DEPARTMENT OF INTERNATIONAL AND EUROPEAN ECONOMIC STUDIES

ATHENS UNIVERSITY OF ECONOMICS AND BUSINESS

Reverse matching for ex-ante policy evaluation

George Planiteros

Working Paper Series

22-06

January 2022

Reverse matching for ex-ante policy evaluation

George Planiteros

January 28, 2022

Abstract

The paper attacks the central policy evaluation question of forecasting the impact of interventions never previously experienced. It introduces treatment effects approach into a cognitive domain not currently spanned by its methodological arsenal. Existing causal effects bounding analysis is adjusted to the ex-ante program evaluation setting. A Monte Carlo experiment is conducted to test how severe the estimates of the proposed approach deviate from the "real" causal effect in the presence of selection and unobserved heterogeneity. The simulation shows that the approach is valid regarding the formulation of the counterfactual states given previous knowledge of the program rules and a sufficiently informative treatment probability. It also demonstrates that the width of the bounds are resilient to several deviations from the conditional independence assumption.

Keywords: Policy evaluation, forecasting, treatment effects, hypothetical treatment group, bounding and sensitivity analysis.

1 Introduction

The objective of the paper is to nonparametrically recover ex-ante causal effects of interest. It contributes to the econometric program evaluation literature in the following ways: introducing the treatment effects methodological platform into a problem that currently lies beyond its spectrum, that is, forecasting the effects of a policy never previously implemented; making explicit the economic rationale behind the analysis, which is generally absent in the treatment effects approach; adjusting treatment effects bounding analysis to ex-ante policy evaluation settings and modeling uncertainty with respect to both treatment assignment and outcome realization; testing empirically the ability of the approach to address the evaluation problem, or else to define treatment and control groups; and conducting sensitivity analysis to test the resilience of the proposed bounds to different specifications of the model's uncertainty factors.

Policy evaluation literature is mainly concentrated on estimating causal effects of programs already being implemented. Concisely, the problem under consideration is that of evaluating the effect of the exposure of a set of units to a program or policy on some outcome of interest. Reviews of the program evaluation literature can be found in Angrist and Krueger (1999), Wooldridge (2002), Imbens (2004), Angrist and Pischke (2009), Imbens and Wooldridge (2009), and more recently in Athey and Imbens (2017) and Abadie and Cattaneo (2018). A thorough analysis of the theoretical background is provided in Heckman and Vytlacil (1999), Heckman et al. (1999), Heckman and Vytlacil (2007a,b) and Abbring and Heckman (2007).

Ex-ante policy evaluation is part of the more general problem of studying the effects of policy changes prior to their implementation. In the early discrete choice literature, the problem was about predicting the demand for a new good, prior to its being introduced in the market (for one of the earliest applications, see McFadden (1977)). A key question in the literature is how can the performance, or else the predictive ability of the models be validated. One of the most common but not always feasible paths is to compare models' forecasts of treatment effects to those obtained from randomized experiments. An early expression of this approach is the study of Moffitt (1979). Other applications can be found in Lumsdaine et al. (1992), Lise et al. (2004), Todd and Wolpin (2006), Attanasio et al. (2012) and more recently in Gechter et al. (2018).

This study is triggered by a key finding emphasized in early papers by Marschak (1953) and Hurwicz (1966) and in the more recent work of Heckman (2000, 2001), Ichimura and Taber (2002) and Carneiro et al. (2002); that is, estimating the effect of a new policy does not necessarily require specifying the complete structure of the model governing decisions.

The proposed approach can be used for ex-ante policy evaluation in observational settings. Two methodological tools are employed: microsimulation, to estimate the effects of the intervention on the distribution of the policy variables, namely the variables affected by the policy under examination, and matching, to process the simulated data and address the evaluation problem; that is, identify treatment and control groups in the pre-treatment period.

Matching methods (Rubin, 1973; Rosenbaum and Rubin, 1983) are developed on the basis of hypothetical control groups, allowing for causal effects inference when their outcomes are compared to those of the actual treatment groups. This rationale is hereby inversed with the introduction of hypothetical treatment groups. This is what justifies the provocative title of the essay.

Given the extended degrees of uncertainty embedded in the ex-ante nature of the study, point estimation is highly improbable to produce accuracy. Existing treatment effects' bounding analysis is adjusted to take into account uncertainty on both treatment assignment and outcome realization.

A Monte Carlo experiment is provided to explore the properties of the proposed approach, to assess, computationally wise, its correctness, and to test if it provides an estimation that complies with the "real" treatment effect, given the data generating process. Put differently, the aim of the experiment is to test empirically the ability of the approach to address the evaluation problem. Moreover, the resilience of the proposed bounds is tested through calibration of the model's uncertainty factors.

What is crucial for the feasibility of the analysis that follows is its ability to identify variation in the policy variables. The proposed method is applicable to interventions the impact of which can be "translated" in monetary terms. Policies or programs deprived from quantitative characteristics cannot produce identifiable variation in the policy variables. The proposed approach is not applicable to this type of cases.

Before proceeding, it should be noted that the terms policy, program and intervention are used interchangeably throughout the paper.

2 Set-up

The analysis follows a nonparametric representation of the econometric program evaluation set-up in order to reduce model dependence, avoid functional form specification and impose weaker identification restrictions.

2.1 The model

The treatment effects approach infers causality ex-post, by comparing the observed differences in outcomes, Y, given variation in treatment status, T. The collection of potential responses is $\mathcal{Y} = \{Y(t,q) : t \in \mathcal{T},\ q \in \mathcal{Q}\}$, where $Y(t,q)$ denotes the observed outcome if treatment t were assigned to agent q.

Following Heckman (2001), Heckman and Vytlacil (2007b), and Todd and Wolpin (2008), we show that causal effect parameters can also be recovered ex-ante, for interventions never previously implemented. Given the absence of any actually imposed treatment, it is the variation in the policy variables that allows inference on the expected treatment effects.

In contrast to the treatment effects' methodological approach, it is economic theory that dictates the analytical steps in what follows. Treatment assignment is considered to be an agent's choice and the analysis is unfolded in a principal - agent setting. In our case, the principal is the authority (state, local government etc.) that introduces a new policy, $p \in \mathcal{P}$, never previously implemented, or amends an existing one, and the agents are the units (individual, family, household) being affected by the policy. The principal "offers" a policy (tax, benefit, subsidy), that induces agents' participation decisions by changing their incentives.

The mechanism that affects agents' incentives can be written as

$$\kappa : \mathcal{Q} \to \mathcal{I}$$

where $\kappa \in K$ defines the rule that maps agents into incentive constraints, $\mathcal{I}$, and $\mathcal{I}$ includes parameters of tax and benefit schedules.

Given the change in incentives induced by the policy, the agents decide upon their participation (or even their participation level). The mechanism that defines participation, or else the treatment assignment mechanism can be described by the rule

$$\sigma : \mathcal{Q} \times \mathcal{K} \times \mathcal{I} \to \mathcal{T}$$

where $\sigma \in \Sigma$ maps the agent $q \in Q$ who faces the incentive constraint $i \in \mathcal{I}$ assigned by mechanism $\kappa \in \mathcal{K}$ into a treatment $t \in T$.

The decision rule for program participation is characterized by a binary choice model analyzed by Matzkin (2008), where the unobservables enter the utility function in non-additive ways (for identification of similar models see Matzkin (1992, 1993, 1994) and Lewbel (2000))

$$T = \begin{cases} t_{p'}, & \text{if } h_{p'}(S, Z_{p'}, v) > h_p(S, Z_p, v) \text{ for all } p' \neq p \\ t_p, & \text{otherwise} \end{cases} \tag{1}$$

where h is a nonparametric utility function of the two alternatives, the existing policy regime, p, and the new intervention, p'. Agents choose to participate when their utility increases, given potential eligibility rules, so that $T = t_{p'}$, otherwise the current policy schedule applies, in which case $T = t_p$. Policies are characterized by a random vector Z, which includes the policy variables that affect agents' choice decision. S represents a vector of agents' policy invariant observable characteristics, and v, an unobservable random vector. Note that S can include all the observed elements of the outcome equation as well as others unique to the choice equation.

The outcome equations are described by (for identification see Matzkin (2003); Blundell and Powell (2003); Imbens and Newey (2009))

$$Y_{t_{p'}} = \mu_{p'}(X, Z_{p'}, \varepsilon) \tag{2}$$

$$Y_{t_p} = \mu_p(X, Z_p, \varepsilon) \tag{3}$$

where Y_{t_p} is the outcome of the agent under the no-treatment status and $Y_{t_{p'}}$ the outcome if the treatment is received (the treatment in this case is the policy under analysis). The outcome is represented by μ, a non-linear, non-separable function of policy invariant observables, X, policy variables, Z, and unobservables ε. Note that ε is a random variable and both the distributions of ε and the function μ are unknown.

The following restrictions are imposed on the choice and outcome equations

Assumption 1. The nonparametric function h belongs to a set of functions that are continuous in $(S, Z_{p'}, v)$ and strictly increasing in v. Likewise, μ belongs to a set of functions that are continuous in $(X, Z_{p'}, \varepsilon)$ and strictly increasing in ε.

Assumption 2. The cumulative distribution function of the policy variable, $F_{Z_{p'}|S,X}$, is absolutely continuous with respect to a Lebesgue measure with a nondegenerate conditional density $f(z_{p'}|s, x)$.

Assumption 3. The support of the conditional distribution of the policy variable given the observables, i.e. $\text{Supp}(F_{Z_{p'}|S,X})$ exists, is bounded and finite.

Assumption 4. $S_p = S$ and $X_p = X$, $\forall\, p$, i.e. S and X are external variables determined outside the model and invariant to counterfactual manipulations of T.

Assumption 5. The cumulative distribution functions of the unobservables given the observables and the policy variable, i.e.$F_{v|S,Z_{p'}}$ and $F_{\varepsilon|X,Z_{p'}}$ are both strictly increasing.

Assumption 6. The unobservables are distributed independently from the policy variables conditional on the policy invariant variables, that is, $(\varepsilon, v) \perp\!\!\!\perp Z_{p'}|S, X$.

Assumption 7. The propensity score is defined as $E(T|S, Z) = P_t(Z)$. Moreover, $Pr(T = t_{p'}|S, Z) = Pr(\{v|h_{p'}(S, Z_{p'}, v) > h_p(S, Z_p, v)\})$, for all $p' \neq p$

Assumption 8. The probability of outcome realization conditional on the treatment status, the observables and the policy variable is $Pr(Y_{t_{p'}}|X, Z_{p'}, T = t_{p'}) = Pr(Y_{t_{p'}}|P_{t_{p'}}(Z), T = t_{p'})$.

Assumptions [1] - [4] are critical for identifying the functions h and μ. Matzkin (1994, 2003) establishes that in nonadditive models two different continuous functions should be restricted to belong to a set of functions so as not to be observationally equivalent. That is, it should be guaranteed that their corresponding inverse functions are not strictly increasing transformations of each other [1]. In addition, it is implied that there is a neighborhood with positive probability, within which each function can attain different values from the other [2] and [3]. Finally, the observables in both functions are not affected by the policy, i.e. they are policy invariant [4]. Assumptions [5] and [6] guarantee that the cumulative distributions of the unobservables, $F_{v|\cdot}$ and $F_{\varepsilon|\cdot}$, are identified whenever h and μ are identified. They also guarantee that, controlling for heterogeneity and selection, variation in h and μ is translated into variation in the values of $T|S, Z_{p'}$ and $Y|X, Z_{p'}$, respectively. Assumption [7] implies that the conditional choice probability depends only on the differences between the utilities of the alternatives p and p' and it is employed to encounter the evaluation problem. Finally, assumption [8] ensures that $Z|X$ enters the model through the propensity score $P_{t_{p'}}(z)$.

Next, the analytical steps for the generation of the counterfactual states and the estimation of the expected causal effects are determined.

2.2 The method

The analysis is unfolded in three stages. The first stage considers simulating the introduction of the new intervention and how the latter transforms the distribution of the variables that affect agents' choices. We assume that the new policy, p', expressed in terms of the existing policy variables, Z_p, can be defined for some agents $q_{j'} \in \mathcal{Q}$, notationally reduced to j' from now on, as

$$Z_{p',j'} = M_{p'}(Z_{p,j'})$$

where $M_{p'}$ is assumed to be a known deterministic transformation of Z_p, given that the rules of the new policy belong to the information set of the analyst and that Z_p contains the same list of variables as $Z_{p'}$.

The above along with the assumptions imposed on the choice equation imply that

$$\text{if } Z_{p',j'} \geq Z_{p,j'} \Rightarrow h_{p'}(S, Z_{p',j'}, v) \geq h_p(S, Z_{p,j'}, v) \text{ for all } p' \neq p, \text{ then } T = t_{p'}$$

which means that variation in the policy variables determines treatment status. Further elaborated, the new policy induces changes in the agents' incentives through its anticipated impact on the policy variables, which increases agents' expected utility and affects their participation decision.

The second stage includes constructing counterfactuals to deal with the evaluation problem. Let $Z_{p,j}$ denote the policy variables for the agents $q_j \in \mathcal{Q}$ (for all $j \neq j'$) who remain unaffected by the treatment and assume common support conditions hold, so that

Assumption 9. $\text{Supp}(Z_{p,j}|S) \cap \text{Supp}(Z_{p',j'}|S) = z \neq \emptyset$

To construct counterfactuals the following rather strong assumption is invoked

Assumption 10. $P_{t_{p'}}(z) = Pr(T = t_{p'}|Z_{p',j'} = z, S)$

$$= Pr(T = t_p|Z_{p,j} = z, S) = P_{t_p}(z)$$

Assumptions [9] and [10] ensure that over the common support of the policy variables, Z, the treatment probability of the expected-to-be-treated, $P_{t_{p'}}(z)$, is equal to that of those agents that are not eligible to participate but having the same values of Z after the policy is being implemented. That is, the known conditional probability of treatment is the same with the estimated one for the non-eligible, for the same values of the policy variables.

Given the above, potential outcomes are being transformed accordingly

Assumption 11. $\underbrace{(Y_{t_{p'},j'}|X, Z_{p',j'} = z, \varepsilon)}_{\text{unobserved}} = (Y_{t_{p'},j'}|X, M_{p'}(Z_{p,j'}) = z, \varepsilon)$

$$= (Y_{t_p,j}|X, Z_{p,j} = z, \varepsilon) = \underbrace{(Y_{t_p,j}|X, P_{t_p}(z), \varepsilon)}_{\text{observed}}$$

Hence, the after-policy response is assumed to be equal to the historical outcome at the after-policy level.

Assumptions [9] - [11] together with the identification assumptions [1] - [8] allow the outcome equations to be written as

$$Y_{t_{p'}} = \mu_{p'}(X, Z_{p'}, \varepsilon) = Y_{t_p,j}|X, P_{t_p}(z), \varepsilon$$

$$Y_{t_p} = \mu_p(X, Z_p, \varepsilon) = Y_{t_p,j'}|X, P_{t_{p'}}(z), \varepsilon$$

The transformation implies that the potential outcome for the hypothetical treatment group is assumed to be equal to that of the expected non-to-be-treated agents, $Y_{t_p,j}$, whereas the potential outcome for the control group is assumed to be equal to that of the expected participants, $Y_{t_p,j'}$, given the conditional choice probability at the after-policy level.

This takes us to the third stage which deals with the estimation of the treatment effects of interest. For example, the Average Treatment Effect (ATE) can be written as

$$\begin{aligned}\text{ATE}(t_{p'}, q|X, Z, \varepsilon) &= E(Y_{t_p,j}|X, Z, \varepsilon) - E(Y_{t_p,j'}|X, Z, \varepsilon) \\ &= E(Y_{t_p,j}|X, Z_{p,j}, \varepsilon) - E(Y_{t_p,j'}|X, Z_{p',j'}, \varepsilon) \\ &= E(Y_{t_p,j}|X, P_{t_p}(z), \varepsilon) - E(Y_{t_p,j'}|X, P_{t_{p'}}(z), \varepsilon)\end{aligned}$$

Table 1 summarizes the three stages of the analysis.

2.3 The bounds

In the ex-ante policy evaluation setting neither treatment assignment nor outcome realization are actually observed. Given the above and taking as a starting point the general bounds of Manski (1989), Heckman and Vytlacil (1999) and Heckman and Vytlacil (2001), bounding analysis is adjusted to settings where treatment effects can

Table 1: The three-stage analysis for ex-ante policy evaluation

Stages	Assumptions
Simulation	$Z_{p',j'} = M_{p'}(Z_{p,j'})$
Counterfactuals	(i) $\text{Supp}(Z_{p,j}\|S) \cap \text{Supp}(Z_{p',j'}\|S) = z \neq \emptyset$;
	(ii) $P_{t_{p'}}(z) = P_{t_p}(z)$; (iii) $(Y_{t_{p'},j'}\|X, Z_{p',j'} = z, \varepsilon) = (Y_{t_p,j}\|X, P_{t_p}(z), \varepsilon)$
Estimation	$\text{ATE}(t_{p'}, q\|X, Z, \varepsilon) = E(Y_{t_p,j}\|X, Z_{p,j}, \varepsilon) - E(Y_{t_p,j'}\|X, Z_{p',j'}, \varepsilon)$

Note: Ex-ante policy evaluation can be analyzed in three stages: (i) Simulation of the policy variables' distribution; (ii) Construction of the counterfactuals, given that common support conditions hold, (iii) Estimation of the treatment effect of interest.

be recovered ex-ante; the bounds are defined by taking into account both sources of uncertainty.

The more general case where both treatment and outcome variables are continuous is examined. The results can be easily transformed into the binary case. A detailed proof of the analysis is provided in Appendix A. The resulting bounds on $E(Y_{t_{p'},q} - Y_{t_p,q}|X, Z_{p'}, \varepsilon)$ are $B^L \leq E(Y_{t_{p'},q} - Y_{t_p,q}|\cdot) \leq B^U$ where

$$B^L = p_t^{sup} p_y^{sup}[E(Y_{t_{p'}}|\cdot)] + (1 - p_t^{sup} p_y^{sup}) y^l - (1 - p_t^{inf} p_y^{inf})[E(Y_{t_p}|\cdot)] - p_t^{inf} p_y^{inf} y^u,$$

$$B^U = p_t^{sup} p_y^{sup}[E(Y_{t_{p'}}|\cdot)] + (1 - p_t^{sup} p_y^{sup}) y^u - (1 - p_t^{inf} p_y^{inf})[E(Y_{t_p}|\cdot)] - p_t^{inf} p_y^{inf} y^l$$

The width of the bounds is

$$W = B^U - B^L = \Big((1 - p_t^{sup} p_y^{sup}) + (p_t^{inf} p_y^{inf})\Big)\Big(y^u - y^l\Big)$$

and does not necessarily include zero.

The width of the bounds is determined to a great extent by the magnitude of p_t and p_y. The former denotes the value at which the participation probability, $P_t(Z) = Pr(T = t_{p'}|S, Z_{p'})$, is evaluated, whereas the latter is the value of the outcome realization probability, $P_y(Z) = Pr(Y = Y_{t_{p'}}|X, Z_{p'}, T = t_{p'})$. [1] Intuitively, the case where $p_t = 1$ is that of full compliance with the treatment assignment mechanism; all eligible agents

[1]The outcome realization probability can be interpreted as the prognostic score of outcome. For a thoroghough analysis see Hansen (2008).

choose to participate in the program. Given the above limiting case, when in addition $p_y = 1$, i.e. the probability that the realized outcome in the after policy period is identical to the ex-ante estimated one equals to 1, which implies that the unobserved characteristics between the two groups are alike, the above bounds are reduced to the point estimation of ATE. The width is also determined by the deviation between y^u and y^l, the greatest and least element of the outcome variable, respectively.

The added value of the proposed bounding analysis is that it decomposes the stochastic elements of the model, allowing for different weights' assignment to different types of decisions. Furthermore, it can be used to examine the sensitivity of the estimation to different values of the unobserved factors, accounting for ranging magnitude and opposite direction effects.

3 The experiment

Monte Carlo simulation is used to assess the validity of the approach in addressing the evaluation problem and its resilience to different deviations from the conditional independence assumption. The experimental setting mimics the introduction of a cash transfer conditional on an income threshold and the number of children in the family, and examines its effect on some behavioral response of interest. For simplicity, we disregard any potential policies' cross effects.

In the context of the simulation exercise it is not feasible to assume a nonparametric outcome representation. Assuming that the common support condition holds, this does not affect the essence of the approach.

3.1 Data-generating process and simulated cases

The data-generating process (DGP) is described by the following set of equations

$$Z_{p'} = M_{p'}(Z_p) = Z_p + 3K \text{ if } T = t_{p'},\ Z_p \text{ otherwise} \tag{4}$$

$$T = t_{p'} \equiv h_{p'}(S, Z_{p'}, v) = \mathbf{1}[S(X, K \geq 1), Z_p \leq 6, v] \tag{5}$$

$$T = t_p \equiv h_p(S, Z_{p'}, v) = \mathbf{1}[S(X, K), Z_{p'} | T = t_{p'}, v] \tag{6}$$

$$Y_{t_{p'}} = 0.3 + 0.4Z_{p'} + 0.4X + e_1 \tag{7}$$

$$Y_{t_p} = 0.1 + 0.2Z_{p'} + 0.3X + e_0 \tag{8}$$

Equation 4 represents the deterministic transformation of the policy variable given program rules: participants increase their income, Z_p, by 3 units times the number of kids, K, in the family. Equations 5 and 6 represent the main idea behind the proposed approach, or else how the evaluation problem is addressed. Equation 5 denotes participation for those agents that according to the program rules are eligible and expected to participate in the program; that is, families with children and income below a specified threshold ($Z_p \leq 6$). Equation 6 represents the hypothetical treatment group. The latter consists of families whose income is at the same level with that of the first group after the simulation of the policy intervention. The outcome equations, 7 and 8 are chosen arbitrarily and constitute slightly modified versions of those used in Cerulli (2014). They are characterized by additive separability in variables and errors.

Population values of X generated as independent draws from a chi-squared distribution as $X \sim \chi^2(3)+2$. Population values of K generated as independent draws from a Poisson distribution as $K \sim$ Pois(2). Population values of the policy variable generated as independent draws from a left truncated at zero normal distrobution as $Z_p \sim$ trN(6, 4). Finally, the error terms in the choice and outcome equations are distributed as $v \sim$ Unif[0, 1], $e_0 \sim$ N(0, 3), and $e_1 \sim$ N(0, 6.5).

The DGP is simulated 1.000 times and a sample size of 5.000 units is used. Each simulation provides an estimation of the ATE and the Average Treatment Effect on the Treated (ATET).

The severity and type of potential bias is investigated through the simulation of six different cases. Case 1a assumes full compliance with the treatment assignment mechanism and unobservable homogeneity, such that $Y \perp\!\!\!\perp T | X, Z_{p'}, e$ and $\rho_{v,e_0} = 0$, $\rho_{v,e_1} = 0$, $\rho_{e_0,e_1} = 0$. Case 1b is same as 1a but $K \sim$ Pois(1) and is exposed to illustrate the relationship between the strength of the instrument and the value of the propensity score and consequently between the strength of the instrument and the width of the bounds. While preserving the assumption of unobservable homogeneity, in case 2a the assumption of full treatment compliance is dropped, so that $T = t_p \equiv \mathbf{1}[S(X, K), Z^*_{p'}|T = t_{p'}, v]$, where $Z^*_{p'} = (i * v) * Z_{p'}$, $i = 2$, $\rho_{v,e_0} = 0.5$, $\rho_{v,e_1} = 0.3$, $\rho_{e_0,e_1} = 0$, and $v \sim$ Unif[], $e \sim$ N(). "Endogeneity" invades the model through an unobservable term v that affects the policy variable and is correlated with the outcome equations' error terms.[2] Case 2b is same as 2a but for trivariate normality of the error terms, so that $v \sim$ N(0,1) and $e \sim$ N(). In case 3a, heterogeneity also enters the model through an unobservable term u that affects the "endogenous" policy variable solely through the outcome equation, so that $Z^{**}_{p'} = Z^*_{p'} + i * u = (i * v) * Z_{p'} + i * u$, where $i = 1$, $u \sim \chi^2(3)$, $e^* = e + u$, and correlation between the unobservables is the same as in 2a and 2b. Finally, case 3b is same as 3a but $i = 2$. The latter case increases the magnitude of unobservable heterogeneity that is present in the model and consequently the potential for bias

3.2 Results

Tables 2 and 3 in Appendix B provide detailed results on the simulated ATE and ATET respectively, as well as on the mean standard errors (columns 2 and 3). They also demonstrate the proportional bias of the point estimates for cases 2 and 3 when compared to the baseline (column 4). Finally, they report the constructed upper and lower bounds for cases 1a and 1b (columns 5 and 6). The reported bounds are constructed by estimating the propensity score, $E(T|S, Z) = P_t(Z)$, and evaluating p^{sup}_t and p^{inf}_t according to the treatment status, at the maximum value of the treatment probability for the expected-to-be-treated and at the minimum for the non-eligible, respectively;

[2]Given the ex-ante nature of the experiment we cannot actually detect or correct for the presence of treatment endogeneity. Yet, its consequences on the composition of the hypothetical treatment and control groups approximate the distortions produced by Type I - Type II errors in statistical hypotheis testing. The distribution of participants is composed not only by eligible participants but also by units that do not participate in the program while eligible (Type I error - false positives) and by units that participate in the program while not eligible (Type II error - false negatives). Different specifications of the error terms' distribution and the degree of their correlation changes the composition of participants in comparison to the baseline case and allows for testing the resilience of the approach.

that is, $p_t^{sup} = max\{Pr(T|T = t_{p'})\}$ and $p_t^{inf} = min\{Pr(T|T = t_p)\}$. At this point, it is also assumed that $p_y = 1$, i.e. that the after policy response is equal to the historical outcome at the after policy level.

When calculated by hand and evaluated at the means of the variables included in the model, the ATE is 2.50. The simulated ATE in the baseline case, 1a, is quite close, at 2.59. The ATET in the baseline case is 2.48. In 1b, K preserves the type of the distribution but not its shape. K's probability mass function is concentrated around unity, increasing the range of the estimated participation probability and narrowing the width of the bounds in comparison to 1a. This comes to verify the anticipated inverse relationship between the range of values of the treatment probability and the width of the bounds. When treatment "endogeneity" is introduced in the model, the constructed bounds for both ATE and ATET in cases 2a and 2b are proven resilient and informative The existence of unobservable heterogeneity, along with treatment "endogeneity", in cases 3a and 3b, increases further the bias of the estimates and questions the resilience of the bounds. It is worth noting that the standard error means in all cases are very close to the standard deviations of the estimators. This means that, in the context of the simulation exercise, the asymptotic distribution of the ATE and ATET approximates their finite-sample distribution well.

Next, sensitivity analysis is used to examine the extent to which uncertainty affects the model. Different values are assigned to the bearers of uncertainty, p_t and p_y, and their impact on the width of the bounds is assessed. In contrast to the previous analysis and the way the bounds were constructed, uncertainty is introduced in the outcome realization as well. Put differently, p_y captures various specifications of the probability the realized outcome in the after policy period to be identical to the ex-ante estimated one. Similarly, the different values assigned on the propensity score represent corresponding deviations from the baseline treatment probability estimate. The outcome of the analysis is presented in detail in Appendix C. Tables 4 and 6 provide the results of the sensitivity analysis for the ATE and ATET lower bound whereas tables 5 and 7 those for the upper bound respectively. We restrict analysis within a specific range of values, that is, the product $(p_t^{sup} p_y^{sup})$ takes values between 1 and 0.9 whereas the product $(p_t^{inf} p_y^{inf})$ between 0 and 0.1.

4 Discussion - Concluding remarks

The paper complements the discussion on ex-ante recovering causal effects of a policy intervention under a nonparametric technique, namely without relying on functional form or distributional assumptions. This does not mean that the proposed approach proceeds without any restrictions; on the contrary, some of the latter are rather strong. The decision on the method to be followed lies on the plausibility assessment of the assumptions invoked. The underlying perception is that prior knowledge of functional forms or distributions is rare and the consequences of parametric misspecification are serious and difficult to mitigate.

Reverse matching provides a natural way to conduct ex-ante policy evaluation. To overcome the impossibility to observe the composition of the actual treatment group in the pre-treatment period, a hypothetical treatment group needs to be constructed. A simulation experiment explores how severe the model estimates deviate from the "real" causal effect under different variations of selection and unobserved heterogeneity. The experiment shows that: (i) The approach can be valid regarding the formulation of the treatment and control groups given previous knowledge of the program rules and a sufficiently informative propensity score; (ii) Not surprisingly, the overall resilience of the bounds depends on the magnitude of the correlation between the unobservables in the choice and in the outcome equation. When both selection and unobservable heterogeneity are present, the estimated treatment effects reach the extreme values of the bounds; (iii) Finally, sensitivity analysis based on assigning different values on the uncertainty parameters of the choice and outcome equations can provide a useful insight on the scale of deviation from the point estimation and the corresponding change in the width of the bounds.

It is worth mentioning that the results of the experiment are not indicative of the predictive ability of the model. The latter depends on the degree of coincidence of the actual participants' distribution with that predicted by the chosen selection mechanism and on the realization of outcomes in the after-policy period. The predictive ability can be empirically tested by comparing the models' estimates with the ex-post evaluated results of the policy under analysis; this validity test constitutes breeding grounds for future analysis.

References

A. Abadie and M. D. Cattaneo. Econometric Methods for Program Evaluation Alberto. *Annual Review of Economics*, 10:465–503, 2018. doi: 10.1146/annurev-economics. URL `https://doi.org/10.1146/annurev-economics-`.

J. H. Abbring and J. J. Heckman. Econometric Evaluation of Social Programs, Part III:. In J. J. Heckman and E. E. Leamer, editors, *Handbook of Econometrics, Volume 6B*, chapter 72, pages 5148–5294. North-Holland, 2007.

J. D. Angrist and A. B. Krueger. Emprirical strategies in Labor Economics. In O. Ashenfelter and D. Card, editors, *Handbook of Labor Economics*, chapter 23, pages 1278–1365. Elsevier Science B.V., third edition, 1999.

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

S. Athey and G. W. Imbens. The state of applied econometrics: Causality and policy evaluation. *Journal of Economic Perspectives*, 31(2):3–32, mar 2017. ISSN 08953309. doi: 10.1257/jep.31.2.3.

O. P. Attanasio, C. Meghir, and A. Santiago. Education choices in Mexico: Using a structural model and a randomized experiment to evaluate PROGRESA. *Review of Economic Studies*, 79(1):37–66, jan 2012. ISSN 1467937X. doi: 10.1093/restud/rdr015.

R. Blundell and J. L. Powell. Endogeneity in nonparametric and semiparametric regression models. *Econometric Society Monographs*, 36:312–357, 2003. doi: 10.1920/wp.cem.2001.0901. URL `http://dx.doi.org/10.1920/wp.cem.2001.0901www.econstor.eu`.

P. Carneiro, K. T. Hansen, and J. J. Heckman. Removing the veil of ignorance in assessing the distributional impacts of social policies. 2002. URL `http://www.nber.org/papers/w8840`.

G. Cerulli. ivtreatreg: A command for fitting binary treatment models with heterogeneous response to treatment and unobservable selection. *The Stata Journal*, 14(3):453–480, 2014.

M. Gechter, C. Samii, R. Dehejia, and C. Pop-Eleches. Evaluating Ex Ante Counterfactual Predictions Using Ex Post Causal Inference. *arXiv preprint arXiv:1806.07016*, 2018.

B. B. Hansen. The Prognostic Analogue of the Propensity Score. *Source: Biometrika*, 95(2):481–488, 2008. doi: 10.1093/biomet/asnOO4.

J. J. Heckman. Causal parameters and policy analysis in economics: A twentieth century retrospective. *Quarterly Journal of Economics*, 115(1):45–97, 2000. ISSN 00335533. doi: 10.1162/003355300554674.

J. J. Heckman. Micro data, heterogeneity, and the evaluation of public policy: Nobel lecture. *Journal of Political Economy*, 109(4):673–748, 2001. ISSN 00223808. doi: 10.1086/322086.

J. J. Heckman and E. J. Vytlacil. Local instrumental variables and latent variable models for identifying and bounding treatment effects. In *Proceedings of the National Academy of Sciences*, volume 96, pages 4730–4734, 1999. URL `www.pnas.org.`

J. J. Heckman and E. J. Vytlacil. Instrumental variables, selection models, and tight bounds on the average treatment effect. In M. Lechner and F. Pfeiffer, editors, *Econometric Evaluation of Labour Market Policies*, pages 1–15. Center for European Economic Research, New York, 2001.

J. J. Heckman and E. J. Vytlacil. Econometric Evaluation of Social Programs, Part I: Causal Models, Structural Models and Econometric Policy Evaluation. In J. J. Heckman and E. E. Leamer, editors, *Handbook of Econometrics, Volume 6B*, chapter 70, pages 4779–4874. North-Holland, 2007a.

J. J. Heckman and E. J. Vytlacil. Econometric Evaluation of Social Programs, Part II: Using the Marginal Treatment Effect to Organize Alternative Econometric Estimators to Evaluate Social Programs, and to Forecast their Effects in New Environments. In J. J. Heckman and E. E. Leamer, editors, *Handbook of Econometrics, Volume 6B*, chapter 71, pages 4875–5143. North-Holland, 2007b.

J. J. Heckman, R. J. LaLonde, and J. A. Smith. The economics and econometrics of active labor market programs. In O. Ashenfelter and D. Card, editors, *Handbook of Labor Economics*, pages 1865–2097. North-Holland, 1 edition, 1999.

L. Hurwicz. On the Structural Form of Interdependent Systems. *Studies in Logic and the Foundations of Mathematics*, 44(C):232–239, 1966. ISSN 0049237X. doi: 10.1016/S0049-237X(09)70590-7.

H. Ichimura and C. Taber. Semiparametric reduced-form estimation of tuition subsidies. *American Economic Review*, 92(2):286–292, 2002.

G. W. Imbens. Nonparametric Estimation of Average Treatment Effects under Exogeneity: A Review. *The Review of Economics and Statistics*, 86(1):4–29, 2004.

G. W. Imbens and W. K. Newey. Identification and Estimation of Triangular Simultaneous Equations Models Without Additivity. *Econometrica*, 77(5):1481–1512, 2009. URL `http://www.nber.org/papers/T0285`.

G. W. Imbens and J. M. Wooldridge. Recent developments in the econometrics of program evaluation. *Journal of Economic Literature*, 47(1):5–86, mar 2009. ISSN 00220515. doi: 10.1257/jel.47.1.5.

A. Lewbel. Semiparametric qualitative response model estimation with unknown heteroscedasticity or instrumental variables. *Journal of Econometrics*, 97(1):145–177, 2000. ISSN 03044076. doi: 10.1016/S0304-4076(00)00015-4.

J. Lise, S. Seitz, and J. Smith. Equilibrium Policy Experiments and the Evaluation of Social Programs. 2004. URL `http://www.nber.org/papers/w10283`.

R. L. Lumsdaine, J. H. Stock, and D. A. Wise. Three models of retirement: computational complexity versus predictive validity. In *Topics in the Economics of Aging*, pages 21–60. University of Chicago Press, 1992.

C. F. Manski. Anatomy of the Selection Problem. *The Journal of Human Resources*, 24 (3):343–360, 1989.

J. Marschak. Economic measurements for policy and prediction. In W. C. Hood and T. G. Koopmans, editors, *Studies in Econometric Method*, chapter 1, pages 1–26. Cowles Commission Research Staff Members, 1953. doi: 10.2307/3497754.

R. L. Matzkin. Nonparametric and distribution-free estimation of the binary threshold crossing and the binary choice models. *Econometrica*, 60(2):239–270, 1992. URL `http://about.jstor.org/terms`.

R. L. Matzkin. Nonparametric identification and estimation of polychotomous choice models. *Journal of Econometrics*, 58:137–168, 1993.

R. L. Matzkin. Restrictions of economic theory in nonparametric methods. In R. F. Engel and D. Mcfadden, editors, *Handbook of Econometrics, Volume 4*, chapter 42, pages 2524–2554. North-Holland, 1994.

R. L. Matzkin. Nonparametric estimation of nonadditive random functions. *Econometrica*, 71(5):1339–1375, 2003.

R. L. Matzkin. Identification in Nonparametric Simultaneous Equations Models. *Econometrica*, 76(5):945–978, 2008. ISSN 0012-9682. doi: 10.3982/ecta5940.

D. McFadden. Quantitative Methods for Analyzing Travel Behaviour of Individuals: Some Recent Developments. Technical report, Cowles Foundation for Research in Economics, Yale University, 1977.

R. A. Moffitt. The Labor Supply Response in the Gary Experiment. *The Journal of Human Resources*, 14(4):477–487, 1979.

P. R. Rosenbaum and D. B. Rubin. The central role of the propensity score in observational studies for causal effects. *Biometrika*, 70(1):41–55, 1983. URL `http://biomet.oxfordjournals.org/`.

D. B. Rubin. Matching to Remove Bias in Observational Studies. *Biometrics*, 29(1): 159–183, 1973.

P. E. Todd and K. I. Wolpin. Assessing the impact of a school subsidy program in Mexico: Using a social experiment to validate a dynamic behavioral model of child schooling and fertility. *American Economic Review*, 96(5):1384–1417, 2006.

P. E. Todd and K. I. Wolpin. Ex Ante Evaluation of Social Programs. *Annales d'Économie et de Statistique*, 91/92:263–291, 2008.

J. M. Wooldridge. *Econometric Analysis of Cross Section and Panel Data*. MIT Press, 2002.

A Bounds for ex-ante treatment effects

We can express the expected outcome after the policy intervention, $E(Y_{t_{p'}})$, in probabilistic terms as follows (for notational simplicity, we keep q implicit)

$$
\begin{aligned}
E\Big[TY|S=s,X=x,Z=z\Big] \\
&= E\Big[Y_{t_{p'}}|P_t(Z)=p_v, T=1, P_y(Z)=p_\varepsilon, Y=Y_{t_{p'}}\Big]p(v,\varepsilon) \\
&= E\Big[Y_{t_{p'}}|P_t(Z)=p_v, h_{p'}(s,z_{p'},v) > h_p(s,z_p,v), \\
&\qquad P_y(Z)=p_\varepsilon, Y_{t_{p'}}=\mu_{p'}(x,z_{p'},\varepsilon)\geq\varepsilon\Big]p_v p_{\varepsilon|v} \\
&= E\Big[Y_{t_{p'}}|P_t(Z)=p_v, P_t(Z)\geq v, \\
&\qquad P_y(Z)=p_\varepsilon, P_y(Z)\geq\varepsilon)\Big]p_v p_{\varepsilon|v} \\
&= E\Big[Y_{t_{p'}}|p_v\geq v, p_\varepsilon\geq\varepsilon\Big]p_v p_{\varepsilon|v} \\
&= \int_0^{p_{\varepsilon|v}}\int_0^{p_v} E\Big[Y_{t_{p'}}|V=v,\mathcal{E}=\varepsilon\Big]dF_{v|S=s,Z=z}(p_v)dF_{\varepsilon|T=1,X=x,Z=z}(p_{\varepsilon|v})
\end{aligned}
$$

where we use the fact that the cumulative distribution functions of the unobservables v and ε are identified when the unknown functions h and μ, respectively, are identified.

The double integral on $E(Y_{t_{p'}})$ provides no information on the distribution of (T,Y,S,X,Z) regarding the "counterfactual" probabilities

$$\int_0^{p_{\varepsilon|v}}\int_{p_v}^1 E(Y_{t_{p'}}|\cdot)dF_{v|S,Z}(\cdot)dF_{\varepsilon|T,X,Z}(\cdot)$$

$$\int_{p_{\varepsilon|v}}^1\int_0^{p_v} E(Y_{t_{p'}}|\cdot)dF_{v|S,Z}(\cdot)dF_{\varepsilon|T,X,Z}(\cdot)$$

$$\int_{p_{\varepsilon|v}}^1\int_{p_v}^1 E(Y_{t_{p'}}|\cdot)dF_{v|S,Z}(\cdot)dF_{\varepsilon|T,X,Z}(\cdot)$$

Assuming that the outcomes are bounded within a specific interval as $[B^L,B^U]=[y^l,y^u]$ and that we can evaluate the double integral on $E(Y_{t_{p'}})$ at $p_v = p_t^{sup(s,z)}$ and $p_{\varepsilon|v} =$

$p_y^{sup(x,z)}$ the above expressions can be bounded as

$$p_y^{sup}(1-p_t^{sup})y^l \leq \int_0^{p_y^{sup}} \int_{p_t^{sup}}^1 E(Y_{t_{p'}}|\cdot)dF_{v|S,Z}(\cdot)dF_{\varepsilon|X,Z}(\cdot) \leq p_y^{sup}(1-p_t^{sup})y^u$$

$$(1-p_y^{sup})p_t^{sup}y^l \leq \int_{p_y^{sup}}^1 \int_0^{p_t^{sup}} E(Y_{t_{p'}}|\cdot)dF_{v|S,Z}(\cdot)dF_{\varepsilon|X,Z}(\cdot) \leq (1-p_y^{sup})p_t^{sup}y^u$$

$$(1-p_y^{sup})(1-p_t^{sup})y^l \leq \int_{p_y^{sup}}^1 \int_{p_t^{sup}}^1 E(Y_{t_{p'}}|\cdot)dF_{v|S,Z}(\cdot)dF_{\varepsilon|X,Z}(\cdot) \leq (1-p_y^{sup})(1-p_t^{sup})y^u$$

Following the same pattern, the expected outcome for the control group, $E(Y_{t_p})$, can be expressed as

$$E[(1-T)Y|S=s,X=x,Z=z] = \int_{p_\varepsilon}^1 \int_{p_v}^1 E[Y_{t_p}|V=v,E=\varepsilon]dF_{v|S=s,Z=z}(p_v)dF_{\varepsilon|X=x,Z=z}(p_\varepsilon)$$

which can be bounded accordingly as

$$(1-p_y^{inf})p_t^{inf}y^l \leq \int_{p_y^{inf}}^1 \int_0^{p_t^{inf}} E(Y_{t_p}|\cdot)dF_{v|S}(\cdot)dF_{\varepsilon|X}(\cdot) \leq (1-p_y^{inf})p_t^{inf}y^u$$

$$p_y^{inf}(1-p_t^{inf})y^l \leq \int_0^{p_y^{inf}} \int_{p_t^{inf}}^1 E(Y_{t_p}|\cdot)dF_{v|S}(\cdot)dF_{\varepsilon|X}(\cdot) \leq p_y^{inf}(1-p_t^{inf})y^u$$

$$p_t^{inf}p_y^{inf}y^l \leq \int_0^{p_y^{inf}} \int_0^{p_t^{inf}} E(Y_{t_p}|\cdot)dF_{v|S}(\cdot)dF_{\varepsilon|X}(\cdot) \leq p_t^{inf}p_y^{inf}y^u$$

Thus $E(Y_{t_{p'}} - Y_{t_p}|S,X,Z)$ can be bounded as $B^L \leq E(Y_{t_{p'}} - Y_{t_p}|\cdot) \leq B^U$ where

$$B^L = p_t^{sup}p_y^{sup}[E(Y_{t_{p'}}|\cdot)] + (1-p_t^{sup}p_y^{sup})y^l - (1-p_t^{inf}p_y^{inf})[E(Y_{t_p}|\cdot)] - p_t^{inf}p_y^{inf}y^u,$$

$$B^U = p_t^{sup}p_y^{sup}[E(Y_{t_{p'}}|\cdot)] + (1-p_t^{sup}p_y^{sup})y^u - (1-p_t^{inf}p_y^{inf})[E(Y_{t_p}|\cdot)] - p_t^{inf}p_y^{inf}y^l$$

and the width of the bounds is

$$\Big((1 - p_t^{sup} p_y^{sup}) + (p_t^{inf} p_y^{inf})\Big)\Big(y^u - y^l\Big)$$

and does not necessarily include zero. Heckman and Vytlacil (2001) show that the bounds on the ATE given a nonparametric selection model are tight.

B Monte Carlo simulation output

Table 2: Monte Carlo simulation output: ATE

	Mean	Mean SE	Bias	LB	UB
1a	2.59	0.23	-	1.12	3.39
	(0.24)	(0.06)		(0.31)	(0.29)
1b	2.37	0.35	-	1.69	2.92
	(0.41)	(0.20)		(0.43)	(0.44)
2a	2.68	0.25	0.03	-	-
	(0.25)	(0.07)	(0.10)		
2b	2.22	0.30	−0.14	-	-
	(0.33)	(0.12)	(0.13)		
3a	2.95	0.19	0.14	-	-
	(0.17)	(0.02)	(0.06)		
3b	4.25	0.17	0.64	-	-
	(0.15)	(0.01)	(0.06)		

Note: The table provides the results on the simulated ATE and on the mean standard errors (columns 2 and 3). It also demonstrates the proportional bias of the point estimates of cases 2 and 3 when compared to 1a, which serves as the baseline (column 4). Finally, it reports the constructed lower and upper bounds for cases 1a and 1b (columns 5 and 6). The ATE when calculated by hand and evaluated at the means of the variables, given the DGP, is 2.50. The reported bounds are constructed by assuming $p_t = Pr(T = t_{p'}|\cdot)$, i.e. uncertainty in the choice equation equals the treatment probability and $p_y = 1$, i.e. the after policy response is equal to the historical outcome at the after policy level.

Table 3: Monte Carlo simulation output: ATET

	Mean	Mean SE	Bias	LB	UB
1a	2.48	0.16	-	1.01	3.27
	(0.16)	(0.01)		(0.26)	(0.23)
1b	2.39	0.16	-	1.70	2.93
	(0.16)	(0.01)		(0.23)	(0.26)
2a	2.63	0.17	0.06	-	-
	(0.13)	(0.01)	(0.05)		
2b	2.12	0.17	−0.14	-	-
	(0.15)	(0.01)	(0.06)		
3a	2.91	0.18	0.18	-	-
	(0.15)	(0.01)	(0.06)		
3b	4.07	0.17	0.64	-	-
	(0.14)	(0.01)	(0.06)		

Note: The table provides the results on the simulated ATET and on the mean standard errors (columns 2 and 3). It also demonstrates the proportional bias of the point estimates of cases 2 and 3 when compared to 1a, which serves as the baseline (column 4). Finally, it reports the constructed lower and upper bounds for cases 1a and 1b (columns 5 and 6) and their bias, i.e. how much do they proportionally deviate from the respective point estimate (columns 7 and 8). The reported bounds are constructed by assuming $p_t = Pr(T = t_{p'}|\cdot)$, i.e. uncertainty in the choice equation equals the treatment probability and $p_y = 1$, i.e. the after policy response is equal to the historical outcome at the after policy level.

C Bounds for the simulated treatment effects

Table 4: **Sensitivity Analysis: ATE Lower Bound (Case 1a)**

$(p_t^{sup}p_y^{sup})$ \ $(p_t^{inf}p_y^{inf})$	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09	0.10
1.00	2.59	2.50	2.41	2.32	2.23	2.14	2.06	1.97	1.88	1.79	1.70
	(0.00)	(−0.39)	(−0.77)	(−1.16)	(−1.54)	(−1.93)	(−2.31)	(−2.70)	(−3.09)	(−3.47)	(−3.86)
0.99	2.53	2.44	2.35	2.26	2.17	2.08	1.99	1.90	1.81	1.73	1.64
	(−0.27)	(−0.65)	(−1.04)	(−1.42)	(−1.81)	(−2.20)	(−2.58)	(−2.97)	(−3.35)	(−3.74)	(−4.12)
0.98	2.47	2.38	2.29	2.20	2.11	2.02	1.93	1.84	1.75	1.66	1.57
	(−0.54)	(−0.92)	(−1.31)	(−1.69)	(−2.08)	(−2.46)	(−2.85)	(−3.24)	(−3.62)	(−4.01)	(−4.39)
0.97	2.41	2.32	2.23	2.14	2.05	1.96	1.87	1.78	1.69	1.60	1.51
	(−0.80)	(−1.19)	(−1.57)	(−1.96)	(−2.35)	(−2.73)	(−3.12)	(−3.50)	(−3.89)	(−4.27)	(−4.66)
0.96	2.34	2.25	2.16	2.08	1.99	1.90	1.81	1.72	1.63	1.54	1.45
	(−1.07)	(−1.46)	(−1.84)	(−2.23)	(−2.61)	(−3.00)	(−3.39)	(−3.77)	(−4.16)	(−4.54)	(−4.93)
0.95	2.28	2.19	2.10	2.01	1.92	1.83	1.75	1.66	1.57	1.48	1.39
	(−1.34)	(−1.72)	(−2.11)	(−2.50)	(−2.88)	(−3.27)	(−3.65)	(−4.04)	(−4.42)	(−4.81)	(−5.20)
0.94	2.22	2.13	2.04	1.95	1.86	1.77	1.68	1.59	1.50	1.42	1.33
	(−1.61)	(−1.99)	(−2.38)	(−2.76)	(−3.15)	(−3.53)	(−3.92)	(−4.31)	(−4.69)	(−5.08)	(−5.46)
0.93	2.16	2.07	1.98	1.89	1.80	1.71	1.62	1.53	1.44	1.35	1.26
	(−1.87)	(−2.26)	(−2.65)	(−3.03)	(−3.42)	(−3.80)	(−4.19)	(−4.57)	(−4.96)	(−5.35)	(−5.73)
0.92	2.10	2.01	1.92	1.83	1.74	1.65	1.56	1.47	1.38	1.29	1.20
	(−2.14)	(−2.53)	(−2.91)	(−3.30)	(−3.68)	(−4.07)	(−4.46)	(−4.84)	(−5.23)	(−5.61)	(−6.00)
0.91	2.03	1.94	1.85	1.77	1.68	1.59	1.50	1.41	1.32	1.23	1.14
	(−2.41)	(−2.80)	(−3.18)	(−3.57)	(−3.95)	(−4.34)	(−4.72)	(−5.11)	(−5.50)	(−5.88)	(−6.27)
0.90	1.97	1.88	1.79	1.70	1.61	1.52	1.44	1.35	1.26	1.17	1.08
	(−2.68)	(−3.06)	(−3.45)	(−3.83)	(−4.22)	(−4.61)	(−4.99)	(−5.38)	(−5.76)	(−6.15)	(−6.53)

Note: The table provides the results of the sensitivity analysis for the ATE lower bound in the baseline case. Analysis is restricted within a specific range of values for the uncertainty parameters, p_t and p_y; that is, the product $(p_t^{sup}p_y^{sup})$ takes values between 1 and 0.9 whereas the product $(p_t^{inf}p_y^{inf})$ between 0 and 0.1. The upper-left case, where $(p_t^{sup}p_y^{sup})$ gets its largest value while $(p_t^{inf}p_y^{inf})$ its lowest, is that of no uncertainty. The bound is reduced to the baseline point estimate.

Table 5: **Sensitivity Analysis: ATE Upper Bound (Case 1a)**

	$(p_t^{inf}p_y^{inf})$										
$(p_t^{sup}p_y^{sup})$	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09	0.10
1.00	2.59	2.63	2.66	2.70	2.74	2.77	2.81	2.84	2.88	2.92	2.95
	(0.00)	(0.16)	(0.31)	(0.47)	(0.62)	(0.78)	(0.94)	(1.09)	(1.25)	(1.40)	(1.56)
0.99	2.66	2.69	2.73	2.76	2.80	2.84	2.87	2.91	2.94	2.98	3.02
	(0.27)	(0.43)	(0.59)	(0.74)	(0.90)	(1.05)	(1.21)	(1.37)	(1.52)	(1.68)	(1.83)
0.98	2.72	2.75	2.79	2.83	2.86	2.90	2.94	2.97	3.01	3.04	3.08
	(0.55)	(0.70)	(0.86)	(1.02)	(1.17)	(1.33)	(1.48)	(1.64)	(1.79)	(1.95)	(2.11)
0.97	2.78	2.82	2.85	2.89	2.93	2.96	3.00	3.03	3.07	3.11	3.14
	(0.82)	(0.98)	(1.13)	(1.29)	(1.45)	(1.60)	(1.76)	(1.91)	(2.07)	(2.22)	(2.38)
0.96	2.85	2.88	2.92	2.95	2.99	3.03	3.06	3.10	3.13	3.17	3.21
	(1.10)	(1.25)	(1.41)	(1.56)	(1.72)	(1.87)	(2.03)	(2.19)	(2.34)	(2.50)	(2.65)
0.95	2.91	2.94	2.98	3.02	3.05	3.09	3.13	3.16	3.20	3.23	3.27
	(1.37)	(1.52)	(1.68)	(1.84)	(1.99)	(2.15)	(2.30)	(2.46)	(2.62)	(2.77)	(2.93)
0.94	2.97	3.01	3.04	3.08	3.12	3.15	3.19	3.23	3.26	3.30	3.33
	(1.64)	(1.80)	(1.95)	(2.11)	(2.27)	(2.42)	(2.58)	(2.73)	(2.89)	(3.05)	(3.20)
0.93	3.04	3.07	3.11	3.14	3.18	3.22	3.25	3.29	3.32	3.36	3.40
	(1.92)	(2.07)	(2.23)	(2.38)	(2.54)	(2.70)	(2.85)	(3.01)	(3.16)	(3.32)	(3.48)
0.92	3.10	3.14	3.17	3.21	3.24	3.28	3.32	3.35	3.39	3.42	3.46
	(2.19)	(2.35)	(2.50)	(2.66)	(2.81)	(2.97)	(3.13)	(3.28)	(3.44)	(3.59)	(3.75)
0.91	3.16	3.20	3.23	3.27	3.31	3.34	3.38	3.42	3.45	3.49	3.52
	(2.46)	(2.62)	(2.78)	(2.93)	(3.09)	(3.24)	(3.40)	(3.56)	(3.71)	(3.87)	(4.02)
0.90	3.23	3.26	3.30	3.33	3.37	3.41	3.44	3.48	3.51	3.55	3.59
	(2.74)	(2.89)	(3.05)	(3.21)	(3.36)	(3.52)	(3.67)	(3.83)	(3.99)	(4.14)	(4.30)

Note: The table provides the results of the sensitivity analysis for the ATE upper bound in the baseline case. Analysis is restricted within a specific range of values for the uncertainty parameters, p_t and p_y; that is, the product $(p_t^{sup}p_y^{sup})$ takes values between 1 and 0.9 whereas the product $(p_t^{inf}p_y^{inf})$ between 0 and 0.1. The upper-left case, where $(p_t^{sup}p_y^{sup})$ gets its largest value while $(p_t^{inf}p_y^{inf})$ its lowest, is that of no uncertainty. The bound is reduced to the baseline point estimate.

Table 6: **Sensitivity Analysis: ATET Lower Bound (Case 1a)**

$(p_t^{sup}p_y^{sup})$	$(p_t^{inf}p_y^{inf})$ 0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09	0.10
0.10	2.48	2.39	2.30	2.21	2.12	2.03	1.93	1.84	1.75	1.66	1.57
	(0.00)	(−0.55)	(−1.10)	(−1.64)	(−2.19)	(−2.74)	(−3.29)	(−3.83)	(−4.38)	(−4.93)	(−5.48)
0.99	2.42	2.33	2.24	2.15	2.06	1.97	1.87	1.78	1.69	1.60	1.51
	(−0.36)	(−0.91)	(−1.46)	(−2.01)	(−2.55)	(−3.10)	(−3.65)	(−4.20)	(−4.74)	(−5.29)	(−5.84)
0.98	2.36	2.27	2.18	2.09	2.00	1.91	1.82	1.72	1.63	1.54	1.45
	(−0.73)	(−1.27)	(−1.82)	(−2.37)	(−2.92)	(−3.47)	(−4.01)	(−4.56)	(−5.11)	(−5.66)	(−6.20)
0.97	2.30	2.21	2.12	2.03	1.94	1.85	1.76	1.66	1.57	1.48	1.39
	(−1.09)	(−1.64)	(−2.19)	(−2.73)	(−3.28)	(−3.83)	(−4.38)	(−4.92)	(−5.47)	(−6.02)	(−6.57)
0.96	2.24	2.15	2.06	1.97	1.88	1.79	1.70	1.60	1.51	1.42	1.33
	(−1.45)	(−2.00)	(−2.55)	(−3.10)	(−3.64)	(−4.19)	(−4.74)	(−5.29)	(−5.84)	(−6.38)	(−6.93)
0.95	2.18	2.09	2.00	1.91	1.82	1.73	1.64	1.54	1.45	1.36	1.27
	(−1.82)	(−2.37)	(−2.91)	(−3.46)	(−4.01)	(−4.56)	(−5.10)	(−5.65)	(−6.20)	(−6.75)	(−7.29)
0.94	2.12	2.03	1.94	1.85	1.76	1.67	1.58	1.48	1.39	1.30	1.21
	(−2.18)	(−2.73)	(−3.28)	(−3.82)	(−4.37)	(−4.92)	(−5.47)	(−6.01)	(−6.56)	(−7.11)	(−7.66)
0.93	2.06	1.97	1.88	1.79	1.70	1.61	1.52	1.42	1.33	1.24	1.15
	(−2.54)	(−3.09)	(−3.64)	(−4.19)	(−4.74)	(−5.28)	(−5.83)	(−6.38)	(−6.93)	(−7.47)	(−8.02)
0.92	2.00	1.91	1.82	1.73	1.64	1.55	1.46	1.37	1.27	1.18	1.09
	(−2.91)	(−3.46)	(−4.00)	(−4.55)	(−5.10)	(−5.65)	(−6.19)	(−6.74)	(−7.29)	(−7.84)	(−8.38)
0.91	1.94	1.85	1.76	1.67	1.58	1.49	1.40	1.31	1.21	1.12	1.03
	(−3.27)	(−3.82)	(−4.37)	(−4.91)	(−5.46)	(−6.01)	(−6.56)	(−7.11)	(−7.65)	(−8.20)	(−8.75)
0.90	1.88	1.79	1.70	1.61	1.52	1.43	1.34	1.25	1.15	1.06	0.97
	(−3.64)	(−4.18)	(−4.73)	(−5.28)	(−5.83)	(−6.37)	(−6.92)	(−7.47)	(−8.02)	(−8.56)	(−9.11)

Note: The table provides the results of the sensitivity analysis for the ATET lower bound in the baseline case. Analysis is restricted within a specific range of values for the uncertainty parameters, p_t and p_y; that is, the product $(p_t^{sup}p_y^{sup})$ takes values between 1 and 0.9 whereas the product $(p_t^{inf}p_y^{inf})$ between 0 and 0.1. The upper-left case, where $(p_t^{sup}p_y^{sup})$ gets its largest value while $(p_t^{inf}p_y^{inf})$ its lowest, is that of no uncertainty. The bound is reduced to the baseline point estimate.

Table 7: **Sensitivity Analysis: ATET Upper Bound (Case 1a)**

$(p_t^{sup} p_y^{sup})$ \ $(p_t^{inf} p_y^{inf})$	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09	0.10
0.10	2.48	2.51	2.55	2.58	2.62	2.65	2.69	2.72	2.76	2.79	2.83
	(0.00)	(0.21)	(0.43)	(0.64)	(0.85)	(1.07)	(1.28)	(1.49)	(1.71)	(1.92)	(2.13)
0.99	2.54	2.58	2.61	2.65	2.68	2.72	2.75	2.79	2.82	2.86	2.89
	(0.40)	(0.61)	(0.82)	(1.04)	(1.25)	(1.46)	(1.68)	(1.89)	(2.10)	(2.32)	(2.53)
0.98	2.61	2.64	2.68	2.71	2.75	2.78	2.82	2.85	2.89	2.92	2.96
	(0.79)	(1.01)	(1.22)	(1.43)	(1.65)	(1.86)	(2.07)	(2.29)	(2.50)	(2.71)	(2.93)
0.97	2.67	2.71	2.74	2.78	2.81	2.85	2.88	2.92	2.95	2.99	3.03
	(1.19)	(1.41)	(1.62)	(1.83)	(2.05)	(2.26)	(2.47)	(2.69)	(2.90)	(3.11)	(3.33)
0.96	2.74	2.77	2.81	2.84	2.88	2.91	2.95	2.99	3.02	3.06	3.09
	(1.59)	(1.80)	(2.02)	(2.23)	(2.44)	(2.66)	(2.87)	(3.08)	(3.30)	(3.51)	(3.72)
0.95	2.80	2.84	2.87	2.91	2.95	2.98	3.02	3.05	3.09	3.12	3.16
	(1.99)	(2.20)	(2.41)	(2.63)	(2.84)	(3.05)	(3.27)	(3.48)	(3.69)	(3.91)	(4.12)
0.94	2.87	2.91	2.94	2.98	3.01	3.05	3.08	3.12	3.15	3.19	3.22
	(2.38)	(2.60)	(2.81)	(3.02)	(3.24)	(3.45)	(3.66)	(3.88)	(4.09)	(4.30)	(4.52)
0.93	2.94	2.97	3.01	3.04	3.08	3.11	3.15	3.18	3.22	3.25	3.29
	(2.78)	(3.00)	(3.21)	(3.42)	(3.64)	(3.85)	(4.06)	(4.28)	(4.49)	(4.70)	(4.92)
0.92	3.00	3.04	3.07	3.11	3.14	3.18	3.21	3.25	3.28	3.32	3.35
	(3.18)	(3.39)	(3.61)	(3.82)	(4.03)	(4.25)	(4.46)	(4.67)	(4.89)	(5.10)	(5.31)
0.91	3.07	3.10	3.14	3.17	3.21	3.24	3.28	3.31	3.35	3.38	3.42
	(3.58)	(3.79)	(4.00)	(4.22)	(4.43)	(4.64)	(4.86)	(5.07)	(5.28)	(5.50)	(5.71)
0.90	3.13	3.17	3.20	3.24	3.27	3.31	3.34	3.38	3.41	3.45	3.48
	(3.97)	(4.19)	(4.40)	(4.61)	(4.83)	(5.04)	(5.25)	(5.47)	(5.68)	(5.89)	(6.11)

Note: The table provides the results of the sensitivity analysis for the ATET upper bound in the baseline case. Analysis is restricted within a specific range of values for the uncertainty parameters, p_t and p_y; that is, the product $(p_t^{sup} p_y^{sup})$ takes values between 1 and 0.9 whereas the product $(p_t^{inf} p_y^{inf})$ between 0 and 0.1. The upper-left case, where $(p_t^{sup} p_y^{sup})$ gets its largest value while $(p_t^{inf} p_y^{inf})$ its lowest, is that of no uncertainty. The bound is reduced to the baseline point estimate.