda)$ that minimises the cross-validated mean squared error. Our model selects 10 (out of 19) household-level variables relating to demographics, basic amenities and asset ownership – the number of members aged between 15 and 64, the number of educated members, the number of cellular phones,

Figure 3.4: Estimated sorted GATEs from threshold-based candidate estimators of the CATE



Note: Sorted GATEs are the estimated ATEs for each quintile of the population, ranked (in descending order) by predicted CATEs. The bottom quintile (Q1) denotes the least affected population subgroup (i.e. households with the smallest CATEs), and the top quintile (Q5) denotes the most affected subgroup (i.e. households with the largest CATEs). The red dashed line is the candidate-specific ATE estimate. Separate plots are displayed for CATE estimators that consider $V1$ and $V2$.

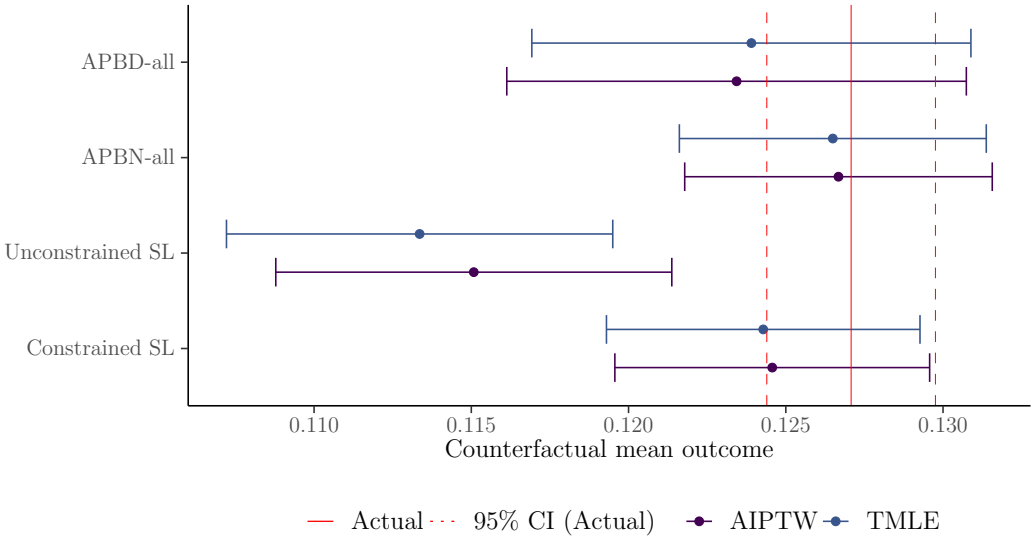
home ownership status, indicators for whether the household has electricity, drinking water, electric/gas cooking fuel and concrete walls, and indicators for whether the household owns a gas canister, vehicle (car or motorcycle) and gold – that form $V2$.

3.7.4 Evaluating the estimated policy rules

Figure 3.5 presents point estimates and 95% confidence intervals of the counterfactual mean outcomes under the super learner estimates of the optimal policy rules and the static rules, in comparison to the mean outcome under the actual policy assignment D_i of 0.127 (95% CI +/- 0.003). Using our primary performance measure, the AIPTW estimator of the mean value of the learned policy, we find that the unconstrained rule has the best performance, with an estimated mean outcome of 0.115 (95% CI +/- 0.006). The constrained rule performs slightly worse with an estimated mean outcome of 0.125 (95% CI +/- 0.005), which is still better than the actual assignment. Assigning all households to APBD generates a mean outcome of 0.123 (95% CI +/- 0.007), and assigning all to APBN gives 0.127 (95% CI +/-

0.004). The TMLE estimator supports our primary findings, with very similar point estimates and confidence intervals.

Figure 3.5: Estimated counterfactual mean outcomes



Note: Point estimates and 95% confidence intervals (CI) are reported for the AIPTW and TMLE estimators. APBD-all (APBN-all) is the static rule that assigns APBD (APBN) to everyone. Unconstrained (constrained) SL is the super learner estimate of the unconstrained (constrained) optimal policy rule. The red solid line denotes the mean outcome under the actual policy assignment D_i.

Table 3.B.2 presents estimates of the counterfactual mean outcomes for the candidate estimators included in the super learner. We find that the majority of our algorithms, including simple linear models, generalised boosted models, causal forests, regression splines and policy trees (of all depths), perform particularly well with estimates in the range of 0.114-0.118. Overall, candidates rules that are based on simpler or restricted functional forms (for example, linear models and and shallow policy trees) outperform those based on more complex models (for example, neural networks and support vector machines). Figure 3.B.3 presents the average weighted contributions of the candidate estimators across the sample-specific super learner estimates of the policy rules. Causal forests (fitted on $V1$) and generalised linear models (fitted on $V2$) receive the largest weightings in the respective unconstrained and constrained rules.

Table 3.4 presents descriptive statistics of the learned subgroups under the super learner estimates of the optimal policy rule for selected variables in X. We compare the characteristics of households under the optimal policy to those under the actual assignment. Compared to households actually assigned to APBD, those that are counterfactually assigned to APBD under the unconstrained rule are more likely to be urban-based, where health care facilities are even more accessible. They are slightly less likely to reside in the regions of Maluku-Papua and Sumatera, where the current APBD allocation is concentrated, but in Sulawesi, Bali, and Java. The opposite is true for APBN, in that compared to the actual assignment, the unconstrained rule counterfactually assigns APBN to less urban households, particularly in the regions of Sumatera and Maluku-Papua, where health care is less accessible. There are also some important socioeconomic differences that strengthen this urban-rural distinction. Compared to the actual assignment, households counterfactually assigned to APBD under the rule have more educated members, their homes are more likely to have basic amenities and assets, and they are also more likely to access technology (for example, internet and cellphones). Households counterfactually assigned to APBN are slightly worse off in terms of the same features. With regard to demography, APBD-assigned households under the rule are smaller, with heads of households that are older (aged 45 and over) and more likely to be employed in the secondary and tertiary sectors, compared to the current assignment. On the contrary, APBN-assigned households under the rule are larger, with heads of households that are younger (aged below 45) and working in the primary sector. Under the constrained rule, we find similar differences between households that are assigned to the two schemes compared to the actual assignment. However, certain differences in characteristics are more pronounced. For example, households that are counterfactually assigned to APBD are even more likely to be educated and have access to technology, while the opposite is true for households that are counterfactually assigned to APBN.

Figures 3.B.4-3.B.9 plot the learned policy trees that are candidates in the super learner estimate of the unconstrained optimal policy. The splitting covariates support our findings from the subgroup analyses that households counterfactually assigned to APBD under the rule are socioeconomically richer than those assigned

to APBN, with better household amenities (for example, protected drinking water, and cooking and toilet facilities) and assets (for example, vehicles and cellphones). The shallower trees (of depth 1) only split on the availability of household amenities, while the deeper trees (of depths 1 and 2) introduce characteristics associated with household members (for example, the household size and the number of members in work or with education).

3.8 Discussion

In this paper, we considered the problem of assigning Indonesia's two subsidised health insurance schemes to eligible households from the perspective of the state government. We learned optimal policy rules that efficiently assign households to treatment (PBI-APBD) and control (PBI-APBN) using an objective function where utility gains correspond to a reduction in the expected probability of households incurring catastrophic expenditure from health spending. We found that, although average treatment effects are small, the financial impact of being enrolled into APBD over APBN varies with observed covariates, implying that utility gains could be maximised with CATE-based policy rules, which we estimate using the super learner. Our unconstrained optimal policy rule that targets a selection of preferred covariates has a lower expected risk of catastrophic expenditure than the actual assignment and the static rules. Under a 10% budget constraint, the optimal policy also outperforms the actual (unconstrained) assignment and the static rule that assigns APBN to all. Although the positive effects of the estimated rules appear modest in percentage terms, the reduction in the total number of uninsured households facing catastrophic health expenditure from implementing these rules compared to the actual assignment strategy could be substantial at the population level. For example, in a population of approximately 245 million (in 2017) with 22.5% uninsured (=55 million) (as reported by Mahendradhata et al. (2017)), the relative reduction in the number of uninsured facing catastrophic health expenditure is approximately 660,000 for the unconstrained rule, and 110,000 for the constrained rule. One of the main differentiating features between the assignment strategy under the estimated rules and the actual assignment is geography, particularly the urban-rural distinc-

tion, which is known to be linked to the availability of health services, and a key determinant of health spending (Agustina et al., 2019; Johar et al., 2018; Sambodo et al., 2021). An optimal policy rule that shifts a small proportion of urban enrollees from APBN to APBD could improve outcomes, assuming that this strategy is chosen over one that improves the availability of health services in less-developed regions. Socioeconomic differences are also present within the assignment strategy. Households that are counterfactually assigned to APBD are more likely to be better off than those assigned to APBN, in terms of characteristics that can be easily verified by decision makers. For example, whether the household has electricity, toilet facilities or a vehicle. The wider goal of JKN is to achieve full population coverage, which would require generalising the policy rules estimated for the PBI-insured population to the uninsured. Based on our previous work that compares the uninsured and PBI-insured population in terms of a diverse set of observable characteristics, we know that the two populations are largely similar, but differences do exist for a few characteristics. For example, the uninsured are slightly better off than the insured group, which explains why they did not meet the PBI eligibility criteria in the first place. However, after weighting the populations by the inverse of the propensity score, we found that they are balanced on all of our included covariates. In terms of unobservable characteristics, there may be slight differences in health risk due to the social determinants of health, but again the socioeconomic differences between the two populations is very limited, so this not likely to drastically affect the generalisability of the policy rule. Assigning the uninsured population to the appropriate PBI scheme could be supported by the results from our classification analyses and visual decision trees.

Potential criticisms of relying on ensembling tools for policy learning are that their underlying methods are based on a black-box, meaning that the exact contribution of each covariate to prediction is unclear. The use of machine learning in policy decisions has raised some concerns about the potential ethical and equity implications (Kube et al., 2019). The learned rule may need to be constrained to belong to a policy class that imposes restrictions on the functional form, for example, for added interpretability (Kitagawa and Tetenov, 2018; Zhou et al., 2022). In general, the requirement for human interpretability to justify policy decisions could result in a

sub-optimal rule in terms of the objective function. If the expected loss in utility between an optimal, uninterpretable rule and a nearly-optimal, interpretable rule is small, decision makers may choose to opt for the latter. Learning optimal policies using the super learner offers a structured solution to addressing the trade-off between optimality and interpretability, as the candidate library can include a diverse set of estimators according to the decision maker's preferences. In our assignment problem, we find that a policy that belongs to a class of interpretable and transparent rules (for example, simple linear models and shallow decision trees) may achieve similar or even better outcomes in finite samples than the black-box ensemble. Further, the comparable performance of candidate estimators that consider $V1$ and $V2$ also implies that a policy rule that targets only the most important predictors of treatment effect heterogeneity, which we identify using regularisation on the CATE model, can also produce an interpretable solution. Since our policy rule is based on decision criteria that were curated by policymakers, and we can, to some extent, explain the construction of the rule, we hope that the potential impacts of algorithmic bias on our findings are reduced (Panch et al., 2019).

Our findings highlight the potential of combining ex-ante and ex-post targeting measures to improve the delivery of social programmes. Evaluating the progress of implemented policies, through an assessment of impact heterogeneity, could strengthen policy corrections and guide future decisions, by potentially identifying a superior assignment strategy that achieves larger expected benefits. We acknowledge that policymaking at the national level is incredibly complex. Health care may be one of several objectives of the state, so aligning targeting criteria across the social welfare function is important. Improving equity is a policy priority in most health care systems, including Indonesia (Johar et al., 2018). We do not directly encode equity constraints into our objective function, although in theory this could be possible with some algorithmic modifications. For example, an equity constraint could assign centrally-funded insurance to all households that receive social security, regardless of their observed covariates. We do, however, include proxies for socioeconomic status in our pre-specified targeting criteria that could indirectly address equity considerations. Since our learned policy trees define an assignment strategy that is primarily based on socioeconomic disparities, policymakers could use these to

guide fairness decisions. Learning policy rules can be time and resource-intensive. A well-constructed rule relies on gathering sufficient data and having an appropriately skilled workforce to conduct the analyses. Sustaining utility gains would require regularly updating the rule using new information that reflects any changes in covariate distributions, particularly if these changes are in response to the rule itself. However, this is unlikely in our empirical problem given the relatively comparable characteristics between treated and control populations.

Our study has some limitations. We assume no unobserved confounding, given our rich set of included controls, and rely on a doubly robust adjustment for observed confounding to estimate our target causal parameter. Kreif et al. (2022) use a similar selection on observables approach to evaluate the average and heterogeneous impacts of JKN on maternal health care utilisation and infant mortality. Other evaluations of JKN adjust for observed confounding using propensity score weighting and matching methods (Anindya et al., 2020; Erlangga et al., 2019a).[24] We previously discussed our justification for restricting the sample to households that are enrolled into either PBI scheme, and excluding uninsured households. By controlling for head-of household characteristics, household assets and health care service accessibility, we try to minimise unobserved confounding that may explain the process of selection into APBD or APBN, which can create bias in our estimates of the treatment effects, and consequently, the policy rules. Kallus and Zhou (2021) propose a framework for optimising the minimax regret obtained from a learned policy rule against a baseline policy (for example, the actual assignment), in the presence of unobserved confounding. Their approach currently applies to IPTW estimators of the policy value, which they plan to extend to doubly robust estimators. Other approaches for reducing the impact of unmeasured confounding could include instrumental variables analysis and panel data methods, which were not possible within the confines of our data. We also make explicit the fact that we trim observations with extreme propensity scores, which are becoming more common in larger datasets. We acknowledge that trimming methods are sensitive to the pre-defined

[24]Erlangga et al. (2019a) combine propensity score matching with difference-in-differences analysis to also address unobserved confounding. However, they use an alternative data source (Indonesian Family Life Survey) and explore the impact of JKN on health care utilisation, not financial protection.

cut-off points, and can result in a substantial sample size reduction (Li et al., 2019). Recent developments in overlap weighting, which up-weights households with the most overlap and down-weights households in the propensity score tails, has been proposed as a better alternative (Li et al., 2018). The latest super learner algorithm, however, does not enable overlap weighting, so we are restricted to trimming methods in our study. Lastly, we encode a 10% budget constraint into our objective function to demonstrate how we can generate a constrained policy rule from the perspective of the state government, who reportedly give 10% of their health care budget to local governments. We are aware that this constraint does not take into account any additional sources of APBD funding. A truly representative policy rule with constraints would require a consultation with policymakers to understand the exact budgetary restrictions on PBI. In addition, the associated costs of rolling out PBI to additional groups may be an important consideration for policymakers. Learning optimal policy rules that maximise population outcomes while also minimising costs is a natural extension of this work (Lakkaraju and Rudin, 2017; Xu et al., 2020). Lastly, in our super learner library of candidate estimators, we include rules that are based on the CATE function (threshold- and tree-based rules) and static rules. In theory, other candidate estimators that do not rely on estimating the CATE function could be included. We leave this for future research.

Table 3.4: Classification analysis of selected variables in X, for households assigned to PBI-APBN and PBI-APBD under the learned policy rules, estimated using the unconstrained and resource-constrained super learners.

	Actual assignment		**Unconstrained rule**				**Constrained rule**			
	APBN (n=39,367)	APBD (n=22,375)	APBN (n=23,598)		APBD (n=38,144)		APBN (n=58,168)		APBD (n=3,574)	
	Mean	Mean	Mean	Diff	Mean	Diff	Mean	Diff	Mean	Diff
Head of household characteristics										
Male	0.814	0.850	0.828	0.013	0.827	-0.023	0.827	0.012	0.834	-0.016
Age 0-24	0.013	0.018	0.015	0.002	0.015	-0.003	0.015	0.002	0.013	-0.006
Age 25-44	0.307	0.391	0.340	0.033	0.336	-0.055	0.338	0.030	0.341	-0.051
Age 45-64	0.508	0.461	0.487	-0.021	0.494	0.033	0.492	-0.017	0.487	0.025
Age 65+	0.154	0.114	0.140	-0.014	0.139	0.025	0.139	-0.015	0.144	0.030
Married	0.770	0.805	0.781	0.011	0.783	-0.021	0.782	0.012	0.791	-0.014
In employment	0.855	0.892	0.867	0.012	0.869	-0.023	0.869	0.014	0.862	-0.030
Employment: primary sector	0.437	0.524	0.469	0.032	0.469	-0.055	0.469	0.032	0.457	-0.067
Employment: secondary sector	0.063	0.047	0.058	-0.005	0.057	0.010	0.057	-0.006	0.056	0.010
Employment: tertiary sector	0.354	0.321	0.340	-0.014	0.344	0.022	0.342	-0.012	0.348	0.027
Literate: Latin/Arabic letters	0.914	0.907	0.913	-0.001	0.910	0.004	0.911	-0.003	0.911	0.004
Compulsory education	0.684	0.683	0.683	-0.001	0.684	0.001	0.684	-0.001	0.687	0.004
Had a cellphone in previous 3 months	0.585	0.610	0.594	0.009	0.593	-0.016	0.594	0.009	0.587	-0.023
Used internet in previous 3 months	0.117	0.106	0.116	-0.001	0.112	0.005	0.114	-0.003	0.107	0.001
Household characteristics										
Number of members	3.538	3.693	3.590	0.053	3.596	-0.097	3.593	0.055	3.618	-0.075
Number of productive members (aged 15-64)	2.366	2.340	2.358	-0.008	2.356	0.015	2.356	-0.010	2.365	0.025
Number of children in school	0.394	0.352	0.376	-0.018	0.381	0.029	0.378	-0.016	0.388	0.037
Number of members in employment	1.684	1.695	1.689	0.004	1.688	-0.007	1.688	0.004	1.693	-0.002
Number of members with compulsory education	2.211	2.136	2.179	-0.032	2.187	0.051	2.182	-0.029	2.218	0.082
Number of rooms	5.851	5.796	5.836	-0.014	5.828	0.031	5.830	-0.021	5.845	0.049
Location: urban	0.372	0.315	0.349	-0.023	0.352	0.038	0.351	-0.021	0.355	0.041
Household asset ownership										
Electricity	0.958	0.908	0.942	-0.016	0.938	0.031	0.940	-0.018	0.935	0.027
Purchases drinking water	0.596	0.559	0.586	-0.010	0.581	0.022	0.583	-0.014	0.580	0.021
Private/shared toilet	0.794	0.790	0.793	-0.001	0.792	0.003	0.792	-0.002	0.804	0.015
Refrigerator	0.407	0.422	0.408	0.001	0.415	-0.007	0.413	0.006	0.407	-0.015
Gas canister >5.5kg	0.062	0.053	0.059	-0.003	0.058	0.005	0.058	-0.003	0.059	0.006
Car/motorcycle	0.638	0.660	0.648	0.009	0.646	-0.015	0.646	0.007	0.652	-0.008
Gold/jewellery >10g	0.131	0.139	0.135	0.005	0.133	-0.006	0.134	0.003	0.131	-0.007
District characteristics										
Easy access: secondary care	0.793	0.681	0.755	-0.038	0.750	0.069	0.752	-0.040	0.750	0.069
Easy access: community care	0.979	0.924	0.960	-0.019	0.958	0.034	0.959	-0.020	0.957	0.034
Easy access: primary care	0.887	0.815	0.866	-0.021	0.858	0.043	0.861	-0.026	0.857	0.042
Easy access: maternity care	0.867	0.781	0.839	-0.028	0.834	0.052	0.836	-0.031	0.828	0.047
Region										
Sumatera	0.311	0.360	0.328	0.017	0.329	-0.030	0.329	0.018	0.317	-0.042
Jakarta	0.003	0.000	0.002	-0.001	0.002	0.002	0.002	-0.001	0.003	0.003
Java	0.312	0.270	0.296	-0.015	0.297	0.027	0.296	-0.016	0.306	0.036
Bali,NTB,NTT	0.069	0.030	0.055	-0.014	0.054	0.024	0.054	-0.015	0.064	0.034
Kalimantan	0.072	0.088	0.079	0.007	0.077	-0.011	0.078	0.006	0.073	-0.015
Sulawesi	0.165	0.075	0.135	-0.030	0.131	0.056	0.133	-0.032	0.129	0.054
Maluku-Papua	0.069	0.177	0.104	0.036	0.110	-0.067	0.108	0.039	0.108	-0.069

Note: Sample means are reported for selected variables included in X for households in the trimmed 70% data partition that are assigned to PBI-APBD and PBI-APBN under the actual assignment, and the optimal unconstrained and constrained policy rules (estimated using the super learner). The absolute differences ("Diff") in covariate means for households that are counterfactually assigned to APBN and APBD under the estimated optimal policy rules compared to the actual assignment are also reported. Coloured cells denote the p-values from Welch's two-sample t-test on whether these differences in covariate means are significant. Two-sample t-test is conducted at a significance level of 0.05. If p-value on t-statistic <0.05: cell colour is green; white otherwise.

Appendix 3.A List of all covariates used for confounder adjustment

Head of household-level (binary)

Male

Marital status: married

Age 0-24

Age 25-44

Age 45-64

Age 65+

Has a national identity number

Literate: Latin/Arabic letters

Educated (at the compulsory level)

Travelled domestically for tourism in 2016

Had a cellphone in previous 3 months

Used internet in previous 3 months

Employment status: in employment

Employment sector: primary

Employment sector: secondary

Employment sector: tertiary

Household-level (count)

Educated (at the compulsory level)

Employment status: in employment

Productive members (aged 15-64)

Children at school

Size (members)

Size (families)

Rooms

Household-level (binary)

Location: urban area

Home occupancy status: owner

Has a second residence

Roof material: concrete/tile

Wall material: concrete

Floor material: marble/granite/ceramic/parquet/vinyl/carpet

Toilet facility: private/shared

Protected drinking water source

Purchases drinking water

Electricity

Cooking fuel: gas/electric

Sewage disposal: septic tank/sewage system

Experienced a natural diaster in previous year

Natural tourism in residential area

Savings account

Goods ownership: gas (over 5.5kg)

Goods ownership: refrigerator

Goods ownership: air conditioning

Goods ownership: radiator

Goods ownership: landline

Goods ownership: computer

Goods ownership: gold (over 10g)

Goods ownership: boat

Goods ownership: car/motorcycle

Goods ownership: television

Goods ownership: land

District-level (binary)

Easy access to primary health care

Easy access to community health care

Easy access to maternal health care

Easy access to secondary (hospital) health care

Easy access to pharmacy

Regional-level (binary)

Region: Sumatera

Region: Jakarta

Region: Jawa

Region: Bali, NTB, NTT

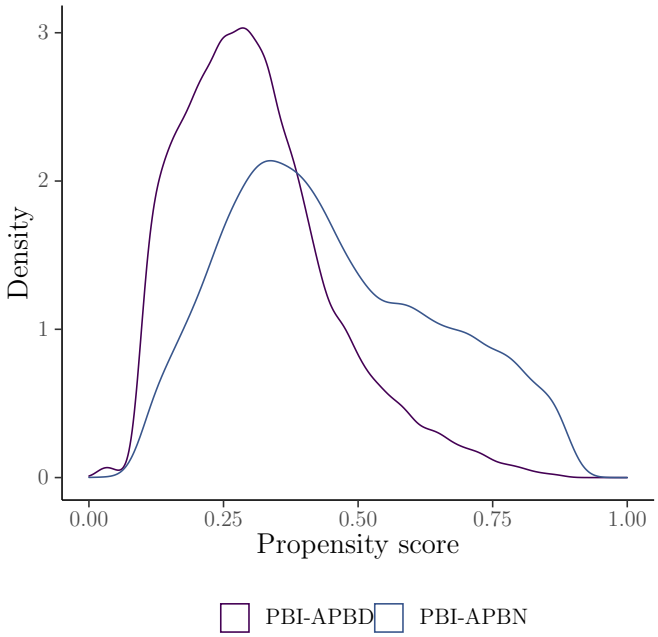
Region: Kalimantan

Region: Sulawesi

Region: Maluka-Papua

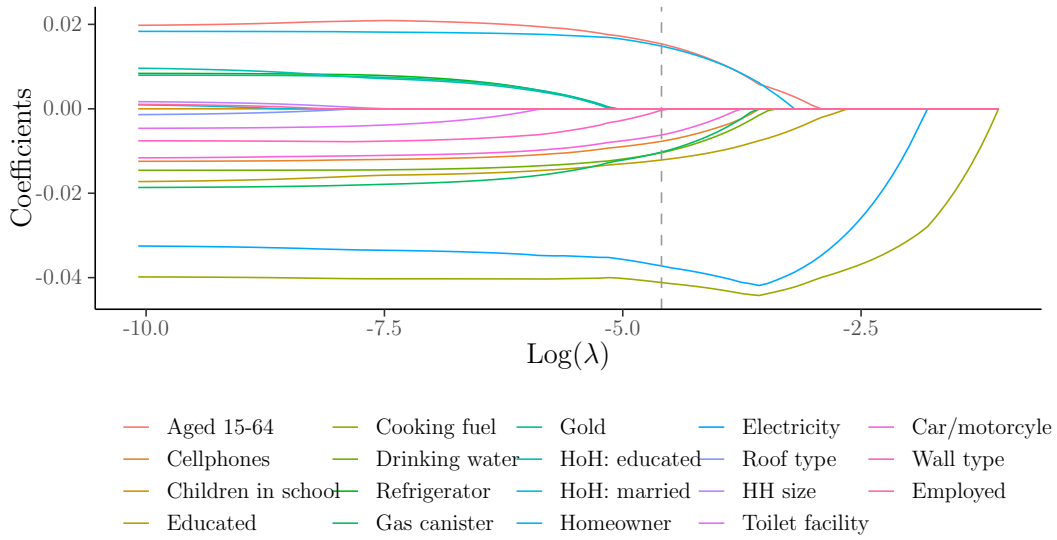
Appendix 3.B Additional figures/tables

Figure 3.B.1: Overlap plot



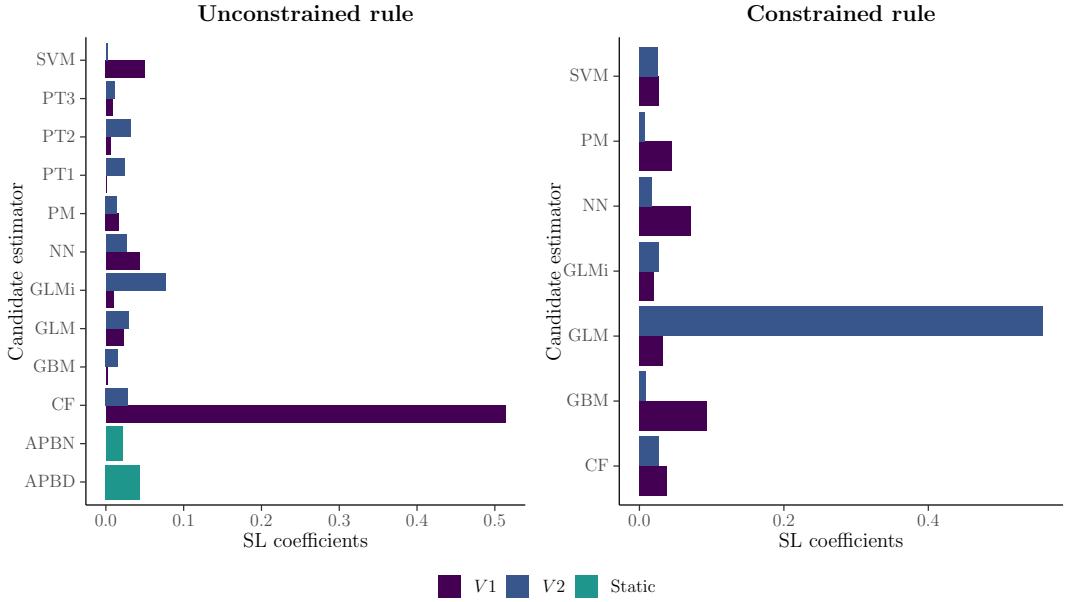
Note: Density plot showing the distribution of propensity scores for households enrolled into PBI-APBD and PBI-APBN in the trimmed 70% data partition.

Figure 3.B.2: Variable trace plot of adaptive LASSO fit



Note: Plot shows the coefficients on $V1$, as a function of the $\log(\lambda)$ values used in the cross-validated adaptive LASSO model. The grey dashed line represents the value of $\log(\lambda)$ that minimises the cross-validated mean squared error. See Table 2 for a detailed description of the covariates in $V1$. HoH = head of household.

Figure 3.B.3: Weighted contributions of candidate estimators to the super learner



Note: Average contribution of each candidate estimator in the super learner across cross-fitted samples are displayed. Static rules are not included in the candidate library for the constrained rule.

Table 3.B.1: Results from heterogeneity test using difference-in-means estimator

	V1				V2			
	Est	SE	Unadj p-val	Adj p-val	Est	SE	Unadj p-val	Adj p-val
CF								
Q2-Q1	-0.055	0.000	0.000	0.000	-0.054	0.000	0.000	0.000
Q3-Q1	-0.073	0.000	0.000	0.000	-0.072	0.000	0.000	0.000
Q4-Q1	-0.088	0.000	0.000	0.000	-0.083	0.000	0.000	0.000
Q5-Q1	-0.113	0.000	0.000	0.000	-0.101	0.000	0.000	0.000
GBM								
Q2-Q1	-0.011	0.000	0.000	0.000	-0.011	0.000	0.000	0.000
Q3-Q1	-0.017	0.000	0.000	0.000	-0.018	0.000	0.000	0.000
Q4-Q1	-0.019	0.000	0.000	0.000	-0.020	0.000	0.000	0.000
Q5-Q1	-0.021	0.000	0.000	0.000	-0.022	0.000	0.000	0.000
GLM								
Q2-Q1	-0.039	0.000	0.000	0.000	-0.037	0.000	0.000	0.000
Q3-Q1	-0.057	0.000	0.000	0.000	-0.055	0.000	0.000	0.000
Q4-Q1	-0.070	0.000	0.000	0.000	-0.067	0.000	0.000	0.000
Q5-Q1	-0.088	0.000	0.000	0.000	-0.081	0.000	0.000	0.000
GLMi								
Q2-Q1	-0.072	0.000	0.000	0.000	-0.053	0.000	0.000	0.000
Q3-Q1	-0.104	0.000	0.000	0.000	-0.073	0.000	0.000	0.000
Q4-Q1	-0.133	0.000	0.000	0.000	-0.091	0.000	0.000	0.000
Q5-Q1	-0.185	0.000	0.000	0.000	-0.121	0.000	0.000	0.000
SVM								
Q2-Q1	-0.088	0.001	0.000	0.000	-0.077	0.001	0.000	0.000
Q3-Q1	-0.124	0.001	0.000	0.000	-0.106	0.001	0.000	0.000
Q4-Q1	-0.156	0.001	0.000	0.000	-0.125	0.001	0.000	0.000
Q5-Q1	-0.227	0.001	0.000	0.000	-0.163	0.001	0.000	0.000
NN								
Q2-Q1	-0.032	0.001	0.000	0.000	-0.048	0.000	0.000	0.000
Q3-Q1	-0.036	0.001	0.000	0.000	-0.052	0.000	0.000	0.000
Q4-Q1	-0.037	0.001	0.000	0.000	-0.080	0.000	0.000	0.000
Q5-Q1	-0.064	0.001	0.000	0.000	-0.127	0.000	0.000	0.000
PM								
Q2-Q1	-0.053	0.000	0.000	0.000	-0.040	0.000	0.000	0.000
Q3-Q1	-0.077	0.000	0.000	0.000	-0.058	0.000	0.000	0.000
Q4-Q1	-0.098	0.000	0.000	0.000	-0.069	0.000	0.000	0.000
Q5-Q1	-0.132	0.000	0.000	0.000	-0.086	0.000	0.000	0.000

Note: Table reports estimates and standard errors of the differences in sorted GATEs (for quintiles of predicted CATEs) between the lowest quintile (Q1) and higher quintiles (Q2-Q5). Unadj p-val does not correct for multiple hypothesis testing. Adj p-val uses the Romano-Wolf procedure to correct for multiple hypothesis testing.

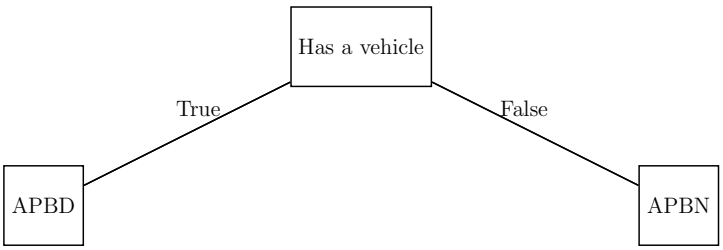
Table 3.B.2: Counterfactual mean outcomes for candidate estimators included in the super learner

	AIPTW		**TMLE**	
	Est	SE	Est	SE
Static rules				
APBD-all	0.123	0.004	0.124	0.003
APBN-all	0.127	0.002	0.126	0.002
Threshold-based rules				
GLM-$V1$	0.114	0.003	0.113	0.003
GLMi-$V1$	0.117	0.003	0.115	0.003
PM-$V1$	0.116	0.003	0.114	0.003
NN-$V1$	0.122	0.003	0.120	0.003
SVM-$V1$	0.120	0.003	0.122	0.003
GBM-$V1$	0.114	0.003	0.112	0.003
CF-$V1$	0.118	0.003	0.116	0.003
GLM-$V2$	0.114	0.003	0.112	0.003
GLMi-$V2$	0.114	0.003	0.113	0.003
PM-$V2$	0.114	0.003	0.113	0.003
NN-$V2$	0.119	0.003	0.117	0.003
SVM-$V2$	0.126	0.003	0.127	0.003
GBM-$V2$	0.115	0.003	0.112	0.003
CF-$V2$	0.117	0.003	0.116	0.003
Tree-based rules				
PT1-$V1$	0.116	0.003	0.114	0.003
PT2-$V1$	0.116	0.003	0.115	0.003
PT3-$V1$	0.117	0.003	0.115	0.003
PT1-$V2$	0.116	0.003	0.114	0.003
PT2-$V2$	0.116	0.003	0.114	0.003
PT3-$V2$	0.116	0.003	0.114	0.003

Note: Point estimates (Est) and standard errors (SE) are reported for the AIPTW and TMLE estimators.

Figure 3.B.4: Depth 1 policy trees (fitted on $V1$)

(a) Sample 1



(b) Sample 2

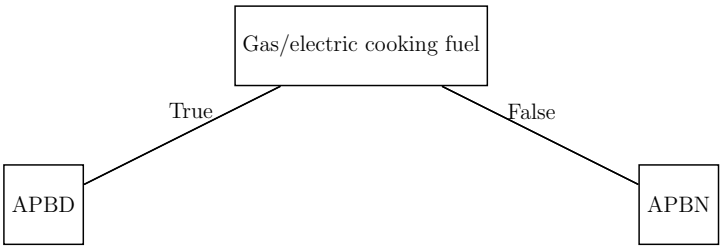
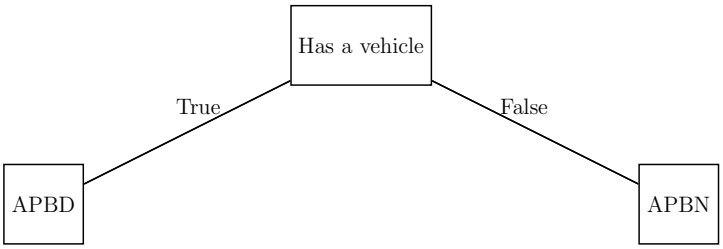


Figure 3.B.5: Depth 1 policy trees (fitted on $V2$)

(a) Sample 1



(b) Sample 2

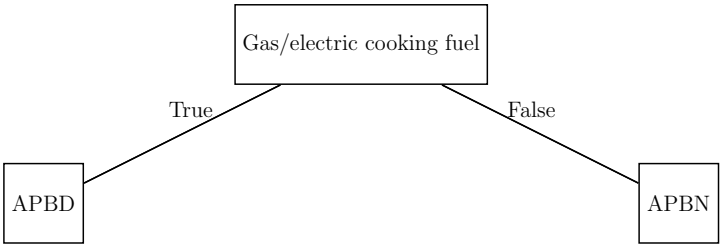
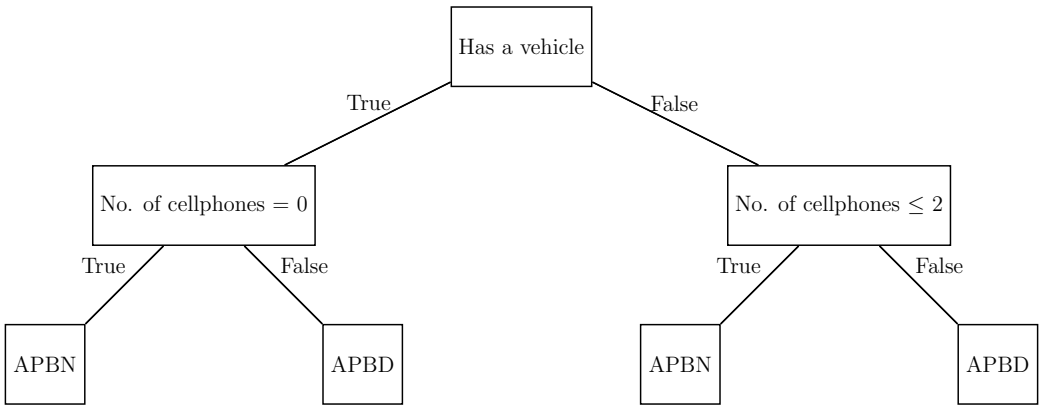


Figure 3.B.6: Depth 2 policy trees (fitted on $V1$)

(a) Sample 1



(b) Sample 2

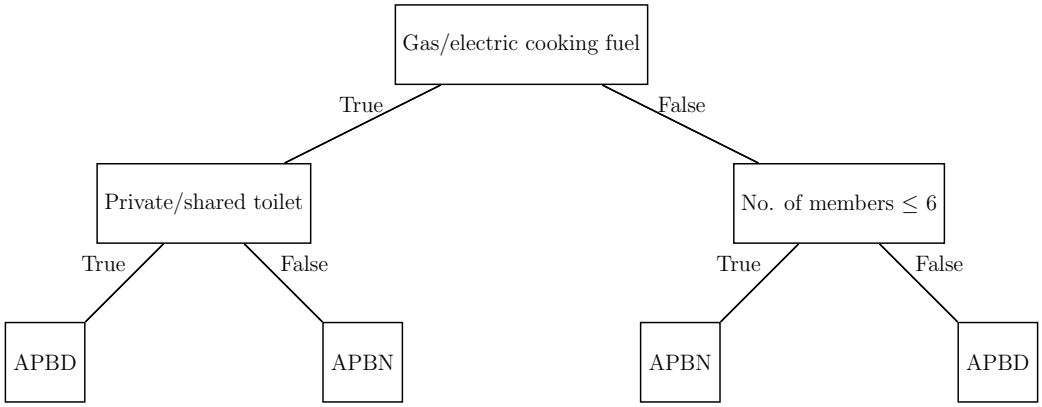
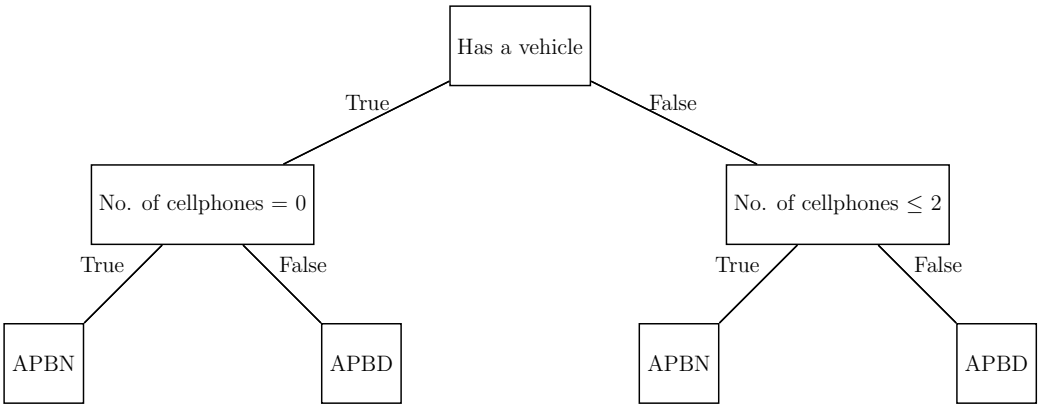


Figure 3.B.7: Depth 2 policy trees (fitted on $V2$)

(a) Sample 1



(b) Sample 2

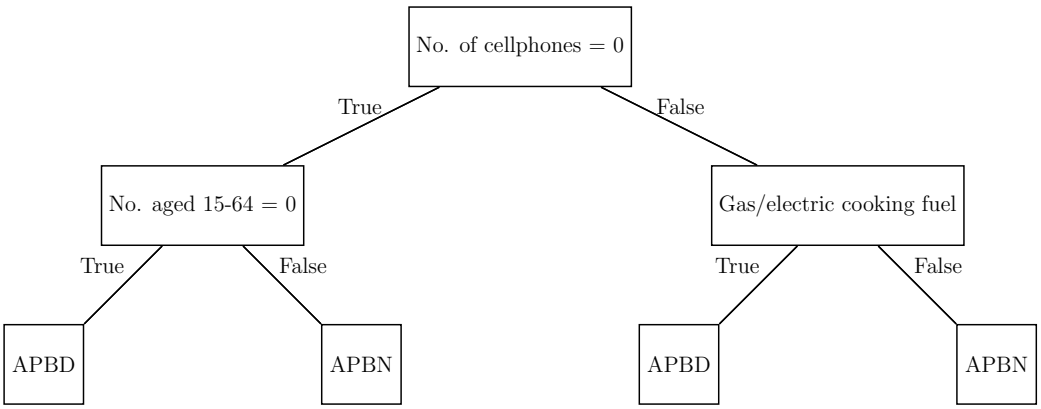
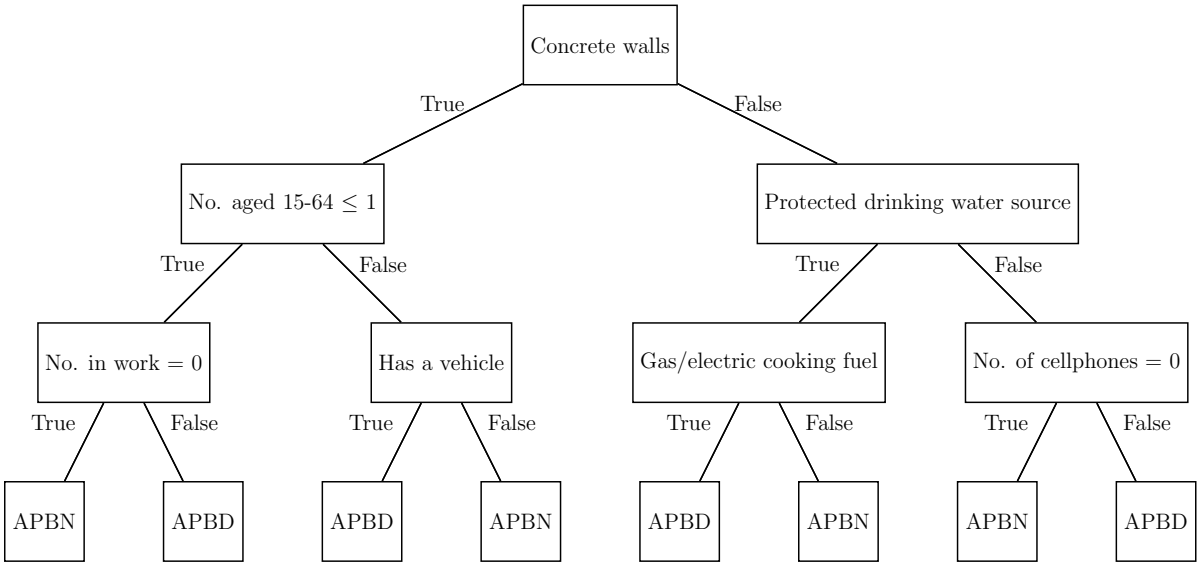


Figure 3.B.8: Depth 3 policy trees (fitted on $V1$)

(a) Sample 1



(b) Sample 2

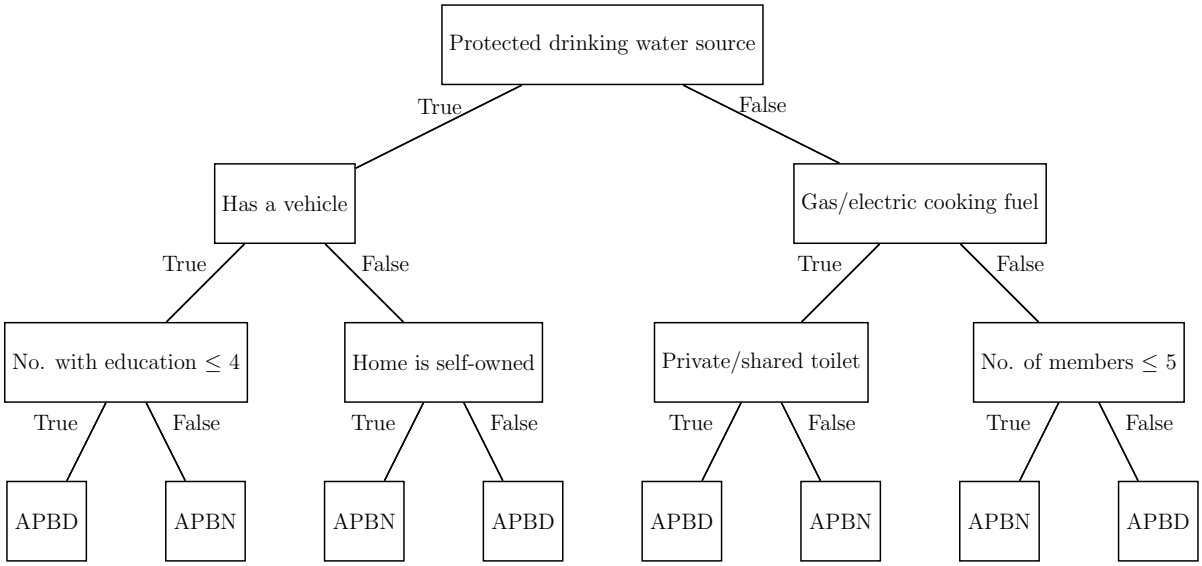
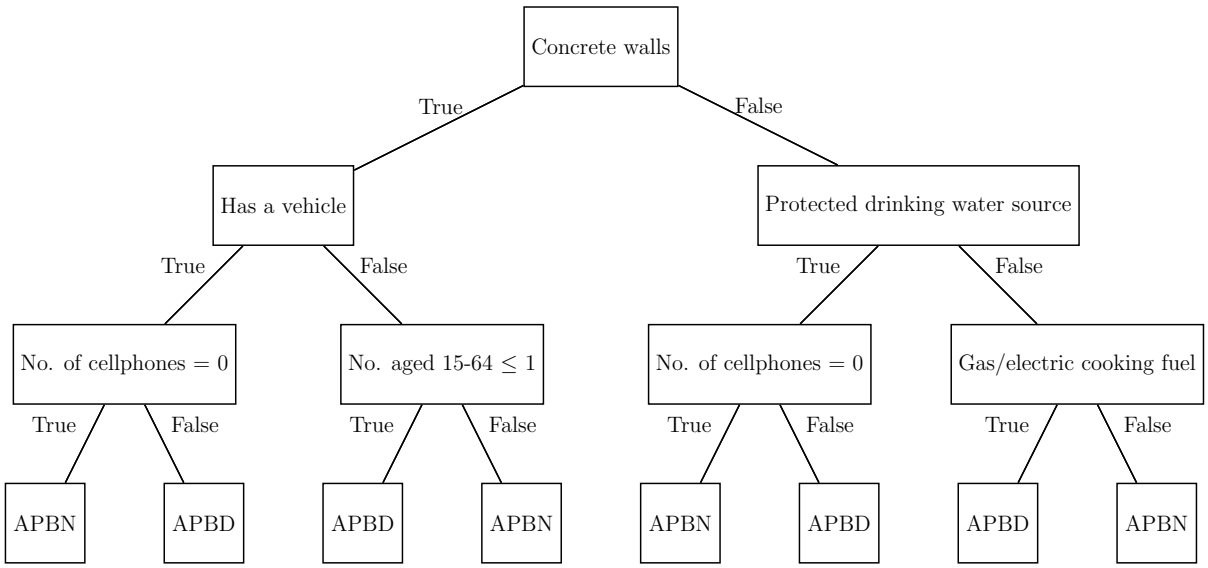
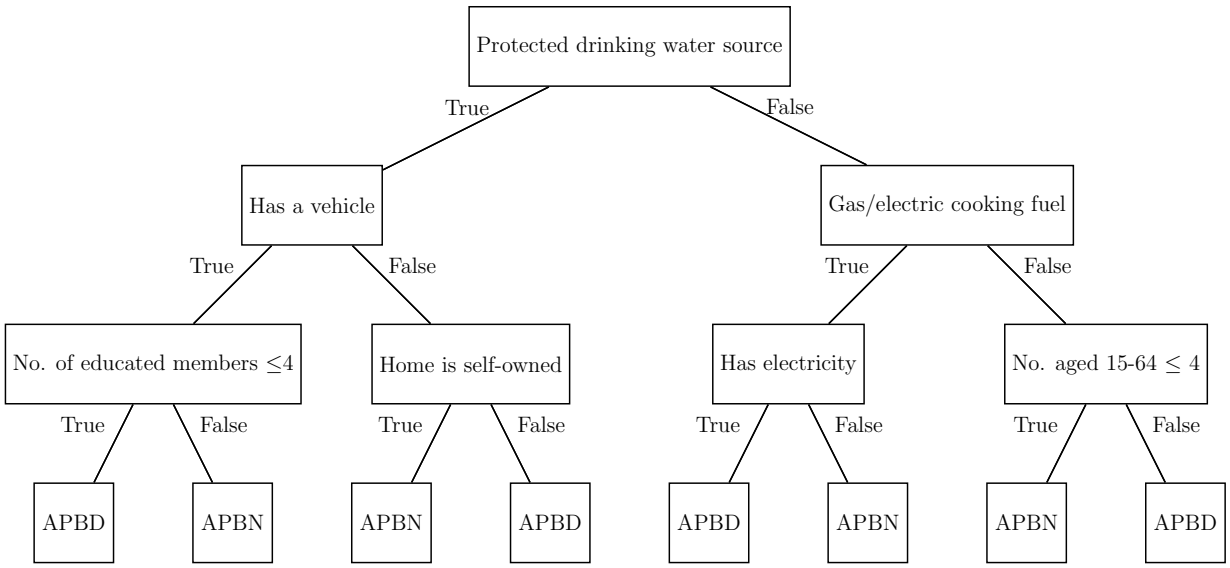


Figure 3.B.9: Depth 3 policy trees (fitted on $V2$)

(a) Sample 1



(b) Sample 2



Appendix 3.C Optimal policy learning with resource constraints

Luedtke and van der Laan (2016a) model the problem of estimating optimal resource constrained policy rules by imposing a constraint κ on the maximum proportion of units that can be treated, and defining a set of solutions that satisfy κ. The optimal policy rule is the optimal solution among the set of solutions that respect the constraint.

The formal theorem starts by defining S_P as the survival function of the CATE function $\hat{\tau}(V_i)$, i.e. the probability that $\hat{\tau}(V_i)$ is greater than some varying threshold T: $T \mapsto P(\tau(V_i) > T$.

Then, let

$$\eta := \inf\{T : S_P(T) \leq \kappa\}$$
$$T_P := \max\{\eta_P, 0\},$$

where η identifies the largest threshold value for which the survival probability is less than κ.

The optimal policy rule can be defined as follows:

$$\hat{\pi}^* := \begin{cases} \frac{\kappa - S_{P(T_P)}}{P(\hat{\tau}(V_i) = T_P)}, & \text{if } \hat{\tau}(V_i) = T_P \text{ and } T_P > 0 \\ \mathrm{I}(\hat{\tau}(V_i) > T_P), & \text{otherwise.} \end{cases}$$

Discussion

This thesis explores the recent methodological developments in heterogeneous treatment effect estimation using causal machine learning, and applies these methods to evaluate the average and heterogeneous impacts of Indonesia's national health insurance programme, the JKN, on two important targets of universal health coverage schemes; the utilisation of health care services and the financial protection from catastrophic health expenditures. The treatment effect estimates from the policy evaluation are used to learn optimal policy rules that efficiently allocate JKN to the eligible population according to their observed characteristics, in a way that maximises welfare, defined as reductions in catastrophic health expenditure.

Chapter 1 reviews the current literature on the use of causal machine learning methods to evaluate the heterogeneous impacts of binary treatments (or policies) in a selection on observables framework. It focuses on three promising algorithms – the X-learner, the R-learner and causal forests – and describes their relative strengths and weaknesses based on the policy setting and study design. To demonstrate their application, the chapter concludes with a case study evaluation of the subsidised JKN scheme on the utilisation of inpatient health care, which has the characteristics of a typical health policy evaluation: a binary outcome and a large, diverse set of confounders and effect modifiers. The results from the case study stress the importance of looking beyond average impacts and highlight the benefits of supplementing theory-driven subgroup analyses with a data-driven approach. From a methodological perspective, the key takeaway is that for policy evaluations which assume no unobserved confounding, the R-learner is a suitable approach since it can flexibly estimate heterogeneous treatment effects using any method that is formulated as a loss-minimisation problem, while also achieving strong empirical performance and asymptotic guarantees (Nie and Wager, 2021). In particular, an implementation of the R-learner using causal forests generates estimates of the conditional average treatment effect (CATE) function that are robust to observed confounding; a feature that supports the selection of causal forests as the method of choice for the

evaluation of JKN in Chapter 2.

This chapter contributes to the causal machine learning and policy evaluation literature, given the rapid increase in the number of methods available for estimating treatment effect heterogeneity, and limited guidance on which methods to use. The selection of these three methods was based on their popularity among applied researchers at the time of writing, combined with their relative ease of implementation given their associated `R` software packages. However, as with any rapidly evolving field, newer methods, as well as extensions to existing methods, have since emerged. For example, the DR-Learner by Kennedy (2020), where the DR refers to the double robustness property of the learner, extends the meta-learner framework to incorporate double machine learning and cross-fitting into the estimation task (Chernozhukov et al., 2017). Given its double robustness property, it is an appropriate estimator for policy evaluations where confounding adjustment is crucial. For other settings where the assumption of no unobserved confounding cannot be satisfied, and repeated cross-sectional data is available, Nie et al. (2019) propose a non-parametric heterogeneous treatment effect estimator using a two-period difference-in-differences design that compares the changes in outcomes over time; before and after the policy is introduced. Future evaluations of the JKN policy could implement this type of estimator to reduce the effects of remaining confounding in the CATE estimates, provided that repeated cross-sectional data is available and that the treated and control groups can be identified in the pre-treatment period.

Chapter 2 evaluates the the impact of being enrolled into the subsidised JKN programme (PBI), compared to being uninsured, on health care utilisation, measured by the demand for inpatient and outpatient care. This study extends the case study example from the previous chapter, which also considers a measure of health care utilisation (inpatient demand) as the outcome, for a more detailed evaluation of JKN. A prediction algorithm, the super learner, is used to estimate the nuisance functions, and a causal algorithm, causal forests, is used to estimate CATEs. The findings show that, although on average insurance take up is associated with an increase in demand for health care, these positive impacts are not consistent across the entire population. Some parts of the population do not change their health care

utilisation, and in some instances, even decrease their demand for outpatient care. The additional decomposition of the outcome model enables a separate exploration of the participation and consumption effects of insurance on health care demand. The findings show that, for those that decide to access care as a result of insurance, the increase in the amount of care consumed is particularly large. The results from the subgroup analyses find larger policy effects among urban respondents, where health care facilities are more accessible. The data-adaptive methods identify important effect modifiers (e.g. marital and employment status, technology usage), in addition to those motivated by theory (e.g. age, household expenditure and health care accessibility) that drive heterogeneity in treatment effects.

This chapter highlights the potential for combining predictive and causal machine learning to improve the estimation of causal parameters, in particular the CATE function. The study leverages upon several recent innovations in the literature to estimate treatment effect heterogeneity and to summarise the findings in a meaningful, constructive way. For example, incorporating sample splitting and cross-fitting, which is not routinely performed in policy evaluation studies. This chapter contributes to the applied causal machine learning literature by demonstrating alternative ways of exploring and interpreting treatment effect heterogeneity, and highlighting how data-driven subgroup analyses can complement theoretically-motivated analyses. The decomposition of the outcome model also generates novel findings on the participation and consumption effects of insurance on health care demand. More generally, the outputs from this chapter echo those from the previous chapter that more careful considerations of policy impacts beyond the average can improve future policymaking. A natural extension of this work relates closely to Chapter 3, in that the estimated CATEs could be used to learn an optimal policy rule that efficiently assigns subsidised insurance to the eligible population. However, since the aim is to leave no one uninsured, a relevant decision question would be to choose between different modalities of subsidised health insurance. Secondly, an important target of subsidised health insurance is to reduce catastrophic health expenditure, which is the main outcome measure in Chapter 3.

Chapter 3 considers the problem of learning optimal policy rules that assign In-

donesia's two subsidised health insurance schemes, PBI-APBD and PBI-APBN, to eligible households, with the aim of minimising the expected population risk of incurring catastrophic health expenditures. The study finds varying impacts on the population of being enrolled into APBD over APBN, which justifies learning CATE-based policy rules. The super learner is used to estimate optimal policy rules, using a diverse candidate library of threshold- and tree-based rules, as well as simple static rules, and the learned rules are evaluated using doubly robust estimators of the counterfactual mean outcome. The unconstrained and resource-constrained rules outperform the actual policy assignment and the static rule that assigns APBN to everyone. The static rule that assigns APBD to everyone performs slightly better than the constrained rule, but is outperformed by the unconstrained rule. Through characterising households that are counterfactually assigned to APBD and APBN, it is evident that regional variation, particularly the urban-rural distinction, is a key differentiating factor between the two populations.

This chapter consolidates the ideas introduced in the previous chapters into a tangible output that can be used to influence policy. In terms of contributions, it is the first paper to consider tree-based and super learner-based policy learning within the same framework. It is also one of the first evaluations of a system-level health policy that learns optimal policy rules from heterogeneous treatment effect estimation. In recent years, the policy learning literature has expanded considerably, however the methods are still very much in development. Within the statistics community, progress is being made on adapting the current methodology on policy learning from observational data to address certain challenges, such as unobserved confounding and overlap issues (Kallus, 2020; Kallus and Zhou, 2021). From a computational perspective, researchers are trying to develop better algorithms that can find optimal solutions faster and can incorporate other features, such as continuous treatments, a large number of confounders and effect modifiers, and pre-imposed constraints (e.g. equity and budget) on the objective function (Amram et al., 2022; Bertsimas et al., 2019; Kallus, 2017; Kallus and Zhou, 2018b; Liu et al., 2021).

The overall findings from this thesis generate some important considerations for Indonesian policymakers. Most notably that a variety of demographic, socioeconomic

and regional factors contribute to the heterogeneous impacts of JKN on health care utilisation and financial protection. If the only objective of policymaking is to maximise welfare without any ethical or equity constraints, assigning health insurance based on an estimated optimal policy rule would be an easy solution. In reality, targeting policies to populations that are expected to benefit the most, or that live in a certain region, or are of a certain socioeconomic status, for example, is unlikely to happen. A more realistic output of this work is for policymakers to explore why health insurance is more effective in some parts of the population, and whether the intended beneficiaries of the policy are included in this group. Essentially, this work could be used to generate hypotheses for future research into the current shortcomings of the programme, rather than for immediate decision making. For example, a common finding from this work and also from the wider literature is the disparity in policy impacts between urban and rural populations, where health care accessibility varies substantially. Addressing supply-side limitations through better health infrastructure in rural areas is an obvious strategy that could improve the overall effectiveness of the policy (Maulana et al., 2022).

Although this thesis focuses on recent advances in data-driven methods for identifying and summarising treatment effect heterogeneity, it is important to acknowledge the role of theory when using these methods. The need for theoretical justification can arise in two forms: firstly, in the model selection for heterogeneous treatment effect estimation; and secondly, in the interpretation of model outputs within the contextual setting of health and social policymaking. To address the first problem of choosing the best model, the incorporation of the super learner into the estimation task allows the user to specify a diverse library of candidate algorithms, which reduces the risk of model misspecification. The selection of confounders and effect modifiers that form the covariate space is guided by a combination of economic theory around health insurance and existing evidence around JKN and universal health coverage policies more broadly. Furthermore, in Chapter 3, the potential targeting criteria that are accessed by the machine learning algorithm to learn optimal policy rules are solely based on policymakers' preferences that are captured according to the original eligibility criteria for JKN. To address the second problem of interpretation, the outputs from the data-driven estimators in this thesis are, for the most

part, successful in identifying many of the effect modifiers that are specified a priori according to theory, in addition to identifying new effect modifiers that are not typically considered in subgroup analyses within health insurance evaluation studies. Overall, the thesis highlights the importance of aligning theory- and data-driven approaches in order to make policy recommendations that are reasonably justified.

The policy evaluations in this thesis rely on cross-sectional, non-randomised data to estimate the causal parameters using a selection on observables framework. The assumption of no unobserved confounding is made on the basis that the data source (SUSENAS) is particularly rich, and includes a large, diverse set of demographic, socioeconomic and geographic characteristics that may explain the relationship between the policy and the outcome, and can be included as controls for confounding adjustment. Despite this, the unconfoundedness assumption is inherently unverifiable and the steps taken to reduce the risk of unobserved confounding creating bias in the causal effect estimates is not guaranteed. When designing the study, the possibility of using an instrumental variable was considered to address some of these concerns, but a suitable one that satisfies the exclusion criteria could not be identified. Another approach would be to follow the same individuals or households over multiple time periods, ideally pre- and post-JKN, to exploit changes in outcomes before and after the policy implementation. The SUSENAS data, however, has a repeated cross-sectional design and information from only a single time period could be obtained, given cost constraints. Further, the health-related variables that are included in the survey have frequently changed over time, which would not be conducive to a longitudinal study, if the intended target outcomes are affected. More general limitations of earlier versions of the SUSENAS survey data have been reported, particularly in relation to the expenditure variable not reflecting true out-of-pocket expenses (Johar et al., 2018). An important shortcoming of the data that is relevant to the JKN policy evaluation problem is the limited number of health-related variables that are included in the survey. In particular, the exclusion of detailed information on co-morbidities is notable given the relationship between increased health risk and self-selection into insurance. In addition, the included measures of health care utilisation do not differentiate between desirable (e.g. preventative health care use) and undesirable (e.g. emergency hospital admission)

health care use. This strongly motivates the argument for exploring other outcomes, such as financial risk protection. Nonetheless, although improvements in health care utilisation and financial protection are signs of a well-functioning health insurance system, monitoring the associated improvements in population health are equally as important (Moreno-Serra and Smith, 2012). The SUSENAS data does not include health status indicators that would support this type of evaluation. A separate, more detailed health survey (known as RISKESDAS) was conducted in 2018 that can be integrated with the 2018 version of SUSENAS. However, RISKESDAS data is not routinely collected and there are some sampling differences between the two surveys, which makes integration more complicated. To monitor and evaluate health policies more effectively would require more comprehensive, frequent and accurate data collection, which has been recognised and some progress is being made in this direction (Asmanto, 2019).

In this thesis, alternative approaches to statistical inference for heterogeneous treatment effects have been introduced. If the CATE function is estimated reasonably well, it can be used to learn optimal policy rules and to identify subgroups with significantly different treatment effects (for example, by estimating average treatment effects (ATEs) for subgroups sorted on some scoring rule). However, it is important to make a distinction between these separate but closely related estimation and inferential tasks. ML algorithms, such as causal forests, can be used in the estimation of "individual" treatment effects for everyone in the sample (and if needed, out of sample), which can be used to classify the sample population into several groups based on the magnitude of these estimated effects. However, statistical inference for subgroups that are discovered by black-box ML algorithms can be particularly challenging. There is an extensive literature exploring alternative approaches to subgroup identification using a heterogeneous treatment effect estimation framework - see Loh et al. (2019) for a comparative review of various tree- and non-tree-based methods. More recent approaches by Chernozhukov et al. (2018b) and Imai and Li (2022) rely on generic ML algorithms for heterogeneous treatment effect estimation. Some differences exist between the two approaches but in general, statistical inference for subgroup identification does not require making any assumptions about the properties of the ML algorithms, but requires the random sampling of units and a

repeated sampling framework.

As machine learning methods become increasingly proficient at learning personalised treatment effects and policy rules, the process of decision making, as we currently know it, will fundamentally change over time. In the health and social sector, where resource constraints are prevalent, sophisticated algorithms are already being used in public policymaking (examples include Bhattacharya and Dupas (2012); Kube et al. (2019); McCarthy et al. (2017); Mukhopadhyay et al. (2016)). However, there are well-founded concerns about the ethical and fairness implications of replacing human decision making with data-driven methods, particularly for the allocation of shared public resources (O'Neil, 2017; Reddy et al., 2020). There are two problems at hand. The first is learning a rule that allocates policy according to observed characteristics that are morally questionable (e.g. age, sex, religion). This can be partially addressed by decision makers carefully selecting targeting criteria based on their preferences, as was demonstrated in Chapter 3, hence minimising the risk of a fully data-driven solution. The second is learning a rule that is based on a "black-box" algorithm that could inadvertently create adverse effects, for example, biases that perpetuate inequities (Cockx et al., 2019). Kube et al. (2019) highlight some examples from the literature where this has been the case. A proposed solution is for greater transparency and interpretability in the decision making process. The downside is that human interpretability is often achieved at the expense of finding an optimal solution. The question for decision makers is whether the loss in utility is worth the added explainability of the rule.

Applying machine learning tools to evaluations of large-scale policies requires the researcher to make a number of decisions. Some examples of decisions that are made in this work include: which model to use for a specific prediction or estimation task, which variables to control for in the nuisance models, which variables to include in the CATE models, how many times to perform sample-splitting, how many folds to use in K-fold cross-validation, and so on. The problem with using real-world data to inform these decisions is that the so-called "truth" (for example, the true nuisance or CATE functions) is unknown. If decisions are made that cannot be tested (such as parameter or model selection), and this results in a violation of the

underlying assumptions or a misspecification in the models being estimated, the accuracy and validity of the results could be affected. Simulation studies that are designed to be consistent with the specific settings of a real-world case study, can provide some guidance in this context (Morris et al., 2019). Future extensions of this work could simulate different scenarios where certain parameters are known, so that important decisions are backed by evidence, rather than an ad-hoc approach. An advantage of incorporating the super learner into the workflow is that it allows researchers to consider many different strategies for learning a function, and chooses the best strategy using its internal validation process. Even when the true (causal) parameter is unknown (for example, the optimal policy rule), it uses the machine learning logic of out-of-sample prediction, but optimised in a way that reduces the bias in the target parameter. For prediction problems, such as nuisance estimation, the cross-validation process is used in the usual way, targeting good out-of-sample predictions. However, the super learner itself requires a number of choices to be made to ensure that it is well-specified. Recent guidance from Phillips et al. (2022) can help in this regard, and many of their proposed practical considerations are incorporated in this work.

In the vast majority of the policy learning literature, the optimisation typically involves specifying an objective function to be maximised, capturing relevant population outcomes. In some cases, there is an imposed constraint of some form to reflect budget ceilings, but actual costs of the policy are rarely considered. In reality, policymakers usually make allocation decisions based on a cost-effectiveness analysis that considers the trade off between utility gains and incremental costs of the policy, often reported as a single population average. However, the cost quantity is likely to vary across individuals, according to their observed characteristics. In the JKN allocation problem, insuring a person with poor health, who also has good access to health care facilities will be costlier than insuring another person who is in better health, for example. The solution is to learn an optimal policy rule that maximises population outcomes and minimises costs, while also considering both the heterogeneity in treatment effectiveness and costs. Lakkaraju and Rudin (2017) learn cost-effective rules by constructing a decision list – a sequence of if-then-else rules – that maps an individual's observed covariate profile to a pol-

icy decision by sequentially optimising three separate parameters in the objective function: the doubly robust scores, the expected assessment cost of treatment and the expected treatment cost. The optimisation problem is modelled using a Markov Decision Process. Xu et al. (2020) propose a direct optimisation approach using outcome weighted classification, where the cost-effectiveness outcome is the optimisation parameter, and misclassification is measured by the net monetary benefit of the policy. These examples highlight potential directions for future evaluations of JKN and other system-level policies that combine optimal policy learning with economic evaluation.

Personalised policymaking has the potential to become an essential tool for the efficient allocation of health and social policies. The increasing availability of large-scale administrative and survey data, combined with the development of sophisticated machine learning algorithms, have made estimating granular (conditional) treatment effects feasible. The key question, however, is whether decision makers feel comfortable in basing their policy decisions on algorithms that are not always interpretable, and whether the associated equity and ethical implications outweigh the estimated welfare gains from a data-driven approach. Although these considerations are certainly important, it is worth noting that an allocation strategy that is based on data and estimates from a policy evaluation is more likely to achieve better outcomes than one that is not backed by evidence; which is a reflection of the current state of practice in health and social policy allocation.

Bibliography

Abadie, A. and Imbens, G. W. (2006). Large sample properties of matching estimators for average treatment effects. *Econometrica*, 74(1):235–267.

Abadie, A. and Imbens, G. W. (2011). Bias-corrected matching estimators for average treatment effects. *Journal of Business & Economic Statistics*, 29(1):1–11.

Abrevaya, J., Hsu, Y.-C., and Lieli, R. P. (2015). Estimating conditional average treatment effects. *Journal of Business & Economic Statistics*, 33(4):485–505.

Acharya, A., Vellakkal, S., Taylor, F., and others (2013). The impact of health insurance schemes for the informal sector in low-and middle-income countries: a systematic review. *The World Bank Research Observer*, 28(2):236–266.

Agustina, R., Dartanto, T., Sitompul, R., Susiloretni, K. A., Suparmi, Achadi, E. L., Taher, A., Wirawan, F., Sungkar, S., Sudarmono, P., Shankar, A. H., Thabrany, H., and Indonesian Health Systems Group (2019). Universal health coverage in Indonesia: concept, progress, and challenges. *The Lancet*, 393(10166):75–102.

Alatas, V., Banerjee, A., Hanna, R., Olken, B. A., and Tobias, J. (2012). Targeting the poor: evidence from a field experiment in Indonesia. *American Economic Review*, 102(4):1206–1240.

Alkenbrack, S. and Lindelow, M. (2015). The impact of community-based health insurance on utilization and out-of-pocket expenditures in Lao People's Democratic Republic. *Health Economics*, 24(4):379–399.

Allen, D. M. (1974). The relationship between variable selection and data agumentation and a method for prediction. *Technometrics*, 16(1):125–127.

Amit, Y. and Geman, D. (1997). Shape quantization and recognition with randomized trees. *Neural Computation*, 9(7):1545–1588.

Amram, M., Dunn, J., and Zhuo, Y. D. (2022). Optimal policy trees. *Machine Learning*, 111(7):2741–2768.

Andini, M., Ciani, E., de Blasio, G., D'Ignazio, A., and Salvestrini, V. (2018). Targeting with machine learning: an application to a tax rebate program in Italy. *Journal of Economic Behavior & Organization*, 156:86–102.

Angrist, J. D. and Pischke, J.-S. (2008). *Mostly Harmless Econometrics: An Empiricist's Companion*. Princeton University Press.

Anindya, K., Lee, J. T., McPake, B., Wilopo, S. A., Millett, C., and Carvalho, N. (2020). Impact of Indonesia's national health insurance scheme on inequality in access to maternal health services: a propensity score matched analysis. *Journal of Global Health*, 10(1):010429.

Asmanto, A. A. P. (2019). Harmonisation of Susenas and Riskesdas. Technical Report 43, National Team for the Acceleration of Poverty Reduction (TNP2K).

Aspinall, E. (2014). Health care and democratization in Indonesia. *Democratization*, 21(5):803–823.

Assmann, S. F., Pocock, S. J., Enos, L. E., and Kasten, L. E. (2000). Subgroup analysis and other (mis)uses of baseline data in clinical trials. *The Lancet*, 355(9209):1064–1069.

Atan, O., Zame, W. R., and van der Schaar, M. (2018). Learning optimal policies from observational data. `https://arxiv.org/abs/1802.08679`. Accessed: 2023-2-5.

Athey, S. (2015). Machine learning and causal inference for policy evaluation. In *Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pages 5–6.

Athey, S. (2017). Beyond prediction: using big data for policy problems. *Science*, 355(6324):483–485.

Athey, S. (2019). The impact of machine learning on economics. In *The Economics of Artificial Intelligence: An Agenda*, pages 507–547. University of Chicago Press.

Athey, S., Bayati, M., Imbens, G., and Qu, Z. (2019a). Ensemble methods for causal effects in panel data settings. *AEA Papers and Proceedings*, 109:65–70.

Athey, S. and Imbens, G. (2016). Recursive partitioning for heterogeneous causal effects. *Proceedings of the National Academy of Sciences of the United States of America*, 113(27):7353–7360.

Athey, S. and Imbens, G. W. (2015). Machine learning methods for estimating heterogeneous causal effects. *stat*, 1050(5):1–26.

Athey, S. and Imbens, G. W. (2017). The state of applied econometrics: causality and policy evaluation. *Journal of Economic Perspectives*, 31(2):3–32.

Athey, S. and Imbens, G. W. (2019). Machine learning methods that economists should know about. *Annual Review of Economics*, 11(1):685–725.

Athey, S., Tibshirani, J., and Wager, S. (2019b). Generalized random forests. *Annals of Statistics*, 47(2):1148–1178.

Athey, S. and Wager, S. (2019). Estimating treatment effects with causal forests: an application. *Observational Studies*, 5(2):37–51.

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

Bahamyirou, A., Schnitzer, M. E., Kennedy, E. H., Blais, L., and Yang, Y. (2022). Doubly robust adaptive LASSO for effect modifier discovery. *The International Journal of Biostatistics*, 18(2):307–327.

Banerjee, A., Finkelstein, A., Hanna, R., Olken, B. A., Ornaghi, A., and Sumarto, S. (2021). The challenges of universal health insurance in developing countries: experimental evidence from Indonesia's national health insurance. *American Economic Review*, 111(9):3035–3063.

Bang, H. and Robins, J. M. (2005). Doubly robust estimation in missing data and causal inference models. *Biometrics*, 61(4):962–973.

Basu, S., Kumbier, K., Brown, J. B., and Yu, B. (2018). Iterative random forests to discover predictive and stable high-order interactions. *Proceedings of the National Academy of Sciences of the United States of America*, 115(8):1943–1948.

Bell, R. M., Koren, Y., and Volinsky, C. (2010). All together now: a perspective on the Netflix prize. *Chance*, 23(1):24–29.

Bembom, O. and van der Laan, M. J. (2007). A practical illustration of the importance of realistic individualized treatment rules in causal inference. *Electronic Journal of Statistics*, 1:574–596.

Ben-Michael, E., Feller, A., and Rothstein, J. (2020). Varying impacts of letters of recommendation on college admissions: approximate balancing weights for subgroup effects in observational studies. `https://arxiv.org/abs/2008.04394`. Accessed: 2023-2-5.

Bernal, N., Carpio, M. A., and Klein, T. J. (2017). The effects of access to health insurance: evidence from a regression discontinuity design in Peru. *Journal of Public Economics*, 154:122–136.

Bertrand, Crépon, M., Marguerie, B., Premand, A., and Patrick (2017). Contemporaneous and post-program impacts of a public works program: evidence from Côte d'Ivoire. *World Bank*.

Bertsimas, D., Dunn, J., and Mundru, N. (2019). Optimal prescriptive trees. *INFORMS Journal on Optimization*, 1(2):164–183.

Bertsimas, D. and Kallus, N. (2020). From predictive to prescriptive analytics. *Management Science*, 66(3):1025–1044.

Bhandari, A. and Wagner, T. (2006). Self-reported utilization of health care services: improving measurement and accuracy. *Medical Care Research and Review*, 63(2):217–235.

Bhattacharya, D. and Dupas, P. (2012). Inferring welfare maximizing treatment assignment under budget constraints. *Journal of Econometrics*, 167(1):168–196.

Boerma, T., Eozenou, P., Evans, D., Evans, T., Kieny, M.-P., and Wagstaff, A. (2014). Monitoring progress towards universal health coverage at country and global levels. *PLoS Medicine*, 11(9):e1001731.

Bonetti, M. and Gelber, R. D. (2004). Patterns of treatment effects in subsets of patients in clinical trials. *Biostatistics*, 5(3):465–481.

Bookwalter, J. T., Fuller, B. S., and Dalenberg, D. R. (2006). Do household heads speak for the household? A research note. *Social Indicators Research*, 79(3):405–419.

Brand, J. E., Xu, J., Koch, B., and Geraldo, P. (2021). Uncovering sociological effect heterogeneity using tree-based machine learning. *Sociological Methodology*, 51(2):189–223.

Breiman, L. (1996a). Out-of-bag estimation. Technical report, University of California, Berkeley.

Breiman, L. (1996b). Stacked regressions. *Machine Learning*, 24(1):49–64.

Breiman, L. (2001). Random forests. *Machine Learning*, 45(1):5–32.

Breiman, L., Friedman, J. H., Olshen, R. A., and Stone, C. J. (1984). *Classification and Regression Trees*. Wadsworth Books.

Brookes, S. T., Whitely, E., Egger, M., Smith, G. D., Mulheran, P. A., and Peters, T. J. (2004). Subgroup analyses in randomized trials: risks of subgroup-specific analyses; power and sample size for the interaction test. *Journal of Clinical Epidemiology*, 57(3):229–236.

Brookes, S. T., Whitley, E., Peters, T. J., Mulheran, P. A., Egger, M., and Davey Smith, G. (2001). Subgroup analyses in randomised controlled trials: quantifying the risks of false-positives and false-negatives. *Health Technology Assessment*, 5(33):1–56.

Bühlmann, P. and Yu, B. (2002). Analyzing bagging. *Annals of Statistics*, 30(4):927–961.

Carvalho, C., Feller, A., Murray, J., Woody, S., and Yeager, D. (2019). Assessing treatment effect variation in observational studies: results from a data challenge. *Observational Studies*, 5(2):21–35.

Chen, J.-E. and Hsiang, C.-W. (2019). Causal random forests model using instrumental variable quantile regression. *Econometrics*, 7(4):49.

Chen, J.-E., Huang, C.-H., and Tien, J.-J. (2021). Debiased/Double machine learning for instrumental variable quantile regressions. *Econometrics*, 9(2):15.

Chen, Y. and Jin, G. Z. (2012). Does health insurance coverage lead to better health and educational outcomes? Evidence from rural China. *Journal of Health Economics*, 31(1):1–14.

Chernozhukov, V., Chetverikov, D., Demirer, M., Duflo, E., Hansen, C., and Newey, W. (2017). Double/Debiased/Neyman machine learning of treatment effects. *American Economic Review*, 107(5):261–265.

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

Chernozhukov, V., Demirer, M., Duflo, E., and Fernández-Val, I. (2018b). Generic machine learning inference on heterogeneous treatment effects in randomized experiments, with an application to immunization in India. Technical Report 24678, National Bureau of Economic Research.

Chernozhukov, V. and Hansen, C. (2005). An IV model of quantile treatment effects. *Econometrica*, 73(1):245–261.

Chipman, H. A., George, E. I., and McCulloch, R. E. (2010). BART: Bayesian additive regression trees. *Annals of Applied Statistics*, 4(1):266–298.

Clarke, D., Romano, J. P., and Wolf, M. (2020). The Romano–Wolf multiple-hypothesis correction in stata. *The Stata Journal*, 20(4):812–843.

Cockx, B., Lechner, M., and Bollens, J. (2019). Priority to unemployed immigrants? A causal machine learning evaluation of training in Belgium. `https://arxiv.org/abs/1912.12864`. Accessed: 2023-2-5.

Cook, D. I., Gebski, V. J., and Keech, A. C. (2004). Subgroup analysis in clinical trials. *Medical Journal of Australia*, 180(6):289–291.

Coyle, J. R. (2017). *Computational Considerations for Targeted Learning*. PhD thesis. University of California, Berkeley.

Crump, R. K., Hotz, V. J., Imbens, G. W., and Mitnik, O. A. (2009). Dealing with limited overlap in estimation of average treatment effects. *Biometrika*, 96(1):187–199.

Cui, Y., Kosorok, M. R., Sverdrup, E., Wager, S., and Zhu, R. (2020). Estimating heterogeneous treatment effects with right-censored data via causal survival forests. `https://arxiv.org/abs/2001.09887`. Accessed: 2023-2-5.

Dahabreh, I. J., Hayward, R., and Kent, D. M. (2016). Using group data to treat individuals: understanding heterogeneous treatment effects in the age of precision medicine and patient-centred evidence. *International Journal of Epidemiology*, 45(6):2184–2193.

Daoud, A. and Johansson, F. (2019). Estimating treatment heterogeneity of International Monetary Fund programs on child poverty with generalized random forest. `https://doi.org/10.31235/osf.io/awfjt`. Accessed: 2023-2-5.

Dartanto, T., Halimatussadiah, A., Rezki, J. F., Nurhasana, R., Siregar, C. H., Bintara, H., Usman, Pramono, W., Sholihah, N. K., Yuan, E. Z. W., and Soeharno, R. (2020a). Why do informal sector workers not pay the premium regularly? Evidence from the national health insurance system in Indonesia. *Applied Health Economics and Health Policy*, 18(1):81–96.

Dartanto, T., Pramono, W., Lumbanraja, A. U., Siregar, C. H., Bintara, H., Sholihah, N. K., and Usman (2020b). Enrolment of informal sector workers in the national health insurance system in Indonesia: a qualitative study. *Heliyon*, 6(11):e05316.

Dartanto, T., Rezki, J. F., Pramono, W., Siregar, C. H., Bintara, U., and Bintara, H. (2016). Participation of informal sector workers in Indonesia's national health insurance system. *Journal of Southeast Asian Economies*, 33(3):317–342.

Davis, J. M. V. and Heller, S. B. (2017). Using causal forests to predict treatment heterogeneity: an application to summer jobs. *American Economic Review*, 107(5):546–550.

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

Díaz, I. (2020). Machine learning in the estimation of causal effects: targeted minimum loss-based estimation and double/debiased machine learning. *Biostatistics*, 21(2):353–358.

Dow, W. H. and Schmeer, K. K. (2003). Health insurance and child mortality in Costa Rica. *Social Science & Medicine*, 57(6):975–986.

Dutta, A., Ward, K., Setiawan, E., and Prabhakaran, S. (2020). Fiscal space for health in Indonesia: public sector opportunities and constraints in achieving the goals of Indonesia's mid-term development plan (RPJMN) 2020–2024. Technical report, Jakarta: Kementerian PPN/Bappenas.

Efron, B. (2014). Estimation and accuracy after model selection. *Journal of the American Statistical Association*, 109(507):991–1007.

Einav, L. and Levin, J. (2014). Economics in the age of big data. *Science*, 346(6210):1243089.

El-Sayed, A. M., Vail, D., and Kruk, M. E. (2018). Ineffective insurance in lower and middle income countries is an obstacle to universal health coverage. *Journal of Global Health*, 8(2):020402.

Erlangga, D., Ali, S., and Bloor, K. (2019a). The impact of public health insurance on healthcare utilisation in Indonesia: evidence from panel data. *International Journal of Public Health*, 64(4):603–613.

Erlangga, D., Suhrcke, M., Ali, S., and Bloor, K. (2019b). The impact of public health insurance on health care utilisation, financial protection and health status in low- and middle-income countries: A systematic review. *PloS one*, 14(11):e0225237.

Escobar, M.-L., Griffin, C. C., and Paul Shaw, R. (2011). *The Impact of Health Insurance in Low- and Middle-Income Countries*. Brookings Institution Press.

Fan, Q., Hsu, Y.-C., Lieli, R. P., and Zhang, Y. (2022). Estimation of conditional average treatment effects with high-dimensional data. *Journal of Business & Economic Statistics*, 40(1):313–327.

Fan, Q. and Wu, Y. (2020). Endogenous treatment effect estimation with some invalid and irrelevant instruments. `https://arxiv.org/abs/2006.14998`. Accessed: 2023-2-5.

Filho, W. L., Wall, T., Azul, A. M., Brandli, L., and Özuyar, P. G. (2020). *Good Health and Well-Being*. Encyclopedia of the UN Sustainable Development Goals. Springer Publishing.

Fink, G., Robyn, P. J., Sié, A., and Sauerborn, R. (2013). Does health insurance improve health?: Evidence from a randomized community-based insurance rollout in rural Burkina Faso. *Journal of Health Economics*, 32(6):1043–1056.

Foster, D. J. and Syrgkanis, V. (2019). Orthogonal statistical learning. `https://arxiv.org/abs/1901.09036`. Accessed: 2023-2-5.

Foster, J. C., Taylor, J. M. G., and Ruberg, S. J. (2011). Subgroup identification from randomized clinical trial data. *Statistics in Medicine*, 30(24):2867–2880.

Gail, M. and Simon, R. (1985). Testing for qualitative interactions between treatment effects and patient subsets. *Biometrics*, 41(2):361–372.

Garrett, L., Chowdhury, A. M. R., and Pablos-Méndez, A. (2009). All for universal health coverage. *The Lancet*, 374(9697):1294–1299.

Ghebreyesus, T. A. (2017). All roads lead to universal health coverage. *The Lancet Global Health*, 5(9):e839–e840.

Giedion, U., Andrés Alfonso, E., and Díaz, Y. (2013). The impact of universal coverage schemes in the developing world : a review of the existing evidence. Technical report, World Bank.

Green, D. P. and Kern, H. L. (2012). Modeling heterogeneous treatment effects in survey experiments with bayesian additive regression trees. *Public Opinion Quarterly*, 76(3):491–511.

Greene, W. H. (2000). *Econometric Analysis, 4th Edition*. Prentice Hall, Englewood Cliffs.

Gruber, S. and van der Laan, M. J. (2010). A targeted maximum likelihood estimator of a causal effect on a bounded continuous outcome. *The International Journal of Biostatistics*, 6(1):Article 26.

Hahn, J. (1998). On the role of the propensity score in efficient semiparametric estimation of average treatment effects. *Econometrica*, 66(2):315–331.

Hahn, R. P., Murray, J. S., and Carvalho, C. M. (2020). Bayesian regression tree models for causal inference: regularization, confounding, and heterogeneous effects (with discussion). *Bayesian Analysis*, 15(3):965–1056.

Harimurti, P., Pambudi, E., Pigazzini, A., and Tandon, A. (2013). The nuts and bolts of Jamkesmas: Indonesia's government-financed health coverage program for the poor and near-poor. Technical report, World Bank.

Hastie, T., Friedman, J., and Tibshirani, R. (2009). *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. Springer Publishing.

Heckman, J. J. (2001). Micro data, heterogeneity, and the evaluation of public policy: Nobel lecture. *Journal of Political economy*, 109(4):673–748.

Heckman, J. J., Ichimura, H., Smith, J., and Todd, P. (1998). Characterizing selection bias using experimental data. *Econometrica*, 66(5):1017–1098.

Hill, J. and Su, Y.-S. (2013). Assessing lack of common support in causal inference using bayesian nonparametrics: implications for evaluating the effect of breastfeeding on children's cognitive outcomes. *Annals of Applied Statistics*, 7(3):1386–1420.

Hill, J. L. (2011). Bayesian nonparametric modeling for causal inference. *Journal of Computational and Graphical Statistics*, 20(1):217–240.

Holland, P. W. (1986). Statistics and causal inference. *Journal of the American Statistical Association*, 81(396):945–960.

Hothorn, T., Lausen, B., Benner, A., and Radespiel-Tröger, M. (2004). Bagging survival trees. *Statistics in Medicine*, 23(1):77–91.

Imai, K. and Li, M. L. (2022). Statistical inference for heterogeneous treatment effects discovered by generic machine learning in randomized experiments.

Imai, K. and Ratkovic, M. (2013). Estimating treatment effect heterogeneity in randomized program evaluation. *Annals of Applied Statistics*, 7(1):443–470.

Imai, K. and Strauss, A. (2011). Estimation of heterogeneous treatment effects from randomized experiments, with application to the optimal planning of the Get-Out-the-Vote campaign. *Political Analysis*, 19(1):1–19.

Imbens, G. W. and Rubin, D. B. (2015). *Causal Inference for Statistics, Social, and Biomedical Sciences*. Cambridge University Press.

Imbens, G. W. and Wooldridge, J. M. (2009). Recent developments in the econometrics of program evaluation. *Journal of Economic Literature*, 47(1):5–86.

Jacob, D. (2020). Cross-fitting and averaging for machine learning estimation of heterogeneous treatment effects. `https://arxiv.org/abs/2007.02852`. Accessed: 2023-2-5.

Jacob, D. (2021). CATE meets ML. *Digital Finance*, 3(2):99–148.

Johar, M. (2009). The impact of the Indonesian health card program: a matching estimator approach. *Journal of Health Economics*, 28(1):35–53.

Johar, M., Soewondo, P., Pujisubekti, R., Satrio, H. K., and Adji, A. (2018). Inequality in access to health care, health insurance and the role of supply factors. *Social Science & Medicine*, 213:134–145.

Johar, M., Soewondo, P., Pujisubekti, R., Satrio, H. K., and Adji, A. (2019). In data we trust? An analysis of Indonesian socioeconomic survey data. *Bulletin of Indonesian Economic Studies*, 55(1):61–82.

Johnson, M., Cao, J., and Kang, H. (2022). Detecting heterogeneous treatment effects with instrumental variables and application to the Oregon health insurance experiment. *Annals of Applied Statistics*, 16(2).

Jones, A. M. (2000). Health econometrics. In *Handbook of Health Economics*, volume 1, pages 265–344. Elsevier.

Kallus, N. (2017). Recursive partitioning for personalization using observational data. In *Proceedings of the 34th International Conference on Machine Learning*, volume 70 of *Proceedings of Machine Learning Research*, pages 1789–1798. PMLR.

Kallus, N. (2018). Balanced policy evaluation and learning. In *Proceedings of the 32nd International Conference on Neural Information Processing Systems*, NIPS'18, pages 8909–8920, Red Hook, NY, USA. Curran Associates Inc.

Kallus, N. (2020). More efficient policy learning via optimal retargeting. *Journal of the American Statistical Association*, 116(534):646–658.

Kallus, N., Mao, X., and Zhou, A. (2019). Interval estimation of individual-level accusal effects under unobserved confounding. In *Proceedings of the Twenty-Second International Conference on Artificial Intelligence and Statistics*, volume 89 of *Proceedings of Machine Learning Research*, pages 2281–2290. PMLR.

Kallus, N. and Zhou, A. (2018a). Confounding-robust policy improvement. `https://arxiv.org/abs/1805.08593`. Accessed: 2023-2-5.

Kallus, N. and Zhou, A. (2018b). Policy evaluation and optimization with continuous treatments. In *Proceedings of the Twenty-First International Conference on Artificial Intelligence and Statistics*, volume 84 of *Proceedings of Machine Learning Research*, pages 1243–1251. PMLR.

Kallus, N. and Zhou, A. (2021). Minimax-optimal policy learning under unobserved confounding. *Management Science*, 67(5):2870–2890.

Kennedy, E. H. (2020). Optimal doubly robust estimation of heterogeneous causal effects. `https://arxiv.org/abs/2004.14497`. Accessed: 2023-2-5.

Kent, D. M., Steyerberg, E., and van Klaveren, D. (2018). Personalized evidence based medicine: predictive approaches to heterogeneous treatment effects. *BMJ*, 363:k4245.

Kim, T. J., Vonneilich, N., Lüdecke, D., and von dem Knesebeck, O. (2017). Income, financial barriers to health care and public health expenditure: a multilevel analysis of 28 countries. *Social Science & Medicine*, 176:158–165.

Kitagawa, T. and Tetenov, A. (2018). Who should be treated? Empirical welfare maximization methods for treatment choice. *Econometrica*, 86(2):591–616.

Kleinberg, J., Ludwig, J., Mullainathan, S., and Obermeyer, Z. (2015). Prediction policy problems. *American Economic Review*, 105(5):491–495.

Knaus, M. C. (2022). Double machine learning-based programme evaluation under unconfoundedness. *The Econometrics Journal*, 25(3):602–627.

Knaus, M. C., Lechner, M., and Strittmatter, A. (2021). Machine learning estimation of heterogeneous causal effects: empirical monte carlo evidence. *The Econometrics Journal*, 24(1):134–161.

Knaus, M. C., Lechner, M., and Strittmatter, A. (2022). Heterogeneous employment effects of job search programs. *The Journal of Human Resources*, 57(2):597–636.

Kosorok, M. R. and Laber, E. B. (2019). Precision medicine. *Annual Review of Statistics and its Application*, 6(1):263–286.

Kreif, N. and DiazOrdaz, K. (2019). Machine learning in policy evaluation: New tools for causal inference. In Jones, A. M., editor, *The Oxford Encyclopedia of Health Economics*.

Kreif, N., DiazOrdaz, K., Moreno-Serra, R., Mirelman, A., Hidayat, T., and Suhrcke, M. (2022). Estimating heterogeneous policy impacts using causal machine learning: a case study of health insurance reform in Indonesia. *Health Services & Outcomes Research Methodology*, 22(2):192–227.

Kruse, I., Pradhan, M., and Sparrow, R. (2012). Marginal benefit incidence of public health spending: evidence from Indonesian sub-national data. *Journal of Health Economics*, 31(1):147–157.

Kube, A., Das, S., and Fowler, P. J. (2019). Allocating interventions based on predicted outcomes: a case study on homelessness services. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 33 of *33*, pages 622–629.

Künzel, S. R., Sekhon, J. S., Bickel, P. J., and Yu, B. (2019a). Metalearners for estimating heterogeneous treatment effects using machine learning. volume 116 of *116*, pages 4156–4165.

Künzel, S. R., Walter, S. J. S., and Sekhon, J. S. (2019b). Causaltoolbox—Estimator stability for heterogeneous treatment effects. *Observational Studies*, 5(2):105–117.

Lagomarsino, G., Garabrant, A., Adyas, A., Muga, R., and Otoo, N. (2012). Moving towards universal health coverage: health insurance reforms in nine developing countries in Africa and Asia. *The Lancet*, 380(9845):933–943.

Lakkaraju, H. and Rudin, C. (2017). Learning cost-effective and interpretable treatment regimes. In Singh, A. and Zhu, J., editors, *Proceedings of the 20th International Conference on Artificial Intelligence and Statistics*, volume 54 of *Proceedings of Machine Learning Research*, pages 166–175. PMLR.

Lechner, M. (2018). Modified causal forests for estimating heterogeneous causal effects. `https://arxiv.org/abs/1812.09487`. Accessed: 2023-2-5.

Lee, B. K., Lessler, J., and Stuart, E. A. (2010). Improving propensity score weighting using machine learning. *Statistics in Medicine*, 29(3):337–346.

Lee, M.-J. (2017). Extensive and intensive margin effects in sample selection models: Racial effects on wages. *Journal of the Royal Statistical Society: Series A*, 180(3):817–839.

Lee, S., Okui, R., and Whang, Y.-J. (2017). Doubly robust uniform confidence band for the conditional average treatment effect function. *Journal of Applied Econometrics*, 32(7):1207–1225.

Li, F., Morgan, K. L., and Zaslavsky, A. M. (2018). Balancing covariates via propensity score weighting. *Journal of the American Statistical Association*, 113(521):390–400.

Li, F., Thomas, L. E., and Li, F. (2019). Addressing extreme propensity scores via the overlap weights. *American Journal of Epidemiology*, 188(1):250–257.

Limwattananon, S., Neelsen, S., O'Donnell, O., Prakongsai, P., Tangcharoensathien, V., van Doorslaer, E., and Vongmongkol, V. (2015). Universal coverage with supply-side reform: the impact on medical expenditure risk and utilization in Thailand. *Journal of Public Economics*, 121:79–94.

Lin, Y. and Jeon, Y. (2006). Random forests and adaptive nearest neighbors. *Journal of the American Statistical Association*, 101(474):578–590.

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

Liu, Y., Halev, A., and Liu, X. (2021). Policy learning with constraints in model-free reinforcement learning: a survey. In *Proceedings of the Thirtieth International Joint Conference on Artificial Intelligence*. International Joint Conferences on Artificial Intelligence Organization.

Loh, W.-Y., Cao, L., and Zhou, P. (2019). Subgroup identification for precision medicine: A comparative review of 13 methods. *Wiley interdisciplinary reviews. Data mining and knowledge discovery*, 9(5):e1326.

Luedtke, A. R. and van der Laan, M. J. (2016a). Optimal individualized treatments in resource-limited settings. *The International Journal of Biostatistics*, 12(1):283–303.

Luedtke, A. R. and van der Laan, M. J. (2016b). Super-learning of an optimal dynamic treatment rule. *The International Journal of Biostatistics*, 12(1):305–332.

Mackenbach, J. P. (2003). Tackling inequalities in health: the need for building a systematic evidence base. *Journal of Epidemiology and Community Health*, 57(3):162.

Mahendradhata, Y., Trisnantoro, L., Listyadewi, S., Soewondo, P., Marthias, T., Harimurti, P., and Prawira, J. (2017). The Republic of Indonesia health system review. Technical report, WHO Regional Office for South-East Asia.

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

Mardiansjah, F. H., Rahayu, P., and Rukmana, D. (2021). New patterns of urbanization in Indonesia: emergence of non-statutory towns and new extended urban regions. *Environment and Urbanization ASIA*, 12(1):11–26.

Maulana, N., Soewondo, P., Adani, N., Limasalle, P., and Pattnaik, A. (2022). How Jaminan Kesehatan Nasional (JKN) coverage influences out-of-pocket (OOP) payments by vulnerable populations in Indonesia. *PLOS Global Public Health*, 2(7).

Mboi, N. (2015). Indonesia: on the way to universal health care. *Health Systems and Reform*, 1(2):91–97.

McCarthy, S. M., Vayanos, P., and Tambe, M. (2017). Staying ahead of the game: adaptive robust optimization for dynamic allocation of threat screening resources. In *Proceedings of the Twenty-Sixth International Joint Conference on Artificial Intelligence*. International Joint Conferences on Artificial Intelligence Organization.

Mebratie, A. D., Sparrow, R., Yilma, Z., Abebaw, D., Alemu, G., and Bedi, A. S. (2019). The impact of Ethiopia's pilot community based health insurance scheme on healthcare utilization and cost of care. *Social Science & Medicine*, 220:112–119.

Meinshausen, N. (2006). Quantile regression forests. *Journal of Machine Learning Research*, 7(Jun):983–999.

Mihaylova, B., Briggs, A., O'Hagan, A., and Thompson, S. G. (2011). Review of statistical methods for analysing healthcare resources and costs. *Health Economics*, 20(8):897–916.

Miller, G., Pinto, D., and Vera-Hernández, M. (2013). Risk protection, service use, and health outcomes under Colombia's health insurance program for the poor. *American Economic Journal: Applied Economics*, 5(4):61–91.

Montoya, L., Skeem, J., van der Laan, M. J., and Petersen, M. (2021). Performance and application of estimators for the value of an optimal dynamic treatment rule. `https://arxiv.org/abs/2101.12333`. Accessed: 2023-2-5.

Montoya, L. M., van der Laan, M. J., Luedtke, A. R., Skeem, J. L., Coyle, J. R., and Petersen, M. L. (2022). The optimal dynamic treatment rule superlearner: considerations, performance, and application to criminal justice interventions. *The International Journal of Biostatistics*.

Moreno-Serra, R. and Smith, P. C. (2012). Does progress towards universal health coverage improve population health? *The Lancet*, 380(9845):917–923.

Moreno-Serra, R. and Smith, P. C. (2015). Broader health coverage is good for the nation's health: evidence from country level panel data. *Journal of the Royal Statistical Society: Series A*, 178(1):101–124.

Morris, T. P., White, I. R., and Crowther, M. J. (2019). Using simulation studies to evaluate statistical methods. *Statistics in Medicine*, 38(11):2074–2102.

Mukhopadhyay, A., Zhang, C., Vorobeychik, Y., Tambe, M., Pence, K., and Speer, P. (2016). Optimal allocation of police patrol resources using a continuous-time crime model. In *Decision and Game Theory for Security (GameSec 2016)*, pages 139–158.

Mullainathan, S. and Spiess, J. (2017). Machine learning: an applied econometric approach. *Journal of Economic Perspectives*, 31(2):87–106.

Mulyanto, J., Kringos, D. S., and Kunst, A. E. (2019a). The evolution of income-related inequalities in healthcare utilisation in Indonesia, 1993-2014. *PloS One*, 14(6):e0218519.

Mulyanto, J., Kringos, D. S., and Kunst, A. E. (2019b). Socioeconomic inequalities in healthcare utilisation in Indonesia: a comprehensive survey-based overview. *BMJ Open*, 9(7):e026164.

Mulyanto, J., Kunst, A. E., and Kringos, D. S. (2019c). Geographical inequalities in healthcare utilisation and the contribution of compositional factors: a multilevel analysis of 497 districts in Indonesia. *Health & Place*, 60:102236.

Murphy, S. A. (2003). Optimal dynamic treatment regimes. *Journal of the Royal Statistical Society: Series B*, 65(2):331–355.

Muttaqien, M., Setiyaningsih, H., Aristianti, V., Coleman, H. L. S., Hidayat, M. S., Dhanalvin, E., Siregar, D. R., Mukti, A. G., and Kok, M. O. (2021). Why did informal sector workers stop paying for health insurance in Indonesia? Exploring enrollees' ability and willingness to pay. *PloS One*, 16(6):e0252708.

National Team for the Acceleration of Poverty Reduction (Indonesia) (2015). The road to national health insurance (JKN). Technical report.

Neugebauer, R. and van der Laan, M. J. (2005). Why prefer double robust estimators in causal inference? *Journal of Statistical Planning and Inference*, 129(1):405–426.

Newey, W. K. and Robins, J. R. (2018). Cross-fitting and fast remainder rates for semiparametric estimation. https://arxiv.org/abs/1801.09138. Accessed: 2023-2-5.

Nie, X., Lu, C., and Wager, S. (2019). Nonparametric heterogeneous treatment effect estimation in repeated cross sectional designs. https://arxiv.org/abs/1905.11622. Accessed: 2023-2-5.

Nie, X. and Wager, S. (2021). Quasi-oracle estimation of heterogeneous treatment effects. *Biometrika*, 108(2):299–319.

Nugraheni, W. P., Mubasyiroh, R., and Hartono, R. K. (2020). The influence of Jaminan Kesehatan Nasional (JKN) on the cost of delivery services in Indonesia. *PloS One*, 15(7):e0235176.

OECD (2019). Social protection system review of Indonesia. Technical report, OECD Development Pathways, OECD Publishing, Paris.

O'Neil, C. (2017). *Weapons of math destruction: How big data increases inequality and threatens democracy*. Penguin Books.

O'Neill and Weeks, M. (2018). Causal tree estimation of heterogeneous household response to time-of-use electricity pricing schemes. https://arxiv.org/abs/1810.09179. Accessed: 2023-2-5.

Oprescu, M., Syrgkanis, V., and Wu, Z. S. (2019). Orthogonal random forest for causal inference. In *Proceedings of the 36th International Conference on Machine Learning, ICML 2019*, pages 8655–8696. International Machine Learning Society.

Panch, T., Mattie, H., and Atun, R. (2019). Artificial intelligence and algorithmic bias: implications for health systems. *Journal of Global Health*, 9(2):010318.

Petticrew, M., Tugwell, P., Kristjansson, E., Oliver, S., Ueffing, E., and Welch, V. (2012). Damned if you do, damned if you don't: subgroup analysis and equity. *Journal of Epidemiology and Community Health*, 66(1):95–98.

Phillips, R. V., van der Laan, M. J., Lee, H., and Gruber, S. (2022). Practical considerations for specifying a super learner. https://arxiv.org/abs/2204.06139. Accessed: 2023-2-5.

Pohlmeier, W. and Ulrich, V. (1995). An econometric model of the two-part decisionmaking process in the demand for health care. *Journal of Human Resources*, 30(2):339.

Polley, E. C. and van der Laan, M. J. (2010). Super learner in prediction. Technical report, University of California, Berkeley.

Powers, S., Qian, J., Jung, K., Schuler, A., Shah, N. H., Hastie, T., and Tibshirani, R. (2018). Some methods for heterogeneous treatment effect estimation in high dimensions. *Statistics in Medicine*, 37(11):1767–1787.

Prabhakaran, S., Dutta, A., Fagan, T., and Ginivan, M. (2019). Financial sustainability of Indonesia's Jaminan Kesehatan Nasional: Performance, prospects, and policy options. Technical report, Washington, DC: Palladium, Health Policy Plus, and Jakarta, Indonesia: National Team for the Acceleration of Poverty Reduction (TNP2K).

Pratiwi, A. B., Setiyaningsih, H., Kok, M. O., Hoekstra, T., Mukti, A. G., and Pisani, E. (2021). Is Indonesia achieving universal health coverage? Secondary analysis of national data on insurance coverage, health spending and service availability. *BMJ Open*, 11(10):e050565.

Qian, M. and Murphy, S. A. (2011). Performance guarantees for individualized treatment rules. *Annals of Statistics*, 39(2):1180–1210.

Reddy, S., Allan, S., Coghlan, S., and Cooper, P. (2020). A governance model for the application of AI in health care. *Journal of the American Medical Informatics Association*, 27(3):491–497.

Robins, J., Sued, M., Lei-Gomez, Q., and Rotnitzky, A. (2007). Comment: Performance of double-robust estimators when" inverse probability" weights are highly variable. *Statistical Science*, 22(4):544–559.

Robins, J. M. (2004). Optimal structural nested models for optimal sequential decisions. In Lin, D. Y. and Heagerty, P. J., editors, *Proceedings of the Second Seattle Symposium in Biostatistics: Analysis of Correlated Data*, pages 189–326.

Robins, J. M. and Rotnitzky, A. (1995). Semiparametric efficiency in multivariate regression models with missing data. *Journal of the American Statistical Association*, 90(429):122–129.

Robins, J. M., Rotnitzky, A., and Zhao, L. P. (1994). Estimation of regression coefficients when some regressors are not always observed. *Journal of the American Statistical Association*, 89(427):846–866.

Robinson, P. M. (1988). Semiparametric econometrics: a survey. *Journal of Applied Econometrics*, 3(1):35–51.

Rose, S. (2013). Mortality risk score prediction in an elderly population using machine learning. *American Journal of Epidemiology*, 177(5):443–452.

Rosenbaum, P. R. (2002). *Observational Studies*. Springer Series in Statistics. Springer Publishing.

Rosenbaum, P. R. and Rubin, D. B. (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika*, 70(1):41–55.

Rosenbaum, P. R. and Rubin, D. B. (1984). Reducing bias in observational studies using subclassification on the propensity score. *Journal of the American Statistical Association*, 79(387):516–524.

Rubin, D. B. (1974). Estimating causal effects of treatments in randomized and nonrandomized studies. *Journal of Educational Psychology*, 66(5):688–701.

Rubin, D. B. (2009). Should observational studies be designed to allow lack of balance in covariate distributions across treatment groups? *Statistics in Medicine*, 28(9):1420–1423.

Saksena, P., Hsu, J., and Evans, D. B. (2014). Financial risk protection and universal health coverage: evidence and measurement challenges. *PLoS Medicine*, 11(9):e1001701.

Sambodo, N. P., Van Doorslaer, E., Pradhan, M., and Sparrow, R. (2021). Does geographic spending variation exacerbate healthcare benefit inequality? A benefit incidence analysis for indonesia. *Health Policy and Planning*, 36(7):1129–1139.

Semenova, V. and Chernozhukov, V. (2021). Debiased machine learning of conditional average treatment effects and other causal functions. *The Econometrics Journal*, 24(2):264–289.

Semenova, V., Goldman, M., Chernozhukov, V., and Taddy, M. (2017). Estimation and inference about heterogeneous treatment effects in high-dimensional dynamic panels. `https://arxiv.org/abs/1712.09988`. Accessed: 2023-2-5.

Shalit, U., Johansson, F. D., and Sontag, D. (2017). Estimating individual treatment effect: generalization bounds and algorithms. In Precup, D. and Teh, Y. W., editors, *Proceedings of the 34th International Conference on Machine Learning*, volume 70 of *Proceedings of Machine Learning Research*, pages 3076–3085. PMLR.

Sheth, K. (2014). Evaluating health-seeking behaviour, utilization of care, and health risk: evidence from a community based insurance model in India. Technical Report 36, International Labour Office.

Sparrow, R., Budiyati, S., Yumna, A., Warda, N., Suryahadi, A., and Bedi, A. S. (2017). Sub-national health care financing reforms in Indonesia. *Health Policy and Planning*, 32(1):91–101.

Sparrow, R., Suryahadi, A., and Widyanti, W. (2013). Social health insurance for the poor: targeting and impact of Indonesia's Askeskin programme. *Social Science & Medicine*, 96:264–271.

Staub, K. E. (2014). A causal interpretation of extensive and intensive margin effects in generalized tobit models. *The Review of Economics and Statistics*, 96(2):371–375.

Stone, C. J. (1980). Optimal rates of convergence for nonparametric estimators. *Annals of Statistics*, 8(6):1348–1360.

Stone, M. (1974). Cross-validatory choice and assessment of statistical predictions. *Journal of the Royal Statistical Society: Series B*, 36(2):111–147.

Strittmatter, A. (2019). What is the value added by using causal machine learning methods in a welfare experiment evaluation? `https://arxiv.org/abs/1812.06533`. Accessed: 2023-2-5.

Stürmer, T., Webster-Clark, M., Lund, J. L., Wyss, R., Ellis, A. R., Lunt, M., Rothman, K. J., and Glynn, R. J. (2021). Propensity score weighting and trimming strategies for reducing variance and bias of treatment effect estimates: a simulation study. *American Journal of Epidemiology*, 190(8):1659–1670.

Surowiecki, J. (2005). *The wisdom of crowds: Why the many are smarter than the few and how collective wisdom shapes business, economies, societies, and nations*. Doubleday & Co.

Sverdrup, E., Kanodia, A., Zhou, Z., Athey, S., and Wager, S. (2020). policytree: Policy learning via doubly robust empirical welfare maximization over trees. *Journal of Open Source Software*, 5(50):2232.

Taddy, M., Gardner, M., Chen, L., and Draper, D. (2016). A nonparametric bayesian analysis of heterogenous treatment effects in digital experimentation. *Journal of Business & Economic Statistics*, 34(4):661–672.

Taylor, L., Schroeder, R., and Meyer, E. (2014). Emerging practices and perspectives on big data analysis in economics: bigger and better or more of the same? *Big Data & Society*, 1(2):2053951714536877.

Tian, L., Alizadeh, A. A., Gentles, A. J., and Tibshirani, R. (2014). A simple method for estimating interactions between a treatment and a large number of covariates. *Journal of the American Statistical Association*, 109(508):1517–1532.

van der Laan, M. J. (2006). Statistical inference for variable importance. *The International Journal of Biostatistics*, 2(1).

van Der Laan, M. J. and Dudoit, S. (2003). Unified cross-validation methodology for selection among estimators and a general cross-validated adaptive epsilon-net estimator: finite sample oracle inequalities and examples. U.C. Berkeley Division of Biostatistics Working Paper Series.

van der Laan, M. J. and Luedtke, A. R. (2014). Targeted learning of an optimal dynamic treatment, and statistical inference for its mean outcome. U.C. Berkeley Division of Biostatistics Working Paper Series.

van der Laan, M. J. and Luedtke, A. R. (2015). Targeted learning of the mean outcome under an optimal dynamic treatment rule. *Journal of Causal Inference*, 3(1):61–95.

van der Laan, M. J., Polley, E. C., and Hubbard, A. E. (2007). Super learner. *Statistical Applications in Genetics and Molecular Biology*, 6(1).

van der Laan, M. J. and Robins, J. M. (2003). *Unified Methods for Censored Longitudinal Data and Causality*. Springer Series in Statistics. Springer Publishing.

van der Laan, M. J. and Rose, S. (2011). *Targeted Learning: Causal Inference for Observational and Experimental Data*. Springer Series in Statistics. Springer Publishing.

van der Laan, M. J. and Rose, S. (2018). *Targeted Learning in Data Science: Causal Inference for Complex Longitudinal Studies*. Springer Series in Statistics. Springer Publishing.

van der Laan, M. J. and Rubin, D. (2006). Targeted maximum likelihood learning. *The International Journal of Biostatistics*, 2(1).

van der Wielen, N., Channon, A. A., and Falkingham, J. (2018). Does insurance enrolment increase healthcare utilisation among rural-dwelling older adults? Evidence from the national health insurance scheme in Ghana. *BMJ Global Health*, 3(1):e000590.

VanderWeele, T. J., Luedtke, A. R., van der Laan, M. J., and Kessler, R. C. (2019). Selecting optimal subgroups for treatment using many covariates. *Epidemiology*, 30(3):334–341.

Varian, H. R. (2014). Big data: New tricks for econometrics. *Journal of Economic Perspectives*, 28(2):3–28.

Vidyattama, Y., Miranti, R., and Resosudarmo, B. P. (2014). The role of health insurance membership in health service utilisation in Indonesia. *Bulletin of Indonesian Economic Studies*, 50(3):393–413.

Vilcu, I., Probst, L., Dorjsuren, B., and Mathauer, I. (2016). Subsidized health insurance coverage of people in the informal sector and vulnerable population groups: trends in institutional design in Asia. *International Journal for Equity in Health*, 15(1):165.

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

Wager, S., Hastie, T., and Efron, B. (2014). Confidence intervals for random forests: the jackknife and the infinitesimal jackknife. *Journal of Machine Learning Research*, 15(1):1625–1651.

Wagstaff, A. (2008). Measuring financial protection in health. Technical Report 4554, World Bank.

Wagstaff, A. (2010). Estimating health insurance impacts under unobserved heterogeneity: the case of Vietnam's health care fund for the poor. *Health Economics*, 19(2):189–208.

Wagstaff, A., Flores, G., Hsu, J., Smitz, M.-F., Chepynoga, K., Buisman, L. R., van Wilgenburg, K., and Eozenou, P. (2018). Progress on catastrophic health spending in 133 countries: a retrospective observational study. *The Lancet Global Health*, 6(2):e169–e179.

Wagstaff, A. and Lindelow, M. (2008). Can insurance increase financial risk? The curious case of health insurance in China. *Journal of Health Economics*, 27(4):990–1005.

Wagstaff, A., O'Donnell, O., van Doorslaer, E., and Lindelow, M. (2007). *Analyzing Health Equity Using Household Survey Data: A Guide to Techniques and their Implementation*. World Bank Publications.

Wagstaff, A. and van Doorslaer, E. (2003). Catastrophe and impoverishment in paying for health care: with applications to Vietnam 1993-1998. *Health Economics*, 12(11):921–934.

Wang, M. and Barbu, A. (2019). Are screening methods useful in feature selection? An empirical study. *PloS one*, 14(9):e0220842.

Wang, R., Lagakos, S. W., Ware, J. H., Hunter, D. J., and Drazen, J. M. (2007). Statistics in medicine–reporting of subgroup analyses in clinical trials. *New England Journal of Medicine*, 357(21):2189–2194.

Westreich, D., Lessler, J., and Funk, M. J. (2010). Propensity score estimation: neural networks, support vector machines, decision trees (CART), and meta-classifiers as alternatives to logistic regression. *Journal of Clinical Epidemiology*, 63(8):826–833.

Wiseman, V., Thabrany, H., Asante, A., Haemmerli, M., Kosen, S., Gilson, L., Mills, A., Hayen, A., Tangcharoensathien, V., and Patcharanarumol, W. (2018). An evaluation of health systems equity in Indonesia: study protocol. *International Journal for Equity in Health*, 17(1):138.

Wooldridge, J. M. (2010). *Econometric Analysis of Cross Section and Panel Data*. MIT Press.

World Bank (2020). Indonesia public expenditure review 2020: Spending for better results. Technical report, World Bank.

World Health Organization (2010). Health systems financing: The path to universal coverage. Technical report, World Health Organization.

Xu, K., Evans, D. B., Kawabata, K., Zeramdini, R., Klavus, J., and Murray, C. J. L. (2003). Household catastrophic health expenditure: a multicountry analysis. *The Lancet*, 362(9378):111–117.

Xu, Y., Greene, T. H., Bress, A. P., Sauer, B. C., Bellows, B. K., Zhang, Y., Weintraub, W. S., Moran, A. E., and Shen, J. (2020). Estimating the optimal individualized treatment rule from a cost-effectiveness perspective. *Biometrics*, 78(1):337–351.

Yadlowsky, S., Fleming, S., Shah, N., Brunskill, E., and Wager, S. (2021). Evaluating treatment prioritization rules via rank-weighted average treatment effects. `https://arxiv.org/abs/2111.07966`. Accessed: 2023-2-5.

Zhao, Y., Zeng, D., Rush, A. J., and Kosorok, M. R. (2012). Estimating individualized treatment rules using outcome weighted learning. *Journal of the American Statistical Association*, 107(449):1106–1118.

Zhou, Z., Athey, S., and Wager, S. (2022). Offline Multi-Action policy learning: generalization and optimization. *Operations Research*, 71(1):148–183.

Zhou, Z.-H. (2012). *Ensemble methods: foundations and algorithms*. Chapman and Hall/CRC.

Zimmert, M. and Lechner, M. (2019). Nonparametric estimation of causal heterogeneity under high-dimensional confounding. `https://arxiv.org/abs/1908.08779`. Accessed: 2023-2-5.

Zou, H. (2006). The adaptive lasso and its oracle properties. *Journal of the American Statistical Association*, 101(476):1418–1429.