Policy Learning under Endogeneity Using Instrumental Variables*

Yan Liu†

December 18, 2022

Abstract

This paper studies the statistical decision problem of learning an individualized intervention policy when data are obtained from observational studies or randomized experiments with imperfect compliance. Leveraging a continuous instrumental variable, we provide a social welfare criterion that accounts for endogenous treatment selection. To this end, we incorporate the marginal treatment effects (MTE) when identifying treatment effect parameters and consider encouragement rules that affect social welfare through treatment take-up when designing policies. We apply the representation of the social welfare criterion of encouragement rules via the MTE to the Empirical Welfare Maximization (EWM) method and derive convergence rates of the worst-case regret (welfare loss). Using data from the Indonesian Family Life Survey, we apply the EWM encouragement rule to advise on how to encourage upper secondary schooling. Our framework offers interpretability of the optimal policy by explaining why a certain subpopulation is targeted.

Keywords: Endogeneity, instrumental variables, encouragement rules, marginal treatment effects (MTE), empirical welfare maximization (EWM), statistical decision rules

*I am indebted to Hiroaki Kaido for his guidance and encouragement throughout this project. For helpful comments and discussions, I thank Shuowen Chen, Iván Fernández-Val, Bryan Graham, Eric Hardy, Sukjin Han, Toru Kitagawa, Ming Li, Ismael Mourifié, Xiaoxia Shi, and participants at the BU Econometric Seminar, NASMES 2022, IAAE 2022 Annual Conference, ESAM2022, SETA2022, YES 2022, and MEG 2022. The financial support of the IAAE Student Travel Grant is gratefully acknowledged. All errors are my own.

†Department of Economics, Boston University. Email: yanliu@bu.edu.
1 Introduction

An important goal of program evaluation is to inform policy on the assignment of individuals to treatment based on observable characteristics in the presence of treatment effect heterogeneity. This paper considers the statistical decision problem of learning an individualized intervention policy that maximizes social welfare. By leveraging a continuous instrumental variable (IV), we address two issues with policy learning using data from observational studies or randomized experiments with noncompliance, which can be viewed as two sides of the same coin. First, when identifying and estimating causal effects to form the social welfare criterion, we allow for endogeneity (i.e., unobserved confoundedness). Second, when designing and implementing policies, we allow for imperfect enforcement. The remedy is a broader definition of policy interventions that encompasses not only directly altering treatment status but also inducing flows into (and possibly out of) treatment by manipulating the instrument. We refer to the latter as the encouragement rule, a term borrowed from the epidemiology literature.

The first of the above issues concerns the identification of the social welfare criterion in order to evaluate a policy. For example, the policymaker may wish to select a binary treatment rule that assigns college enrollment to maximize average log wages. Each individual is assigned to a treatment based on observable characteristics that affect wages (e.g., parental education, rural/urban residence). The social welfare criterion, in this case, is the average outcome and can be written in terms of the conditional average treatment effect (CATE). The crucial assumption to point identify the CATE from the observed data is unconfoundedness, which is that the treatment is independent of potential outcomes conditional on observable characteristics. However, this assumption breaks down when selection is based on the unobserved components, e.g., ability or motivation, of heterogeneous responses to treatment. To deal with self-selection, we exploit a continuous instrumental variable informative of marginal treatment selection behavior. This enables us to incorporate the concept of marginal treatment effects (MTE) (Björklund and Moffitt, 1987) as a building block of the social welfare criterion. In this sense, we bridge the literatures on the MTE and on statistical decision rules of policy interventions.

The second issue concerns the definition of policies. We note a fundamental asymmetry in
the existing literature: the policymaker acknowledges endogenous treatment selection but designs treatment assignment rules as if they could be fully enforced. We address this asymmetry by expanding the universe of possible policies to include changes in the instrument that affect social welfare through the treatment, which we call encouragement rules. For instance, it is highly costly or impossible to force people to attend college (or not to). In terms of implementation, a more realistic scenario is to provide a scholarship or a tuition subsidy, whereas the tuition fee is commonly used as a continuous instrument for college attendance.

We apply the social welfare criterion of an encouragement rule, identified via the MTE function, to one popular class of statistical decision rules: Empirical Welfare Maximization (EWM) rules (see Hirano and Porter, 2020, Section 2.3). We analyze the properties of the EWM method in terms of regret (loss in expected welfare relative to an oracle optimal rule). To keep the analysis tractable, we restrict our attention to settings in which the policymaker makes binary decisions between two a priori chosen manipulation levels of the instrument and constrains the class of feasible allocations.\(^1\) Such settings are also of practical interest when the policymaker wants to avoid complicated rules or satisfy legal, ethical, or political considerations. We take advantage of the techniques developed by Kitagawa and Tetenov (2018b) to derive convergence rates of the worst-case regret across all unknown states of nature. In addition, we propose an alternative social welfare representation allowing for heterogeneity in treatment choice behavior when there are multiple instruments. As another practically relevant extension, we demonstrate how to incorporate budget constraints, which inherently involve an unknown cost because of imperfect take-up.

We apply the EWM encouragement rule to an empirical dataset from the third wave of the Indonesian Family Life Survey (IFLS) to provide advice on how upper secondary schooling can be encouraged to maximize average adult wages. We find that the optimal policy without budget constraints provides tuition subsidy eligibility to individuals who face relatively high tuition fees and live relatively close to the nearest secondary school. The MTE structure underlying the social welfare criterion allows us to understand how the optimal policy is driven by heterogeneity in treatment take-up and treatment effects.

\(^1\)In principle, it is possible to generalize the regret analysis to multi-action settings. This would necessitate different tools, such as different complexity measures of the policy class, and is beyond the scope of this paper.
Related Literature: This paper is related to four strands of literature, of which we provide a non-exhaustive overview below.

First, the research question is closely related to the literature on statistical treatment rules in econometrics, including Manski (2004), Hirano and Porter (2009), Bhattacharya and Dupas (2012), Kitagawa and Tetenov (2018b), Atty and Wager (2021), Mbakop and Tabord-Meehan (2021), Sun (2021). See also Hirano and Porter (2020) for a recent review. Despite the breadth of the literature, only a few econometric works look into policy choice with observational data when the unconfoundedness assumption does not hold. Kasy (2016) and Byambadalai (2022) focus on cases of partial identification and welfare ranking of policies rather than optimal policy choices. Atty and Wager (2021) assume homogeneous treatment effects so that the CATE on compliers can be extrapolated to those on the entire population. Sasaki and Ura (2020) identify the social welfare criterion via the MTE and demonstrate an application to the EWM framework. Nonetheless, these works implicitly assume the complete enforcement of treatment rules, whereas we consider more realistic policy tools. The only exception we are aware of is Chen and Xie (2022), who use the MTE framework to study the personalized subsidy rule. Their work is complementary to ours in emphasizing different aspects of policy learning. They focus on the oracle optimal policy and its welfare properties without restricting the policy class. In contrast, we restrict the complexity of the policy class and analyze the convergence rate of the regret bound of the estimated policy.

Second, in epidemiology and biostatistics, the problem of treatment choice is termed “individualized treatment rules” (ITR). Cui and Tchetgen (2021) and Qiu, Carone, Sadikova, Petukhova, Kessler, and Luedtke (2021) allow for treatment endogeneity. They achieve point identification by leveraging the “no common effect modifier” assumption outlined in Wang and Tchetgen Tchetgen (2018), which is that there is no unmeasured confounder that is simultaneously an effect modifier of the additive effect of treatment on the outcome and also of the additive effect of the instrument on the treatment. Such identification assumptions largely restrict the heterogeneity of compliance behavior. As an alternative criterion, Pu and Zhang (2021) introduce the notion of “IV-optimality” to estimate the optimal treatment regime based on partial identification of the CATE. Our approach crucially differs from these works on ITR under endogeneity in that we account for imperfect en-
forcement as a consequence of endogeneity. One exception is individualized encouragement rules that manipulate a binary instrument, considered in Qiu et al. (2021). Our framework nests theirs.

Third, if one takes an intention-to-treat perspective, then optimization of the manipulation of a continuous instrument can also be studied under a framework of policy learning with continuous treatments, assuming unconfoundedness. Kallus and Zhou (2018) generalize the inverse propensity weighting method to the continuous treatment setting and propose a kernelized off-policy evaluator. Bertsimas and McCord (2018) focus on regression-based methods and propose predicted cost minimization penalized by variance and bias. Athey and Wager (2021) study binary decisions on whether or not to offer each individual an infinitesimal nudge to the preexisting treatment level. In contrast, our approach treats the variable intervened upon as an instrument and imposes the exclusion restriction that the instrument only affects the outcome through a binary treatment. The MTE structure, justified under the exclusion restriction, allows the policymaker to understand why the optimal policy targets a certain subpopulation and to rigorously extrapolate beyond the observed support of the instrument.

Lastly, our representation of the social welfare criterion can be viewed as a variation of policy relevant treatment effects (PRTE), which adds to the class of policy parameters that can be written as weighted averages of the MTE. Hence, our research also complements the literature on identification, estimation, and inference of PRTE based on the MTE, including Carneiro, Heckman, and Vytlacil (2010, 2011), Mogstad, Santos, and Torgovitsky (2018), Mogstad and Torgovitsky (2018), and Sasaki and Ura (2021) among many others.

**Organization:** The rest of the paper is organized as follows. Section 2 sets up the model, introduces the encouragement rule with a continuous instrument, and derives a representation of the social welfare criterion via the MTE. Section 3 applies the social welfare representation to the EWM method and analyzes the associated theoretical properties. Section 4 discusses extensions incorporating multiple instruments and budget constraints. Section 5 presents an empirical application. Section 6 concludes and discusses some open questions. Mathematical proofs and additional discussions are collected in the Appendix.
2 Encouragement Rules with A Continuous Instrument

2.1 Setup

We consider the canonical program evaluation problem with a binary treatment $D \in \{0, 1\}$ and a scalar, real-valued outcome $Y \in \mathcal{Y} \subset \mathbb{R}$. Outcome production is modeled through the potential outcomes framework (Rubin, 1974):

$$Y = Y(1)D + Y(0)(1 - D),$$

where $(Y(0), Y(1))$ are the potential outcomes under no treatment and under treatment. Let $X \in \mathcal{X} \subset \mathbb{R}^{d_x}$ denote a vector of pretreatment covariates. For instance, in the analysis of returns to college, $D$ is an indicator for college enrollment, $Y$ is the log wage, and $X$ includes observable characteristics that affect wages (e.g., parental education, rural/urban residence). By adopting the potential outcomes model, we implicitly follow the conventional practice of imposing the Stable Unit Treatment Value Assumption (SUTVA), namely that there are no spillover or general equilibrium effects.

A \textit{treatment assignment rule} is defined as a mapping from $X$ to $\{0, 1\}$ and can be indexed by its decision set $G \subset \mathcal{X}$. The policymaker’s objective function is the utilitarian (additive) welfare criterion defined by the average outcome in the population:

$$W(G) = E[Y(1) \cdot 1\{X \in G\} + Y(0) \cdot 1\{X \notin G\}] = E[Y(0)] + E[1\{X \in G\} \cdot \underbrace{E[Y(1) - Y(0)|X]}_{\text{CATE}}].$$

Under the unconfoundedness assumption that $D$ is independent of $(Y(0), Y(1))$ conditional on $X$, the CATE is identified by $E[Y|D = 1, X = x] - E[Y|D = 0, X = x]$. However, the unconfoundedness assumption is violated if, for example, people self-select into college based on unmeasured benefits and costs driven by ability and motivation, both of which also affect wages. As a result, $E[Y(d)|X = x] \neq E[Y|D = d, X = x]$ for $d = 0, 1$ in general, and thus the social welfare criterion is not identified the moments of observables. In this case, it is helpful to assume that there exists an instrument
(i.e., an excluded variable) \( Z \in \mathbb{Z} \subset \mathbb{R} \) that affects the treatment but not the outcome, e.g., a tuition subsidy. We explicitly model the selection into treatment via an additively separable latent index model:

\[
D = 1\{\tilde{\nu}(X, Z) - \tilde{U} \geq 0\},
\]

where \( \tilde{\nu} \) is an unknown function, and \( \tilde{U} \) represents unobservable factors that affect treatment choice.

Let \( P \) denote the joint distribution of \((Y(0), Y(1), D, X, Z)\) and “\( \perp \)” denote (conditional) statistical independence. We adopt the following standard assumptions from the MTE literature (Heckman and Vytlacil, 1999, 2005; Mogstad et al., 2018):

**Assumption 1** (IV restrictions and continuous distribution)

1.1 \( \tilde{U} \perp Z | X \).

1.2 \( E_P[Y(d)|X, Z, \tilde{U}] = E_P[Y(d)|X, \tilde{U}] \) and \( E_P[Y(d)^2] < \infty \) for \( d \in \{0, 1\} \).

1.3 \( \tilde{U} \) is continuously distributed conditional on \( X \).

Assumptions 1.1 and 1.2 impose exogeneity and an exclusion restriction on \( Z \) but allow for arbitrary dependence between \((Y(0), Y(1))\) and \( \tilde{U} \), even conditional on \( X \). Vytlacil (2002) shows that, under Assumption 1(i), the existence of an additively separable selection equation as in (2) is equivalent to the monotonicity assumption used for the local average treatment effects (LATE) model of Imbens and Angrist (1994). The LATE monotonicity assumption restricts choice behavior in the sense that, conditional on \( X \), an exogenous shift in \( Z \) either weakly encourages or weakly discourages every individual to choose \( D = 1 \). Nonetheless, we maintain the selection equation (2) because it allows us to express the welfare contrast under counterfactual policies as a function of identifiable and interpretable objects. Assumption 1.3 allows us to reparameterize the model so that

\[
D = 1\{\nu(X, Z) - U \geq 0\} \quad \text{with} \quad U|X, Z \sim \text{Unif}[0, 1],
\]

7
where \( U \equiv F_{\tilde{V}|X}(\tilde{U}) \) and \( \nu(X, Z) \equiv F_{\tilde{V}|X}(\tilde{\nu}(X, Z)) \). As a consequence,

\[
\nu(x, z) = F_{U|X,Z}(\nu(x, z)) = \Pr_P(D = 1|X = x, Z = z) \equiv p(x, z),
\]

where \( p(x, z) \) is the propensity score.

Endogenous treatment selection also challenges the plausibility of fully mandating the treatment. We expand the universe of possible policies to include changes in the instrument that affect social welfare through the treatment. Let \( M_\alpha : Z \rightarrow \mathbb{R} \) be a user-chosen manipulation function. The manipulation function \( M_\alpha(\cdot) \) is indexed by a scalar \( \alpha \) representing the manipulation level and satisfies \( M_0(Z) = Z \). For example, \( M_\alpha(Z) \) describes a subsidy or discount in \( Z \) so that \( M_\alpha(Z) = Z + \alpha \).

An encourage rule is a mapping \( \alpha : X \times Z \rightarrow \mathbb{R} \) that determines the manipulation level of the instrument for an individual with observable characteristics \( X \) and instrument \( Z \). Our representation of the social welfare criterion in Section 2.2 covers the most general setting without restrictions on \( \alpha \). When we study the finite-sample regret bounds for statistical decision rules in Section 3, we take a stand on the complexity of feasible encourage rules. In particular, we focus on a case in which the policymaker makes a binary decision between two a priori chosen manipulation levels. This binary formulation of encourage rules nests treatment assignment rules (Kitagawa and Tetenov, 2018b; Sasaki and Ura, 2020) as a special case.

### 2.2 Representation of the Social Welfare Criterion via the MTE

Encouragement rule \( \alpha \) induces the treatment choice

\[
D^\alpha = 1\{p(X, M_{\alpha(X,Z)}(Z)) \geq U\},
\]

and the outcome that would be observed under encourage rule \( \alpha \) is

\[
Y^\alpha = Y(1) \cdot D^\alpha + Y(0) \cdot (1 - D^\alpha).
\]

---

\(^2\)This is one of the policy changes commonly considered by works on (marginal) PRTE.
We define the utilitarian (additive) welfare criterion as the average outcome in the population:

\[ W(\alpha) = E_P[Y^{\alpha}] \]

It turns out that \( W(\alpha) \) involves the average treatment effect conditional on selection unobservables \( U \) and observable characteristics \( X \), which is the definition of the MTE:

\[ MTE(u, x) \equiv E_P[Y(1) - Y(0)|U = u, X = x]. \]

The concept of MTE was introduced by Björklund and Moffitt (1987) and extended by Heckman and Vytlacil (1999, 2001, 2005). The MTE is identifiable under suitable assumptions on \( Z \) that do not contradict Assumption 1. Theorem 1 expresses \( W(\alpha) \) as a function of the MTE; a proof is provided in Appendix A.1. In light of Theorem 1, we can identify and estimate the social welfare criterion using existing methods to identify and estimate the MTE (e.g., Carneiro and Lee, 2009; Cornelissen, Dustmann, Raute, and Schönberg, 2016; Brinch, Mogstad, and Wiswall, 2017).

**Theorem 1.** Under Assumption 1, the social welfare criterion is given by

\[
W(\alpha) = E_P[Y] + E_P \left[ \int_0^1 MTE(u, X) \cdot \left( 1\{p(X, M_{\alpha(X,Z)}(Z)) \geq u\} - 1\{p(X, Z) \geq u\} \right) du \right].
\]

In contrast to the intention-to-treat approach, the representation in Theorem 1 unfolds the two forces driving the optimal policy based on the social welfare criterion: the change in treatment take-up when receiving the encouragement and the MTE among those induced to switch treatment status. Note that the welfare contrast \( W(G) - E_P[Y] = E_P[Y^{\alpha}] - E_P[Y] \) measures the total causal effect on \( Y \) when changing from the status quo to encouragement rule \( \alpha \) and can be viewed as an unnormalized version of PRTE (Heckman and Vytlacil, 2001; Carneiro et al., 2010). Corollary 1 gives an alternative representation of \( W(\alpha) \) based on this observation.

**Corollary 1.** Under Assumption 1, when \( E_P[p(X, M_{\alpha(X,Z)}(Z))] - E_P[p(X, Z)] \neq 0 \), the social welfare criterion is given by

\[
W(\alpha) = E_P[Y] + (E_P[p(X, M_{\alpha(X,Z)}(Z))] - E_P[p(X, Z)]) \cdot \text{PRTE}(\alpha),
\]
where

\[
\text{PRTE}(\alpha) = EP\left[\int_0^1 \text{MTE}(u, X) \cdot \omega(u, X, Z; \alpha) \, du\right]
\]

with the weight defined as

\[
\omega(u, x, z; \alpha) = \frac{1\{p(x, M_{\alpha(x,z)}(z)) \geq u\} - 1\{p(x, z) \geq u\}}{EP[p(X, M_{\alpha(X,Z)}(Z))] - EP[p(X, Z)]}.
\]

Corollary 1 clarifies the connection between the social welfare criterion and the PRTE. It decomposes the welfare contrast \( W(\alpha) - EP[Y] \) into the average change in treatment take-up and the PRTE when going from the status quo to encouragement rule \( \alpha \). The PRTE is a weighted average of MTE with weights determined by the change in treatment take-up for individuals with different values of \((U, X, Z)\).

### 2.3 Binary Encouragement Rules

When we study the finite-sample regret bounds for statistical decision rules in Section 3, we focus on the setting in which the policymaker a priori chooses two manipulation levels \((\alpha_0, \alpha_1)\) and makes a binary decision between manipulating the instrument by \(\alpha_0\) or \(\alpha_1\). A binary encouragement rule, indexed by a decision set \(G \subset X \times Z\), is

\[
\alpha^G(X, Z) = \alpha_1 \cdot 1\{(X, Z) \in G\} + \alpha_0 \cdot 1\{(X, Z) \notin G\}.
\]

When \(\alpha_0 = 0\), \(\alpha^G\) can be interpreted as an eligibility criterion for a welfare program, such that eligible individuals receive the same manipulation level of the instrument, and ineligible individuals are left in the status quo. A small number of possible manipulation levels may be preferred also because the policymaker wants to avoid overly unfair redistribution of the instrument, especially when the instrument is a monetary variable.

Corollary 2 specializes Theorem 1 to the binary setting by noting that the level of the instrument induced by \(\alpha^G\) is \(M_{\alpha^G(X,Z)}(Z) = M_{\alpha_1}(Z) \cdot 1\{(X, Z) \in G\} + M_{\alpha_0}(Z) \cdot 1\{(X, Z) \notin G\}\).
**Corollary 2.** Under Assumption 1, we have

\[ W(\alpha^G) = W_0 + E_P \left[ 1\{(X,Z) \in G\} \int_0^1 \text{MTE}(u,X) \cdot (1\{p(X,M_{\alpha_1}(Z)) \geq u\} - 1\{p(X,M_{\alpha_0}(Z)) \geq u\}) \, du \right], \]

where \( W_0 = E_P[Y(1) \cdot 1\{p(X,M_{\alpha_0}(Z)) \geq U\} + Y(0) \cdot 1\{p(X,M_{\alpha_0}(Z)) < U\}] \) does not depend on \( G \).

Finally, we demonstrate that our binary formulation of encouragement rules nests treatment assignment rules that directly assign individuals to a certain treatment status as a special case. Suppose that there exist \((\alpha_0,\alpha_1)\) satisfying

\[ p(X,M_{\alpha_1}(Z)) = 1 \text{ and } p(X,M_{\alpha_0}(Z)) = 0, \text{ for all } X \in X \]

so that \( M_{\alpha_1}(Z) \) and \( M_{\alpha_0}(Z) \) create perfectly strong incentives and disincentives to be in the treated state \((D = 1)\), respectively, across heterogeneous covariate values. In this case, encouragement rules are effectively treatment assignment rules: \( D^{\alpha G} = 1\{(X,Z) \in G\} \). Therefore, Corollary 2 provides a representation of the social welfare criterion for treatment assignment rules via the MTE function:

\[ W(\alpha^G) = E_P[Y(1) \cdot 1\{(X,Z) \in G\} + Y(0) \cdot 1\{(X,Z) \notin G\}] 
= E_P[Y(0)] + E_P \left[ 1\{(X,Z) \in G\} \int_0^1 \text{MTE}(u,X) \, du \right]. \]

This representation coincides with Theorem 1 of Sasaki and Ura (2020).\(^4\) However, such powerful manipulations are hard to justify in practice. For example, consider a selection of the form \( D = 1\{Z \geq \tilde{U}\} \) with \( \tilde{U} \) having full support on \( \mathbb{R} \) and a manipulation of the form \( M_\alpha(Z) = Z + \alpha \). Then, we need to set \( \alpha_1 = \infty \) and \( \alpha_0 = -\infty \) to induce full compliance.

\(^3\)Henceforth, with a slight abuse of notation, we write \( \int_{p(X,M_{\alpha_1}(Z))}^{p(X,M_{\alpha_0}(Z))} \text{MTE}(u,X) \, du = \int_0^1 \text{MTE}(u,X) \cdot (1\{p(X,M_{\alpha_1}(Z)) \geq u\} - 1\{p(X,M_{\alpha_0}(Z)) \geq u\}) \, du \) for brevity.

\(^4\)In this case, \( Z \) is redundant as a targeting variable.
3 Applications to EWM and Theoretical Properties

In this section, we focus on binary encouragement rules described in Section 2.3. For notational simplicity, we denote an encouragement rule and its social welfare by \( G \) and \( W(G) \) in place of \( \alpha^G \) and \( W(D^\alpha^G) \), respectively. Let \( \mathcal{G} \subset \{ G \subset X \times Z \} \) be a class of feasible encouragement rules. In view of Corollary 2, the optimal encouragement rule, if \( P \) is known, is

\[
G^*_P \in \arg \max_{G \in \mathcal{G}} E_P \left[ 1 \{ (X, Z) \in G \} \int_{p(X,M_{\alpha_1}(Z))}^{p(X,M_{\alpha_0}(Z))} \text{MTE}(u, X) \, du \right].
\]

We apply this representation to the EWM framework and investigate the theoretical properties of the resulting statistical decision rules. In Section 3.1, we consider the ideal case in which the mappings \( (u, x) \mapsto \text{MTE}(u, x) \) and \( (x, z) \mapsto p(x, z) \) are known. In Section 3.2, we consider a more realistic case in which the mappings \( (u, x) \mapsto \text{MTE}(u, x) \) and \( (x, z) \mapsto p(x, z) \) are unknown and thus need to be estimated. In Section 3.3, we present a concrete estimator for the MTE.

Some extra notations are needed to facilitate the discussion. Suppose the policymaker observes a random sample \( A_i = (Y_i, D_i, X_i, Z_i) \) of size \( n \). Let \( E_n \) denote the sample average operator, i.e., \( E_n f = \frac{1}{n} \sum_{i=1}^{n} f(A_i) \) for any measurable function \( f \). Let \( E_{P^n} \) denote the expectation taken over different realizations of the random sample. Let \( a \vee b = \max\{a, b\} \).

3.1 EWM with Known MTE and Propensity Score

Assume that the researcher knows the mappings \( (u, x) \mapsto \text{MTE}(u, x) \) and \( (x, z) \mapsto p(x, z) \). We define the EWM encouragement rule as a maximizer of the empirical welfare:

\[
\hat{G}^*_\text{EWM} \in \arg \max_{G \in \mathcal{G}} E_n \left[ 1 \{ (X, Z) \in G \} \int_{p(X,M_{\alpha_1}(Z))}^{p(X,M_{\alpha_0}(Z))} \text{MTE}(u, X) \, du \right].
\]

\(^5\)When the assumption of unconfoundedness holds and the randomization mechanism is known, known propensity score is a realistic scenario. However, in the presence of endogenous treatment selection, known propensity score is more like the oracle and thus of less practical interest.
In line with the literature on statistical treatment rules (Manski, 2004; Kitagawa and Tetenov, 2018b), we evaluate the performance of $\hat{G}_{EWM}$ by its regret (welfare loss):

$$W^*_G - E_{P^n}[W(\hat{G}_{EWM})] = E_{P^n}[W^*_G - W(\hat{G}_{EWM})],$$

where $W^*_G$ denotes the highest attainable welfare within class $G$:

$$W^*_G \equiv \sup_{G \in \mathcal{G}} W(G).$$

To analyze the theoretical properties of EWM encouragement rules, we impose the following regularity assumption, adapted from Kitagawa and Tetenov (2018b).

**Assumption 2 (Boundedness and VC-class)**

2.1 $|\int_0^1 \text{MTE}(u, X) \, du| \leq \bar{M} < \infty$ almost surely.

2.2 $\mathcal{G}$ has a finite VC-dimension $V < \infty$ and is countable.

Assumption 2 is adapted from Assumption 2.1 (BO)–(VC) of Kitagawa and Tetenov (2018b). Assumption 2.1 requires a bounded integral of the MTE over $[0, 1]$. A sufficient condition is that the outcome variable is bounded by some constant: $Y \in [y, \overline{y}]$ where $-\infty < y \leq \overline{y} < \infty$. We let $\mathcal{P}(\bar{M})$ denote the class of data generating processes (DGPs) that satisfy Assumptions 1 and 2.1.

Assumption 2.2 restricts the complexity of the class $\mathcal{G}$ of candidate encouragement rules in terms of VC-dimension. Interested readers can refer to van der Vaart and Wellner (1996) for the definition and textbook treatment of VC-dimension. We now give two examples of $\mathcal{G}$ that satisfy Assumption 2.2.

**Example 1. Linear Eligibility Score:** Assume that the eligibility score is linear in a subvector of $(x, z)$, $v \in \mathcal{V} \subset \mathbb{R}^{d_v}$ with $d_v \leq d_x + d_z$ and consider the class of decision sets defined as half-spaces generated by the eligibility score:

$$\mathcal{G}_{LES} = \{ G : G = \{(x, z) \in \mathcal{X} \times \mathcal{Z} : \lambda_0 + \lambda' v \geq 0 \}, (\lambda_0, \lambda') \in \mathbb{R}^{d_v+1} \}.$$
For example, individuals are assigned scholarship eligibility if a linear function of their GPA and parental income exceeds some threshold. \( G_{\text{LES}} \) has VC-dimension \( d_v + 1 \). The EWM method searches over all possible linear coefficients.

**Example 2. Threshold Allocations:** Consider the following class of decision sets:

\[
G_{\text{TA}} = \{ G : G = \{(x, z) \in \mathcal{X} \times \mathcal{Z} : s_k v_k \leq \bar{v}_k \text{ for } k \in \{1, \ldots, d_v\}, \bar{v} \in \mathbb{R}^{d_v}, s \in \{-1, 1\}^{d_v}\} \}
\]

For example, individuals are assigned scholarship eligibility if their GPA and parental income are respectively above and below some specific thresholds. \( G_{\text{TA}} \) has VC-dimension \( d_v \). The EWM method searches over all possible thresholds and directions.

Corollary 3 shows that the convergence rate of the worst-case regret for the EWM encouragement rule is no slower than \( n^{-1/2} \). It is a counterpart of Theorem 2.1 of Kitagawa and Tetenov (2018b); a proof is provided in Appendix A.2.

**Corollary 3.** Under Assumptions 1 and 2, we have

\[
\sup_{P \in \mathcal{P}(\bar{M})} E_{P^n}[W_{G}^* - W(\hat{G}_{\text{EWM}})] \leq 2C_1 \bar{M} \sqrt{\frac{V}{n}}
\]

where \( C_1 \) is a universal constant defined in Lemma A.4 of Kitagawa and Tetenov (2018a).

### 3.2 EWM with Unknown MTE and Propensity Score

We consider situations where the mappings \((u, x) \mapsto \text{MTE}(u, x)\) and \((x, z) \mapsto p(x, z)\) are unknown and need to be estimated from the empirical data. Given an estimator \( \hat{p}(x, z) \) for \( p(x, z) \) and an estimator \( \hat{\text{MTE}}(u, x) \) for \( \text{MTE}(u, x) \), we define the *feasible EWM encouragement rule* as a maximizer of empirical welfare obtained by plugging in these estimators:

\[
\hat{G}_{\text{FEWM}} \in \arg \max_{G \in \mathcal{G}} E_n \left[ 1\{ (X, Z) \in G \} \int_{\hat{p}(X, M_{\alpha_1}(Z))}^{\hat{p}(X, M_{\alpha_0}(Z))} \hat{\text{MTE}}(u, X) \, du \right]. \tag{4}
\]

To analyze the theoretical properties of feasible EWM encouragement rules, we impose the following assumptions.
**Assumption 3** (Propensity score estimation)

3.1 For each \((x, z) \in X \times Z\), \(\hat{p}(x, z) \in [0, 1]\).

3.2 For a class \(\mathcal{P}_p\) of DGPs, there exists a sequence \(\psi_n \to \infty\) such that

\[
\limsup_{n \to \infty} \sup_{P \in \mathcal{P}_p} \psi_n E_P \left[ \sup_{(x, z) \in X \times Z} |\hat{p}(x, z) - p(x, z)| \right] < \infty, \tag{5}
\]

where \(\bar{Z} \equiv Z \cup \left( \bigcup_{z \in Z} M_{\alpha_0}(z) \right) \cup \left( \bigcup_{z \in Z} M_{\alpha_1}(z) \right)\).

**Assumption 4** (MTE function and estimation) For a class \(\mathcal{P}_{\text{MTE}}\) of DGPs,

4.1 \(u \mapsto \text{MTE}(u, x)\) is continuous for each \(x \in X\);

4.2 \(E_P[\text{C}(X)^2] < \infty\), where \(\text{C}(x) \equiv \sup_{u \in [0, 1]} |\text{MTE}(u, x)|\);

4.3 there exists a sequence \(\phi_n \to \infty\) such that

\[
\limsup_{n \to \infty} \sup_{P \in \mathcal{P}_{\text{MTE}}} \phi_n E_P \left[ E_n \left[ \int_0^1 \left( \hat{\text{MTE}}(u, X) \, du - \text{MTE}(u, X) \right) \, du \right] \right] < \infty. \tag{6}
\]

Assumption 3.1 restricts the propensity score estimator to the unit interval. Assumption 3.2 concerns a sup-norm convergence rate \(\psi_n^{-1}\) in expectation for the propensity score estimator uniformly over \(\mathcal{P}_p\). Assumptions 4.1 and 4.2 are mild conditions on the MTE function. Assumption 4.3 concerns a convergence rate \(\phi_n^{-1}\) in expectation for the average absolute estimation error of \(\int_0^1 \hat{\text{MTE}}(u, X) \, du\) uniformly over \(\mathcal{P}_{\text{MTE}}\).

Theorem 2 derives the uniform convergence rate upper bound for the feasible EWM encouragement rule. It is a counterpart of Theorem 2.5 of Kitagawa and Tetenov (2018b); a proof is provided in Appendix A.3. Typically, \(\hat{\text{MTE}}(u, x)\) involves \(\hat{p}(x, z)\) as a generated regressor, rendering \(\phi_n^{-1}\) slower than \(\psi_n^{-1}\) in general. Hence, with a nonparametric propensity score estimator, \(\phi_n^{-1}\) will generally determine the convergence rate.

**Theorem 2.** Under Assumptions 1–4, we have

\[
\sup_{P \in \mathcal{P}_p \cap \mathcal{P}_{\text{MTE}} \cap \bar{P}(\bar{M})} E_P \left[ W^*_G - W(\hat{G}_{\text{FEWM}}) \right] \leq O(\phi_n^{-1} \lor \psi_n^{-1} \lor n^{-1/2}).
\]
3.3 Estimation of the MTE and the Propensity Score

To implement the feasible EWM encouragement rule, we need to specify estimators for the MTE function and the propensity score. The support of \( p(X, M_{\alpha_0}(Z)) \) and \( p(X, M_{\alpha_1}(Z)) \) may not be contained in the support of \( p(X, Z) \). Hence, we choose estimation methods that allow for extrapolation away from the variation in treatment induced by the available instrument.

For the propensity score estimator \( \hat{p}(X, Z) \), we consider global specifications such as logit and series regressions. In Appendix C, we derive a sup-norm convergence rate in expectation for series estimators using regression splines or polynomials. In order for a series estimator \( \hat{p}(X, Z) \) to satisfy Assumption 3.1, one can use a trimmed version of \( \hat{p}(X, Z) \) as in Carneiro and Lee (2009, Eq. (4.2)).

To estimate the MTE function, we follow Cornelissen et al. (2016) and Brinch et al. (2017) by considering a parametric regression function

\[
E_P[Y| p(X, Z) = u, X = x] = x' \beta_0 + x' (\beta_1 - \beta_0) u + \sum_{k=2}^{K} \gamma_k u^k. \tag{7}
\]

Define the regressor as

\[
\hat{W} \equiv ((1 - \hat{p}(X, Z)) X', \hat{p}(X, Z) X', \hat{p}(X, Z)^2, \ldots, \hat{p}(X, Z)^K)' .
\]

The OLS estimator for \( \theta = (\beta_0', \beta_1', \gamma_2, \ldots, \gamma_K)' \) is obtained by regressing \( Y \) on \( \hat{W} \):

\[
\hat{\theta} = (\hat{\beta}_0', \hat{\beta}_1', \hat{\gamma}_2, \ldots, \hat{\gamma}_K)' = E_n[\hat{W}\hat{W}']^{-1}E_n[\hat{W}Y].
\]

Then, the MTE can be estimated by

\[
\hat{\text{MTE}}(u, x) = x'(\hat{\beta}_1 - \hat{\beta}_0) + \sum_{k=2}^{K} k\hat{\gamma}_k u^{k-1}.
\]

As a result, the integral in the objective function in (4) is given by

\[
\int_{\hat{p}(X, M_{\alpha_0}(Z))}^{\hat{p}(X, M_{\alpha_1}(Z))} \hat{\text{MTE}}(u, X) \, du = X'(\hat{\beta}_1 - \hat{\beta}_0)(\hat{p}(X, M_{\alpha_1}(Z)) - \hat{p}(X, M_{\alpha_0}(Z))).
\]

16
\[ + \sum_{k=2}^{K} \hat{\gamma}_k (\hat{p}(X, M_{\alpha_1}(Z)) - \hat{p}(X, M_{\alpha_0}(Z)))\]

which is easy to compute. For this concrete estimator, Proposition 1 of Sasaki and Ura (2020) shows that, if \( \mathcal{P}_p \) is a class of DGPs such that (5) holds, then (6) holds with \( \phi_n = \psi_n \) and \( \mathcal{P}_{MTE} = \mathcal{P}_p \).

4 Extensions

We consider two empirically relevant extensions to the baseline setup in Section 2.1. In Section 4.1, we allow for the presence of other instruments in addition to the one that can be manipulated. In Section 4.2, we incorporate budget constraints. As a further extension, we consider encouragement rules with a binary instrument in Appendix B.

4.1 Multiple Instruments

In practice, the policymaker can observe multiple instruments, but suppose only one of them can be used as the tool for policy intervention. For example, tuition subsidies and proximity to a college are two prominent instruments for college enrollment, but only the former can serve as an encouragement. More generally, we allow \( Z \) to be \( L \)-dimensional. Let \( Z_1 \) be the instrument that can be intervened upon, \( Z_{-1} \equiv (Z_2, \ldots, Z_L)' \in \mathbb{R}^{L-1} \) be a vector of additional instruments, and \( Z = (Z_1, Z_{-1}')' \). We can write \( M_\alpha(Z) = (M_{\alpha,1}(Z_1), Z_{-1}')' \). For instance, \( M_\alpha(Z) = Z + \alpha e_1 \) and \( M_{\alpha,1}(Z_1) = Z_1 + \alpha \), where \( e_j \) is the \( j \)-th unit vector.

We propose the representation of the social welfare criterion

\[ W(\alpha) = E_P[Y] + E_P \left[ \int_0^{u_1} \text{MTE}_1(u_1, X, Z_{-1}) \cdot (1\{p(X, M_{\alpha}(X,Z))(Z) \geq u_1\} - 1\{p(X, Z) \geq u_1\}) \, du_1 \right] \tag{8} \]

where \( \text{MTE}_\ell \) is equivalent to the MTE function using the \( \ell \)-th instrument and conditioning on the other instruments as covariates. The analysis in Section 3 then applies.

---

\(^6\)To legitimize the parametric regression function, we implicitly impose a set of assumptions, including additive separability of outcomes in the observed and unobserved component, independence of the shape of the MTE curve across \( X \), and functional form assumptions. See Cornelissen et al. (2016) for detailed discussions of these assumptions.
To formally define the instrument-specific MTE in (8), we construct a selection equation for each marginal potential treatment defined as

\[ D_\ell(z_\ell) \equiv D(z_\ell, Z_{-\ell}) = 1\{p(X, z_\ell, Z_{-\ell}) \geq U_\ell\}, \]  

(9)

where \( U_\ell \) can be interpreted as a latent proneness to take the treatment, which is measured against the incentive (or disincentive) created by the \( \ell \)th instrument. We replace Assumption 1 with the following set of assumptions:

**Assumption 1’** (Instrument-specific IV restrictions and continuous distribution) For each \( \ell = 1, \ldots, L \),

1.1 \( U_\ell \perp Z_\ell|X, Z_{-\ell} \).

1.2 \( E_P[Y(d)|X, Z, U_\ell] = E_P[Y(d)|X, Z_{-\ell}, U_\ell] \) and \( E_P[Y(d)^2] < \infty \) for \( d \in \{0, 1\} \).

1.3 \( U_\ell|X, Z_{-\ell} \sim \text{Unif}[0, 1] \).

Equation (9) and Assumption 1’ only impose restrictions along each margin of selection and thus are agnostic about unobserved heterogeneity in the marginal rate of substitution across instruments.\(^7\) We can now define the instrument-specific MTE in (8) as

\[ \text{MTE}_\ell(u_\ell, x, z_{-\ell}) \equiv E_P[Y(1) - Y(0)|U_\ell = u_\ell, X = x, Z_{-\ell} = z_{-\ell}]. \]

### 4.2 Budget Constraints

Manipulating the instrument can be costly, especially when the instrument is a monetary variable such as price. In practice, the policymaker often faces budget constraints and wants to prioritize encouragement for the individuals who will benefit the most. Incorporating budget constraints is of particular interest when the treatment effect is intrinsically positive. For example, Dupas (2014) documents an experiment in Kenya that randomly assigned subsidized prices for a new

\(^7\)Mogstad, Torgovitsky, and Walters (2021) use a random utility model to demonstrate that in the presence of multiple instruments, (2) and Assumption 1 imply homogeneity in the marginal rate of substitution. In contrast, (9) and Assumption 1’ do not impose such implicit homogeneity.
health product. The treatment and outcome were indicators for the product’s purchase and usage, respectively. The product was not available outside the experiment, so the potential outcome if not treated is identically equal to zero. Hence, the first-best decision rule was to assign the treatment, or an encouragement that induced one-way flows into treatment, to everyone. However, to preserve financial resources, in this scenario, the policymaker may wish to exclude individuals who are not likely to increase product usage, for example, because of low disease risks in their neighborhood.

Let $C_{\alpha} : \mathcal{X} \times \mathcal{Z} \to \mathbb{R}^+$ be a user-chosen cost function indexed by $\alpha$. For example, $C_{\alpha}(X, Z) = |M_{\alpha}(Z) - Z|$ is a direct measure of manipulation costs.\footnote{Depending on the context, additional costs can be embedded in the experimental design. For example, in the experiment documented in Thornton (2008), besides the monetary incentives for learning HIV results, there were considerably high costs for testing, counseling/giving results, and selling condoms (see Table 12).} For encouragement rule $\alpha$, we define its budget by aggregating the costs for individuals who actually take up the treatment: $B(\alpha) = E_P[C_{\alpha(X,Z)}(X, Z) \cdot p(X, M_{\alpha(X,Z)}(Z))]$. Consider settings in which the cost of implementing any encouragement rule cannot exceed $\kappa$. The policymaker now solves a constrained optimization problem:

$$\max_{\alpha : \mathcal{X} \times \mathcal{Z} \to \mathbb{R}} W(\alpha) \text{ s.t. } B(\alpha) \leq \kappa.$$ 

As in Section 3, we specialize to binary encouragement rules when discussing the performance of statistical decision rules and denote the cost function by $B(G)$ in place of $B(\alpha^G)$. Given a class $\mathcal{G}$ of feasible encouragement rules,\footnote{We implicitly assume that there exists $G \in \mathcal{G}$ such that $B(G) \leq \kappa$.} the constrained optimal encouragement rule, if $P$ is known, is

$$G^*_P \in \arg \max_{G \in \mathcal{G}} W(G) \text{ s.t. } B(G) \leq \kappa,$$

where

$$B(G) = E_P[1\{(X, Z) \in G\} \cdot C_{\alpha_1}(X, Z) \cdot p(X, M_{\alpha_1}(Z)) + 1\{(X, Z) \notin G\} \cdot C_{\alpha_0}(X, Z) \cdot p(X, M_{\alpha_0}(Z))].$$

In the presence of endogenous treatment selection, the policymaker has only imperfect information
about whether an eligibility criterion satisfies the budget constraint. Hence, despite the context-specific forms of $W(G)$ and $B(G)$, the analysis falls into the framework of Sun (2021) in that the policymaker faces a harsh budget constraint, and both the benefit and cost of implementing an eligibility criterion are unknown and need to be estimated.\footnote{If the policymaker wants to account for cost without imposing a fixed budget, she can redefine the social welfare criterion as $W(G) - B(G)$. Then, the analysis in Section 3 applies.} It is helpful to introduce two desirable properties for statistical decision rules in the current setting: \textit{asymptotic optimality} and \textit{asymptotic feasibility}.

\textbf{Definition 1.} A statistical decision rule $\hat{G}$ is uniformly asymptotically optimal over the class $\mathcal{P}$ of DGPs if, for any $\epsilon > 0,$

$$\lim_{n \to \infty} \sup_{P \in \mathcal{P}} \Pr_{P^n}\{W(\hat{G}) - W(G^*_P) < -\epsilon\} = 0.$$ 

A statistical decision rule $\hat{G}$ is uniformly asymptotically feasible over the class $\mathcal{P}$ of DGPs if, for any $\epsilon > 0,$

$$\lim_{n \to \infty} \sup_{P \in \mathcal{P}} \Pr_{P^n}\{B(\hat{G}) - \kappa > \epsilon\} = 0.$$ 

Intuitively, with a large enough sample size, asymptotic optimality imposes that $\hat{G}$ is unlikely to achieve strictly lower welfare than $G^*_P,$ while asymptotic feasibility imposes that $\hat{G}$ is unlikely to strictly violate the budget constraint. It is possible to construct a statistical decision rule that allows both undesirable events to occur, but with diminishing probability, and thus simultaneously achieves both asymptotic optimality and asymptotic feasibility.\footnote{There is an asymmetry in Sun (2021)'s definition of these two properties. Asymptotic optimality only requires the population welfare of a statistical decision rule to concentrate around the optimal value from below. By contrast, asymptotic feasibility requires the statistical decision rule to satisfy the population budget constraint without any slackness and thus is extremely sensitive to sampling uncertainty. In consequence, Sun (2021) proves the negative result that it is impossible for any statistical decision rule to uniformly satisfy both properties.} A natural candidate is the \textit{budget-constrained EWM encouragement rule} defined as a solution to the sample version of the population constrained optimization problem:

$$\hat{G}_{\text{BEWM}} \in \arg \max_{G \in \mathcal{G}} \hat{W}_n(G) \text{ s.t. } \hat{B}_n(G) \leq \kappa,$$

(10)
where

\[
\hat{W}_n(G) = E_n \left[ 1\{(X, Z) \in G\} \int_{\hat{p}(X, M_{\alpha_1}(Z))} \hat{MTE}(u, X) \, du \right],
\]

\[
\hat{B}_n(G) = E_n [1\{(X, Z) \in G\} \cdot C_{\alpha_1}(X, Z) \cdot \hat{p}(X, M_{\alpha_1}(Z)) + 1\{(X, Z) / \in G\} \cdot C_{\alpha_0}(X, Z) \cdot \hat{p}(X, M_{\alpha_0}(Z))].
\]

We set \( \hat{G}_{BEWM} = \emptyset \) if no \( G \in \mathcal{G} \) satisfies \( \hat{B}_n(G) \leq \kappa \). To analyze the theoretical properties of the budget-constrained EWM encouragement rule, we impose the following assumptions.

**Assumption 5** (Boundedness and VC-class)

5.1 \(|\int_0^1 \text{MTE}(u, X) \, du| \leq \tilde{M} < \infty, C_{\alpha_1}(X, Z) \leq \tilde{M} < \infty, \text{and } C_{\alpha_0}(X, Z) \leq \tilde{M} < \infty \) almost surely.

5.2 \( \mathcal{G} \) has a finite VC-dimension \( V < \infty \) and is countable.

We let \( \mathcal{P}'(\tilde{M}) \) denote the class of DGPs that satisfy Assumptions 1 and 5.1. Theorem 3 asserts that the budget-constrained EWM encouragement rule satisfies both properties in Definition 1 over a reasonable class of DGPs; a proof is provided in Appendix A.4.

**Theorem 3.** Suppose Assumptions 1, 3, 4, and 5 hold. Then the budget-constrained EWM encouragement rule \( \hat{G}_{BEWM} \) is uniformly asymptotically optimal and uniformly asymptotically feasible over \( \mathcal{P}'(\tilde{M}) \cap \mathcal{P}_p \cap \mathcal{P}_{\text{MTE}} \).

## 5 Empirical Application

In this section, we apply the feasible EWM encouragement rule and the budget-constrained EWM encouragement rule to provide guidance on how to encourage upper secondary schooling, using data from the third wave of the Indonesian Family Life Survey (IFLS) fielded from June through November 2000. Carneiro, Lokshin, and Umapathi (2017) used this dataset to study the returns to upper secondary schooling in Indonesia. We follow Carneiro et al. (2017) in restricting our sample
to males aged 25–60 who are employed and who have non-missing reported wage and schooling information. This subsample consists of 2,104 individuals.\textsuperscript{12}

We specify the relevant variables in our framework. The outcome $Y$ is the log of hourly wages constructed from self-reported monthly wages and hours worked per week. The treatment $D$ is an indicator of attendance of upper secondary school or higher, corresponding to 10 or more years of completed education. The first instrument $Z_1$ that can be manipulated is the lowest fee per continuing student, in thousands of rupiah, among secondary schools in the community of current residence.\textsuperscript{13} The second instrument $Z_2$ is the distance, in kilometers, from the office of the community head of current residence to the nearest secondary school, which we define as the secondary school closest to the office of the community head.\textsuperscript{14} We collect $Z = (Z_1, Z_2)'$. The covariates $X$ include age, age squared, an indicator of rural residence, distance from the office of the community head of residence to the nearest health post, and indicators for religion, parental education, and the province of residence. Table D.1 in Appendix D presents sample averages for these variables.

We focus on binary encouragement rules that assign eligibility for receiving a manipulation of $Z_1$ based on $(Z_1, Z_2)$. We consider the class of linear encouragement rules:

$$G_{LES} = \{ G : G = \{(x, z) \in X \times Z : \lambda_0 + \lambda_1 \cdot Z_1 + \lambda_2 \cdot Z_2 > 0 \}, \lambda_0, \lambda_1, \lambda_2 \in \mathbb{R} \}.$$ 

We specify the manipulation function as $M_\alpha(Z) = (M_{\alpha,1}(Z_1), Z_2)'$, where $M_{\alpha,1}(Z_1) = (Z_1 - \alpha) \cdot 1\{Z_1 \geq \alpha\}$. Here, $M_\alpha(Z)$ describes a tuition subsidy of up to $\alpha$. The policymaker a priori sets $\alpha_0 = 0$ and chooses from $\alpha_1 \in \{2.5, 22.25\}$, which correspond to the median and maximum of $Z_1$, respectively. We specify the cost function as $C_\alpha(X, Z) = |M_\alpha(Z) - Z|$ and the budget constraint

\textsuperscript{12}The subsample used in Carneiro et al. (2017) does not contain the tuition fee variable, which plays a central role in our framework as the manipulatable instrument. Hence, we followed their descriptions to construct our subsample from raw data downloaded from the RAND Corporation website.

\textsuperscript{13}The term “community” refers to the lowest-level administrative division in Indonesia. A community can either be a desa (village) or a kelurahan (urban community).

\textsuperscript{14}The validity of an instrument constructed in this way can be controversial. Each individual’s tuition and proximity are based on their current residence rather than their residence at the time of the secondary schooling decision. Educated individuals may move to more urban areas with more schools and higher tuition fees. Nonetheless, we note that the instrumental variable independence assumption for unrestricted instruments has testable implications, which are the generalized instrumental inequalities proposed by Kédagni and Mourifié (2020). Their tests suggest that the independence assumption between individuals’ potential earnings and college tuition at age 17 is not rejected by the data from the 1979 National Longitudinal Survey of Youth used in Heckman, Tobias, and Vytlacil (2001).
as $\kappa = 0.28$, which is about one-tenth of the average hourly wage.

We estimate the propensity score $p(X, Z)$ from a logit regression of $D$ on $X$, $Z$, $Z_1 \cdot Z_2$, and interactions between $Z$ and $X$.\textsuperscript{15} We trim observations for which $\hat{p}(X, Z)$ is below 0.0408 or above 0.9863 to restrict our estimates to the common support of $\hat{p}(X, Z)$.\textsuperscript{16} As demonstrated in Section 4.1, we estimate the instrument-specific MTE using the parametric regression function (7) with $Z_2$ controlled for as covariates. Using leave-one-out cross-validation with the mean-squared error criterion, we select a second-order polynomial in $\hat{p}$ (i.e., $K = 2$) and thus a linear MTE curve. Figure D.1 in Appendix D shows that the MTE estimates based on a second-order polynomial are more consistent with a local quadratic regression than polynomials of higher orders.

We compute the feasible EWM encouragement rule $\hat{G}_{\text{FEWM}}$ in (4) and the budget-constrained EWM encouragement rule $\hat{G}_{\text{BEWM}}$ in (10) using the CPLEX mixed integer optimizer. Based on the decomposition result in Corollary 1, we report in Table 1 the estimated welfare gains, the average change in treatment take-up, and the PRTE of alternative encouragement rules, as well as the estimated proportion of eligible individuals. The seemingly favorable tuition subsidy has little effect on overall upper secondary school attendance when applied to everyone, resulting in a welfare gain of only a small magnitude. The feasible EWM encouragement rule and the budget-constrained EWM encouragement rule achieve higher welfare gains by targeting a subpopulation with both a greater increase in treatment take-up and higher PRTE.

\textsuperscript{15}This specification of propensity score is an adaption of that considered by Carneiro et al. (2017) and Sasaki and Ura (2021), who use a single instrument $Z_2$.

\textsuperscript{16}In doing so, we delete 46 observations, corresponding to 2.19 percent of the sample.
Table 1: Estimated Welfare Gain of Alternative Encouragement Rules

| Policy          | Share of Eligible Population | Avg. Change in Est. Welfare Gain | Treatment Take-up | PRTE |
|-----------------|------------------------------|---------------------------------|-------------------|------|
| Panel A: $\alpha_1 = 2.5$ |                              |                                 |                   |      |
| $\hat{G}_{FEWM}$       | 0.394                        | 0.0140                          | 0.0177            | 0.791|
| $\hat{G}_{BEWM}$ ($\kappa = 0.28$) | 0.283                        | 0.0101                          | 0.0140            | 0.722|
| $G = X \times Z$           | 1                            | -0.0003                         | 0.0021            | -0.119|

Panel B: $\alpha_1 = 22.25$

| Policy          | Share of Eligible Population | Avg. Change in Est. Welfare Gain | Treatment Take-up | PRTE |
|-----------------|------------------------------|---------------------------------|-------------------|------|
| $\hat{G}_{FEWM}$       | 0.392                        | 0.0234                          | 0.0324            | 0.721|
| $\hat{G}_{BEWM}$ ($\kappa = 0.28$) | 0.288                        | 0.0091                          | 0.0144            | 0.635|
| $G = X \times Z$           | 1                            | 0.0056                          | 0.0140            | 0.400|

We plot the feasible EWM encouragement rule and the budget-constrained EWM encouragement rule in Panels A and B of Figure 1, respectively. The shaded areas indicate the subpopulations to whom the tuition subsidy should be assigned. For both subsidy levels, the feasible EWM encouragement rule gives eligibility to individuals facing relatively high tuition fees and living relatively close to the nearest secondary school. The subpopulations targeted by the budget-constrained EWM encouragement rule shrink to the left. When the subsidy level $\alpha_1$ is increased from 2.5 to 22.25, the budget-constrained EWM encouragement rule tends to prioritize individuals facing relatively low tuition fees.
Figure 1: Targeted Subpopulation under Alternative Encouragement Rules

Figure 2 gives a sense of why the subpopulations indicated by the shaded areas in Figure 1 are targeted. It visualizes how the impact of going from the status quo to a full tuition waiver (i.e., \( \alpha_1 = 22.25 \)) differs across individuals with different values of \((Z_1, Z_2)\) in terms of treatment take-up and treatment effects. Panel A of Figure 2 displays the level sets of changes in average treatment take-up conditional on \((Z_1, Z_2)\), namely the level sets of \((z_1, z_2) \mapsto E_n[\hat{p}(X, M_{\alpha_1}(z)) - \hat{p}(X, z)]\). We can see that basically only individuals with low \(Z_2\) are induced into treatment. Panel B of Figure 2 displays the level sets of average treatment effects for those induced to switch treatment status conditional on \((Z_1, Z_2)\), namely the level sets of

\[
(z_1, z_2) \mapsto \frac{E_n \left[ \int \hat{p}(X, M_{\alpha_1}(z)) \cdot \text{MTE}_1(u_1, X, z_2) \, du_1 \right]}{E_n[\hat{p}(X, M_{\alpha_1}(z)) - \hat{p}(X, z)]}.
\]

We can see that individuals with low \(Z_1\) have negative treatment effects for \(Z_2\) on the left of the zero contour line for changes in treatment take-up. Put together, absent budget constraints, individuals in the upper-left corner are prioritized for the full tuition fee waiver. On the other hand, contour lines where \(Z_1\) is high are rather steep in both panels, implying that individuals with higher \(Z_1\) induce higher manipulation costs for the same level of welfare gains. Consequently, the optimal policy under budget constraints trades off welfare gains against costs and gives up individuals with high \(Z_1\).
6 Conclusion

In this paper, we propose a framework for policy learning that allows for endogenous treatment selection by leveraging a continuous instrumental variable. To deal with failure of unconfoundedness when identifying the social welfare criterion, we incorporate the MTE function. To deal with imperfect compliance when designing policies, we consider encouragement rules in addition to treatment assignment rules. We apply the representation of the social welfare criterion of encouragement rules to the EWM method and derive convergence rates of the worst-case regret. We also consider extensions allowing for multiple instruments and budget constraints. We illustrate the EWM encouragement rule using data from the Indonesian Family Life Survey.

There are several potential extensions and directions for future research. First, our point identification of the social welfare criterion imposes several restrictive parametric assumptions for extrapolations. To relax these assumptions, one might consider incorporating the approach to partial-identifying policy parameters proposed by Mogstad et al. (2018). Second, to improve the convergence rate of the regret, the technique of a doubly robust estimator of the social welfare criterion proposed by Athey and Wager (2021) could be adapted from inverse-propensity weighting to par-
tially linear models.

References

Athey, S. and Wager, S. (2021), “Policy Learning With Observational Data,” *Econometrica*, 89, 133–161.

Bertsimas, D. and McCord, C. (2018), “Optimization over Continuous and Multi-dimensional Decisions with Observational Data,” in *Advances in Neural Information Processing Systems*, eds. Bengio, S., Wallach, H., Larochelle, H., Grauman, K., Cesa-Bianchi, N., and Garnett, R., Curran Associates, Inc., vol. 31.

Bhattacharya, D. and Dupas, P. (2012), “Inferring Welfare Maximizing Treatment Assignment under Budget Constraints,” *Journal of Econometrics*, 167, 168–196.

Björklund, A. and Moffitt, R. (1987), “The Estimation of Wage Gains and Welfare Gains in Self-Selection Models,” *The Review of Economics and Statistics*, 69, 42–49.

Boucheron, S., Lugosi, G., and Massart, P. (2013), *Concentration Inequalities: A Nonasymptotic Theory of Independence*, Oxford University Press.

Brinch, C. N., Mogstad, M., and Wiswall, M. (2017), “Beyond LATE with a Discrete Instrument,” *Journal of Political Economy*, 125, 985–1039.

Byambadalai, U. (2022), “Identification and Inference for Welfare Gains without Unconfoundedness,” ArXiv: 2207.04314.

Carneiro, P., Heckman, J. J., and Vytlacil, E. (2010), “Evaluating Marginal Policy Changes and the Average Effect of Treatment for Individuals at the Margin,” *Econometrica*, 78, 377–394.

Carneiro, P., Heckman, J. J., and Vytlacil, E. J. (2011), “Estimating Marginal Returns to Education,” *American Economic Review*, 101, 2754–81.
Carneiro, P. and Lee, S. (2009), “Estimating Distributions of Potential Outcomes Using Local Instrumental Variables with an Application to Changes in College Enrollment and Wage Inequality,” *Journal of Econometrics*, 149, 191–208.

Carneiro, P., Lokshin, M., and Umapathi, N. (2017), “Average and Marginal Returns to Upper Secondary Schooling in Indonesia,” *Journal of Applied Econometrics*, 32, 16–36.

Chen, X. and Christensen, T. M. (2015), “Optimal Uniform Convergence Rates and Asymptotic Normality for Series Estimators under Weak Dependence and Weak Conditions,” *Journal of Econometrics*, 188, 447–465, heterogeneity in Panel Data and in Nonparametric Analysis in honor of Professor Cheng Hsiao.

Chen, Y.-C. and Xie, H. (2022), “Personalized Subsidy Rules,” ArXiv: 2202.13545.

Cornelissen, T., Dustmann, C., Raute, A., and Schönb erg, U. (2016), “From LATE to MTE: Alternative Methods for the Evaluation of Policy Interventions,” *Labour Economics*, 41, 47–60, sOLE/EALE conference issue 2015.

Cui, Y. and Tchetgen, E. T. (2021), “A Semiparametric Instrumental Variable Approach to Optimal Treatment Regimes Under Endogeneity,” *Journal of the American Statistical Association*, 116, 162–173, pMID: 33994604.

Dupas, P. (2014), “Short-Run Subsidies and Long-Run Adoption of New Health Products: Evidence From a Field Experiment,” *Econometrica*, 82, 197–228.

Heckman, J., Tobias, J. L., and Vytlacil, E. (2001), “Four Parameters of Interest in the Evaluation of Social Programs,” *Southern Economic Journal*, 68, 211–223.

Heckman, J. J., Urzua, S., and Vytlacil, E. (2006), “Understanding Instrumental Variables in Models with Essential Heterogeneity,” *The Review of Economics and Statistics*, 88, 389–432.

Heckman, J. J. and Vytlacil, E. (2001), “Policy-Relevant Treatment Effects,” *The American Economic Review*, 91, 107–111.
Heckman, J. J. and Vytlacil, E. J. (1999), “Local Instrumental Variables and Latent Variable Models for Identifying and Bounding Treatment Effects,” Proceedings of the National Academy of Sciences of the United States of America, 96, 4730–4734.

Hirano, K. and Porter, J. R. (2009), “Asymptotics for Statistical Treatment Rules,” Econometrica, 77, 1683–1701.

— (2020), “Chapter 4 - Asymptotic Analysis of Statistical Decision Rules in Econometrics,” in Handbook of Econometrics, Volume 7A, eds. Durlauf, S. N., Hansen, L. P., Heckman, J. J., and Matzkin, R. L., Elsevier, pp. 283–354.

Imbens, G. W. and Angrist, J. D. (1994), “Identification and Estimation of Local Average Treatment Effects,” Econometrica, 62, 467–475.

Kallus, N. and Zhou, A. (2018), “Policy Evaluation and Optimization with Continuous Treatments,” in Proceedings of the Twenty-First International Conference on Artificial Intelligence and Statistics, eds. Storkey, A. and Perez-Cruz, F., PMLR, vol. 84 of Proceedings of Machine Learning Research, pp. 1243–1251.

Kasy, M. (2016), “Partial Identification, Distributional Preferences, and the Welfare Ranking of Policies,” The Review of Economics and Statistics, 98, 111–131.

Kitagawa, T. and Tetenov, A. (2018a), “Supplement to ‘Who Should Be Treated? Empirical Welfare Maximization Methods for Treatment Choice’,” Econometrica Supplemental Material, 86.

— (2018b), “Who Should Be Treated? Empirical Welfare Maximization Methods for Treatment Choice,” Econometrica, 86, 591–616.

Kédagni, D. and Mourifié, I. (2020), “Generalized Instrumental Inequalities: Testing the Instrumental Variable Independence Assumption,” Biometrika, 107, 661–675.
Manski, C. F. (2004), “Statistical Treatment Rules for Heterogeneous Populations,” *Econometrica*, 72, 1221–1246.

Mbakop, E. and Tabord-Meehan, M. (2021), “Model Selection for Treatment Choice: Penalized Welfare Maximization,” *Econometrica*, 89, 825–848.

Mogstad, M., Santos, A., and Torgovitsky, A. (2018), “Using Instrumental Variables for Inference About Policy Relevant Treatment Parameters,” *Econometrica*, 86, 1589–1619.

Mogstad, M. and Torgovitsky, A. (2018), “Identification and Extrapolation of Causal Effects with Instrumental Variables,” *Annual Review of Economics*, 10, 577–613.

Mogstad, M., Torgovitsky, A., and Walters, C. R. (2021), “The Causal Interpretation of Two-Stage Least Squares with Multiple Instrumental Variables,” *American Economic Review*, 111, 3663–98.

Pu, H. and Zhang, B. (2021), “Estimating Optimal Treatment Rules with an Instrumental Variable: A Partial Identification Learning Approach,” *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 83, 318–345.

Qiu, H., Carone, M., Sadikova, E., Petukhova, M., Kessler, R. C., and Luedtke, A. (2021), “Optimal Individualized Decision Rules Using Instrumental Variable Methods,” *Journal of the American Statistical Association*, 116, 174–191, pMID: 33731969.

Rubin, D. B. (1974), “Estimating Causal Effects of Treatments in Randomized and Nonrandomized Studies,” *Journal of Educational Psychology*, 66, 688–701.

Sasaki, Y. and Ura, T. (2020), “Welfare Analysis via Marginal Treatment Effects,” ArXiv: 2012.07624.

— (2021), “Estimation and Inference for Policy Relevant Treatment Effects,” *Journal of Econometrics*.

Sun, L. (2021), “Empirical Welfare Maximization with Constraints,” ArXiv: 2103.15298.

Thornton, R. L. (2008), “The Demand for, and Impact of, Learning HIV Status,” *American Economic Review*, 98, 1829–63.
van der Vaart, A. and Wellner, J. (1996), *Weak Convergence and Empirical Processes: With Applications to Statistics*, Springer Series in Statistics, Springer.

Vytlacil, E. (2002), “Independence, Monotonicity, and Latent Index Models: An Equivalence Result,” *Econometrica*, 70, 331–341.

Wang, L. and Tchetgen Tchetgen, E. (2018), “Bounded, Efficient and Multiply Robust Estimation of Average Treatment Effects Using Instrumental Variables,” *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 80, 531–550.
Appendix A  Proofs

For a binary encouragement rule $G$ described in Section 2.3, define the welfare contrast relative to the baseline policy that allocates everyone to $\alpha_0$ and its empirical analogue as

$$\bar{W}(G) = EP\left[1\{(X, Z) \in G\} \int_{p(X, M_{\alpha_1}(Z))}^{p(X, M_{\alpha_0}(Z))} MTE(u, X) \, du\right] = W(G) - W_0,$$

$$\bar{W}_n(G) = E_n\left[1\{(X, Z) \in G\} \int_{p(X, M_{\alpha_1}(Z))}^{p(X, M_{\alpha_0}(Z))} MTE(u, X) \, du\right].$$

Appendix A.1  Proof of Theorem 1

Proof. We have

$$W(\alpha) = EP[Y(1) \cdot D^\alpha + Y(0) \cdot (1 - D^\alpha)]$$

$$= EP[Y(1) \cdot 1\{p(X, M_{\alpha}(X,Z)(Z)) \geq U\} + Y(0) \cdot 1\{p(X, M_{\alpha}(X,Z)(Z)) < U\}]$$

$$= EP[Y(1) \cdot 1\{p(X, Z) \geq U\} + Y(0) \cdot 1\{p(X, Z) < U\}]$$

$$+ EP[(Y(1) - Y(0)) \cdot (1\{p(X, M_{\alpha}(X,Z)(Z)) \geq U\} - 1\{p(X, Z) \geq U\})]$$

$$= EP[Y] + EP[(Y(1) - Y(0)) \cdot (1\{p(X, M_{\alpha}(X,Z)(Z)) \geq U\} - 1\{p(X, Z) \geq U\})]. \quad (A.1)$$

By the law of iterated expectations, the second term in (A.1) can be written as

$$EP[(Y(1) - Y(0)) \cdot (1\{p(X, M_{\alpha}(X,Z)(Z)) \geq U\} - 1\{p(X, Z) \geq U\})]$$

$$= EP[EP[Y(1) - Y(0)|X, Z, U] \cdot (1\{p(X, M_{\alpha}(X,Z)(Z)) \geq U\} - 1\{p(X, Z) \geq U\})]$$

$$= EP[MTE(U, X) \cdot (1\{p(X, M_{\alpha}(X,Z)(Z)) \geq U\} - 1\{p(X, Z) \geq U\})]$$

$$= EP[EP[MTE(U, X) \cdot (1\{p(X, M_{\alpha}(X,Z)(Z)) \geq U\} - 1\{p(X, Z) \geq U\})|X, Z]]$$

$$= EP\left[\int_{0}^{1} MTE(u, X) \cdot (1\{p(X, M_{\alpha}(X,Z)(Z)) \geq u\} - 1\{p(X, Z) \geq u\}) \, du\right],$$

where the second equality follows from Assumption 1.2 and the definition of the MTE, and the fourth equality follows from (3).
Appendix A.2  Proof of Corollary 3

Proof. Define

\[ f(X, Z; G) = 1\{(X, Z) \in G\} \int_{p(X, M_0(Z))}^{p(X, M_1(Z))} \text{MTE}(u, X) \, du \]

and

\[ \mathcal{F} = \{ f(\cdot; G) : G \in \mathcal{G} \}. \]

By Assumption 2.1, \( \mathcal{F} \) is a class of uniformly bounded functions with \( \| f \|_\infty \leq \bar{M} \) for all \( f \in \mathcal{F} \). By Assumption 2.2 and Lemma A.1 of Kitagawa and Tetenov (2018a), \( \mathcal{F} \) is a VC-subgraph class of functions with VC-dimension less than or equal to \( V \). Then, we can apply Lemma A.4 of Kitagawa and Tetenov (2018a) to obtain

\[ E_P^n \left[ \sup_{f \in \mathcal{F}} |E_n(f) - E_P(f)| \right] \leq C_1 \bar{M} \sqrt{\frac{V}{n}}, \tag{A.2} \]

where \( C_1 \) is a universal constant. Then, we have

\[ \sup_{G \in \mathcal{G}} |\bar{W}_n(G) - \bar{W}(G)| = \sup_{f \in \mathcal{F}} |E_n(f) - E_P(f)|. \tag{A.3} \]

Following the derivations in Kitagawa and Tetenov (2018b, Eq. (2.2)), we have for any \( \hat{G} \in \mathcal{G} \),

\[ W(\hat{G}) - W(\hat{G}_{EWM}) = W(\hat{G}) - \bar{W}_n(\hat{G}_{EWM}) + \bar{W}_n(\hat{G}_{EWM}) - W(\hat{G}_{EWM}) \]

\[ \leq W(\hat{G}) - \bar{W}_n(\hat{G}) + \sup_{G \in \mathcal{G}} |\bar{W}_n(G) - \bar{W}(G)| \]

\[ (\because \bar{W}_n(\hat{G}_{EWM}) \geq \bar{W}_n(\hat{G})) \]

\[ \leq 2 \sup_{G \in \mathcal{G}} |\bar{W}_n(G) - \bar{W}(G)|. \]
It follows from (A.2) and (A.3) that

$$ \sup_{P \in \mathcal{P}(M)} E_P^n [W_G^* - W(\hat{G}_{\text{FEWM}})] \leq 2C_1 M \sqrt{\frac{V}{n}}. $$

\[ \square \]

**Appendix A.3  Proof of Theorem 2**

*Proof.* We have for any $\hat{G} \in \mathcal{G}$,

$$ W(\hat{G}) - W(\hat{G}_{\text{FEWM}}) $$

$$ \leq \hat{W}(\hat{G}) - \hat{W}_n(\hat{G}) + \hat{W}_n(\hat{G}_{\text{FEWM}}) - \hat{W}(\hat{G}_{\text{FEWM}}) $$

$$ + \hat{W}_n(\hat{G}) - \hat{W}_n(\hat{G}_{\text{FEWM}}) + \hat{W}_n(\hat{G}_{\text{FEWM}}) $$

$$ \leq 2 \sup_{\hat{G} \in \mathcal{G}} |\hat{W}_n(\hat{G}) - \hat{W}(\hat{G})| + 2 \sup_{\hat{G} \in \mathcal{G}} |\hat{W}_n(\hat{G}) - \hat{W}_n(\hat{G})|, $$

where $\hat{W}_n(\hat{G})$ is defined as in (11). In view of (A.2) and (A.3), it suffices to show that

$$ \limsup_{n \to \infty} \sup_{P \in \mathcal{P} \cap \mathcal{P}_{\text{MTE}}} (\phi_n \wedge \psi_n) E_P^n \left[ \sup_{\hat{G} \in \mathcal{G}} |\hat{W}_n(\hat{G}) - \hat{W}(\hat{G})| \right] < \infty. \quad (A.4) $$

By the triangle inequality and Assumption 3.1,

$$ \sup_{\hat{G} \in \mathcal{G}} |\hat{W}_n(\hat{G}) - \hat{W}(\hat{G})| \leq E_n \left[ \int_{p(\hat{X},\hat{M}_{\alpha_1}(Z))} \hat{\text{MTE}}(u, X) \, du - \int_{p(\hat{X},\hat{M}_{\alpha_0}(Z))} \hat{\text{MTE}}(u, X) \, du \right] $$

$$ \leq E_n \left[ \int_{p(\hat{X},\hat{M}_{\alpha_1}(Z))} \hat{\text{MTE}}(u, X) - \text{MTE}(u, X) \, du \right] $$

$$ \quad + E_n \left[ \int_{p(\hat{X},\hat{M}_{\alpha_0}(Z))} \text{MTE}(u, X) \, du \right] + E_n \left[ \int_{p(\hat{X},\hat{M}_{\alpha_0}(Z))} \text{MTE}(u, X) \, du \right] $$

$$ \leq E_n \left[ \int_{0}^{1} \left( \text{MTE}(u, X) - \text{MTE}(u, X) \right) \, du \right] $$

$$ \quad + E_n \left[ \int_{p(\hat{X},\hat{M}_{\alpha_1}(Z))} \text{MTE}(u, X) \, du \right] + E_n \left[ \int_{p(\hat{X},\hat{M}_{\alpha_0}(Z))} \text{MTE}(u, X) \, du \right]. $$
For $\tilde{Z} \in \{M_{a_0}(Z), M_{a_1}(Z)\}$,

$$
E_n \left[ \int p(X, \tilde{Z}) MTE(u, X) \, du \right] \leq E_n \left[ \int \hat{p}(X, \tilde{Z}) |MTE(u, X)| \, du \right]
$$

$$
\leq E_n \left[ C(X) |\hat{p}(X, \tilde{Z}) - p(X, \tilde{Z})| \right]
$$

$$
\leq E_n[C(X)] \sup_{(x, z) \in X \times \tilde{Z}} |\hat{p}(x, z) - p(x, z)|
$$

$$
\leq (E_n - E_P)[C(X)] \sup_{(x, z) \in X \times \tilde{Z}} |\hat{p}(x, z) - p(x, z)|
$$

$$
+ E_P[C(X)] \sup_{(x, z) \in X \times \tilde{Z}} |\hat{p}(x, z) - p(x, z)|,
$$

where the second inequality follows from Assumption 4.1. By Assumption 4.2,

$$
\limsup_{n \to \infty} \sup_{P \in \mathcal{P}_{\text{MTE}}} n^{1/2} E_P^n[(E_n - E_P)[C(X)]] < \infty.
$$

Combining the above, by Assumptions 3.2, 4.2 and 4.3, (A.4) holds. \hfill \square

**Appendix A.4  Proof of Theorem 3**

**Proof.** For any $\epsilon > 0$,

$$
\Pr_{P^n} \{ W(G_P^n) - W(\hat{G}_{\text{BEWM}}) > \epsilon \} = \Pr_{P^n} \{ \hat{W}(G_P^n) - \hat{W}(\hat{G}_{\text{BEWM}}) > \epsilon, \hat{B}_n(G_P^n) \leq \kappa \}
$$

$$
+ \Pr_{P^n} \{ \hat{W}(G_P^n) - \hat{W}(\hat{G}_{\text{BEWM}}) > \epsilon, \hat{B}_n(G_P^n) > \kappa \}
$$

$$
= P_1 + P_2.
$$

Noting that $\hat{B}_n(G_P^n) \leq \kappa$ implies $\hat{W}_n(\hat{G}_{\text{BEWM}}) \geq \hat{W}_n(G_P^n)$,

$$
P_1 \leq \Pr_{P^n} \{ \hat{W}(G_P^n) - \hat{W}_n(G_P^n) + \hat{W}_n(\hat{G}_{\text{BEWM}}) - \hat{W}(\hat{G}_{\text{BEWM}}) > \epsilon, \hat{B}_n(G_P^n) \leq \kappa \}
$$

$$
\leq \Pr_{P^n} \left\{ 2 \sup_{G \in \mathcal{G}} |\hat{W}(G) - \hat{W}_n(G)| > \epsilon \right\}.
$$
On the other hand, \( B(G_P^*) \leq \kappa \), and thus

\[
P_2 \leq \Pr_{P^n} \{ \hat{B}_n(G_P^*) > \kappa \} \leq \Pr_{P^n} \{ \hat{B}_n(G_P^*) > B(G_P^*) \}.
\]

Combining the above, we have

\[
\Pr_{P^n} \{ W(G_P^*) - W(\hat{G}_{BEWM}) > \epsilon \} \\
\leq \Pr_{P^n} \left\{ \sup_{G \in \mathcal{G}} |\bar{W}(G) - \hat{W}_n(G)| > \epsilon \right\} + \Pr_{P^n} \{ \hat{B}_n(G_P^*) > B(G_P^*) \} \\
\leq \Pr_{P^n} \left\{ \sup_{G \in \mathcal{G}} |\bar{W}(G) - \hat{W}_n(G)| > \epsilon \right\} + \sup_{\epsilon' > 0} \Pr_{P^n} \{ \hat{B}_n(G_P^*) - B(G_P^*) > \epsilon' \} \\
\leq \Pr_{P^n} \left\{ \sup_{G \in \mathcal{G}} |\bar{W}(G) - \hat{W}_n(G)| > \epsilon \right\} + \sup_{\epsilon' > 0} \Pr_{P^n} \left\{ \sup_{G \in \mathcal{G}} |B(G) - \hat{B}_n(G)| > \epsilon' \right\}.
\]  \hspace{1cm} (A.5)

Also, for any \( \epsilon > 0 \),

\[
\Pr_{P^n} \{ B(\hat{G}_{BEWM}) - \kappa > \epsilon \} \leq \Pr_{P^n} \{ B(\hat{G}_{BEWM}) - \hat{B}_n(\hat{G}_{BEWM}) > \epsilon \} \quad (\because \hat{B}_n(\hat{G}_{BEWM}) \leq \kappa) \\
\leq \Pr_{P^n} \left\{ \sup_{G \in \mathcal{G}} |B(G) - \hat{B}_n(G)| > \epsilon \right\}.
\]  \hspace{1cm} (A.6)

Define

\[
B_n(G) = E_n[1\{(X, Z) \in G\} \cdot |M_{\alpha_1}(Z) - Z| \cdot p(X, M_{\alpha_1}(Z)) + 1\{(X, Z) \notin G\} \cdot |M_{\alpha_0}(Z) - Z| \cdot p(X, M_{\alpha_0}(Z))].
\]

By the triangle inequality, for any \( G \in \mathcal{G} \),

\[
|\bar{W}(G) - \hat{W}_n(G)| \leq |\bar{W}(G) - \bar{W}_n(G)| + |\bar{W}_n(G) - \hat{W}_n(G)|, \]  \hspace{1cm} (A.7)

\[
|B(G) - \hat{B}_n(G)| \leq |B(G) - B_n(G)| + |B_n(G) - \hat{B}_n(G)|. \]  \hspace{1cm} (A.8)
From the same argument as the proof of (A.3), we can apply Lemma A.4 of Kitagawa and Tetenov (2018a) to obtain

\[
\sup_{P \in \mathcal{P}(\hat{M})} E_{P^n} \left[ \sup_{G \in \hat{G}} |W_n(G) - \hat{W}(G)| \right] \leq O(n^{-1/2}), \tag{A.9}
\]

\[
\sup_{P \in \mathcal{P}(\hat{M})} E_{P^n} \left[ \sup_{G \in \hat{G}} |B_n(G) - B(G)| \right] \leq O(n^{-1/2}). \tag{A.10}
\]

On the other hand, note that

\[
\sup_{G \in \hat{G}} |B_n(G) - \hat{B}_n(G)| \leq E_n[C_{\alpha_1}(X, Z) \cdot |\hat{p}(X, M_{\alpha_1}(Z)) - p(X, M_{\alpha_1}(Z))| + C_{\alpha_0}(X, Z) \cdot |\hat{p}(X, M_{\alpha_0}(Z)) - p(X, M_{\alpha_0}(Z))|] \\
\leq 2\hat{M} \sup_{(x, z) \in \mathcal{X} \times \mathcal{Z}} |\hat{p}(x, z) - p(x, z)|,
\]

where the second inequality follows from Assumption 5.1. Hence, by (5),

\[
\limsup_{n \to \infty} \sup_{P \in \mathcal{P}_p} \psi_n E_{P^n} \left[ \sup_{G \in \hat{G}} |B_n(G) - \hat{B}_n(G)| \right] < \infty. \tag{A.11}
\]

By Markov’s inequality and (A.4)–(A.11), we have that, for any \(\epsilon > 0\),

\[
\lim_{n \to \infty} \sup_{P \in \mathcal{P}(\hat{M}) \cap \mathcal{P}_p \cap \mathcal{P}_{MT}} \Pr_{P^n}\{W(\hat{G}_{\text{BEWM}}) - W(G_p^*) < -\epsilon\} = 0,
\]

\[
\lim_{n \to \infty} \sup_{P \in \mathcal{P}(\hat{M}) \cap \mathcal{P}_p \cap \mathcal{P}_{MT}} \Pr_{P^n}\{B(\hat{G}_{\text{BEWM}}) - \kappa > \epsilon\} = 0.
\]

\[\square\]

Appendix B  Encouragement Rules with a Binary Instrument

In many applications, the instrument is binary by construction. For completeness, we present a discussion of cases in which a binary instrument is intervened upon. An encouragement rule is a mapping from \(\mathcal{X}\) to \(\{0, 1\}\) and can be indexed by its decision set \(G \subset \mathcal{X}\) such that the policymaker assigns individuals with \(X \in G\) to encouragement \(Z = 1\). For example, the encouragement could
be eligibility for welfare programs such as the National Job Training Partnership Act (JTPA) or the 401(k) retirement program.

Under (1), (3), and Assumption 1, the treatment choice under encouragement rule $G$ is given by

$$D^G = 1\{p(X, 1) \geq U\} \cdot 1\{X \in G\} + 1\{p(X, 0) \geq U\} \cdot 1\{X \notin G\}.$$  

We can represent the social welfare criterion as a function of the MTE:

$$W(G) = E_P[Y(1) \cdot D^G + Y(0) \cdot (1 - D^G)]$$

$$= W_0 + E_P \left[ 1\{X \in G\} \int_{p(X, 0)}^{p(X, 1)} \text{MTE}(u, X) \, du \right],$$

where $W_0 = E_P[Y(1) \cdot 1\{p(X, 0) \geq U\} + Y(0) \cdot 1\{p(X, 0) < U\}]$ does not depend on $G$. If we further impose

**Assumption B.1** (IV positivity) $0 < \Pr_P(Z = 1|X) < 1$ almost surely,

we can then derive a more convenient representation of the social welfare criterion:

$$W(G) = E_P[(Y(1) \cdot 1\{p(X, 1) \geq U\} + Y(0) \cdot 1\{p(X, 1) < U\}) \cdot 1\{X \in G\}$$

$$+ (Y(1) \cdot 1\{p(X, 0) \geq U\} + Y(0) \cdot 1\{p(X, 0) < U\}) \cdot 1\{X \notin G\}]$$

$$= E_P[E_P[Y|X, Z = 1] \cdot 1\{X \in G\} + E_P[Y|X, Z = 0] \cdot 1\{X \notin G\}]$$

$$= E_P \left[ \frac{Y \cdot Z}{q(X)} \cdot 1\{X \in G\} + \frac{Y \cdot (1 - Z)}{1 - q(X)} \cdot 1\{X \notin G\} \right],$$

where $q(X) = E_P[Z|X]$. Given a feasible class $\mathcal{G} \subset \{G \subset X\}$ of encouragement rules, the optimal encouragement rule, if $P$ is known, is $G^*_P \in \arg\max_{G \in \mathcal{G}} W(G)$. The optimal encouragement rule is identified without observing the treatment variable $D$. Intuitively, the optimization problem amounts to finding the optimal treatment rule that assigns $Z$ rather than $D$, from an intention-to-treat perspective. In this case, Assumption 1(i) warrants unconfoundedness, and the analysis essentially follows the original EWM framework. It is worth noting that $W(G)$ only measures the
indirect effect of treatment via the impact of encouragement upon eventual treatment receipt. For always-takers and never-takers, no encouragement rule can possibly alter their outcomes.

We proceed to present two extensions analogous to those in Section 4. In Appendix B.1, we allow for multiple instruments. In Appendix B.2, we incorporate a budget constraint.

**Appendix B.1 Multiple Instruments**

Consider a setting in which there are $L$ instruments available. Let $Z_1$ be the binary instrument that can be manipulated, $Z_{-1} \equiv (Z_2, \ldots, Z_L)' \in \mathbb{Z}_{-1} \subset \mathbb{R}^{L-1}$ be a vector of additional instruments, and $Z = (Z_1, Z_{-1})'$. The policymaker can choose an encouragement rule based on $(X, Z_{-1})$, and the set of encouragement rules is indexed by $G \subset \mathcal{X} \times \mathbb{Z}_{-1}$. Under (1), (9), Assumption 1’, and $0 < \Pr_P(Z = 1 | X, Z_{-1}) < 1$ almost surely, the social welfare criterion is given by

$$W(G) = E_P \left[ \frac{YZ_1}{q(X, Z_{-1})} \cdot 1\{(X, Z_{-1}) \in G\} + \frac{Y(1 - Z_1)}{1 - q(X, Z_{-1})} \cdot 1\{(X, Z_{-1}) \notin G\} \right],$$

where $q(X, Z_{-1}) = E_P[Z_1 | X, Z_{-1}]$. The additional instruments $Z_{-1}$ are treated equivalently to covariates $X$. With further assumptions imposed on the selection behavior, we can produce an interpretation of which subpopulation benefits from the encouragement rule. In particular, we assume

**Assumption B.2** (Component-wise monotonic propensity score) For each $x \in \mathcal{X}$, $z \mapsto p(x, z)$ is component-wise monotonic in $z$.

In the case of two binary instruments, under (9) and Assumptions 1’ and B.2 with the direction of component-wise monotonicity normalized to be increasing, we can partition the population into the six well-defined compliance groups presented in Table B.1.
Table B.1: Compliance Groups (Mogstad et al., 2021, Proposition 4)

| Name                      | $D(0, 0)$ | $D(0, 1)$ | $D(1, 0)$ | $D(1, 1)$ |
|---------------------------|-----------|-----------|-----------|-----------|
| Never-takers (nt)         | N         | N         | N         | N         |
| Always-takers (at)        | T         | T         | T         | T         |
| $Z_1$ compliers (1c)      | N         | N         | T         | T         |
| $Z_2$ compliers (2c)      | N         | T         | N         | T         |
| Eager compliers (ec)      | N         | T         | T         | T         |
| Reluctant compliers (rc)  | N         | N         | N         | T         |

Notes: “T” indicates treatment, and “N” indicates non-treatment.

The social welfare criterion can be written as

\[
W(G) = E_P[Y(D(0, Z_2)) + E_P[(Y(1) - Y(0))(D(1, Z_2) - D(0, Z_2))] \cdot 1\{(X, Z_2) \in G\}]
\]

\[
= E_P[Y(D(0, Z_2))] + E_P[E_P[(Y(1) - Y(0))(D(1, 1) - D(0, 1))] | X] Pr_P(Z_2 = 1 | X) \cdot 1\{(X, 1) \in G\}]
\]

\[
+ E_P[E_P[(Y(1) - Y(0))(D(1, 0) - D(0, 0))] | X] Pr_P(Z_2 = 0 | X) \cdot 1\{(X, 0) \in G\}].
\]

Denote by $H \in \mathcal{H} = \{\text{nt, at, 1c, 2c, ec, rc}\}$ the compliance group identity. For each $h \in \mathcal{H}$, let

\[
\pi_h(X) = Pr_P(H = h | X)
\]

be the conditional population share of compliance group $h$ and $\Delta_h(X) = E_P[Y(1) - Y(0)|X, H = h]$ be the compliance-group-specific CATE. Then,

\[
E_P[(Y(1) - Y(0))(D(1, 1) - D(0, 1))] | X] = \pi_{1c}(X)\Delta_{1c}(X) + \pi_{ec}(X)\Delta_{ec}(X),
\]

\[
E_P[(Y(1) - Y(0))(D(1, 0) - D(0, 0))] | X] = \pi_{1c}(X)\Delta_{1c}(X) + \pi_{ec}(X)\Delta_{ec}(X).
\]

Therefore, besides $Z_1$ compliers, the encouragement rule would affect the outcomes of reluctant compliers with $Z_2 = 1$ and the outcomes of eager compliers with $Z_2 = 0$. 
Appendix B.2  Budget Constraints

The causal mechanism that $Z$ affects $Y$ through $D$ will play a role if we impose a resource constraint. Suppose that the proportion of the population that could receive treatment cannot exceed $\kappa \in (0, 1)$. Then we face a constrained optimization problem:

$$\max_{G \in \mathcal{G}} W(G) \text{ s.t. } E_P[1\{p(X, 1) \geq U\} \cdot 1\{X \in G\}] \leq \kappa. \quad (B.1)$$

One simple idea is to enforce random rationing, as considered by Kitagawa and Tetenov (2018b). Assume that, if $G$ violates the resource constraint, then the encouragement is randomly allocated to a fraction $\frac{E_P[1\{p(X, 1) \geq U\} \cdot 1\{X \in G\}]}{E_P[1\{p(X, 1) \geq U\}]}$ of the assigned recipients with $X \in G$ independently of $(X, Y(1), Y(0))$. Under (1), (3), and Assumptions 1 and B.1, we can define the resource-constrained welfare criterion as

$$W^\kappa(G) = W_0 + E_P \left[ \left( \frac{YZ}{q(X)} - \frac{Y(1-Z)}{1-q(X)} \right) \cdot 1\{X \in G\} \right] \min \left\{ 1, \frac{\kappa}{E_P \left[ \frac{DZ}{q(X)} \cdot 1\{X \in G\} \right]} \right\}.$$ 

The optimal encouragement rule under random rationing is then $G^*_{p, \kappa} \in \arg \max_{G \in \mathcal{G}} W^\kappa(G)$.

When $\mathcal{G}$ is unrestricted, (B.1) coincides with the constrained optimization problem studied by Qiu et al. (2021). They note that this corresponds to a fractional knapsack problem and has a closed-form solution, which assigns encouragement to recipients with $X \in G$ in decreasing order of benefit-cost ratio $[YZ \cdot q(X)] - [Y(1-Z) \cdot q(X)] / [DZ \cdot q(X)]$ until the resources are exhausted. The resulting encouragement rule uses the limited resources more efficiently than $G^*_{p, \kappa}$. If restrictions are placed on $\mathcal{G}$, there may be no closed-form solution to (B.1).

Appendix C  Sup-norm Convergence Rate in Expectation for Series Estimators

Let us introduce some notations. Let $\lambda_{\min}(\cdot)$ and $\lambda_{\max}(\cdot)$ denote the smallest and largest eigenvalues, respectively, of a matrix. The exponent $-$ denotes the Moore–Penrose generalized inverse. $\| \cdot \|$ denotes the Euclidean norm when applied to vectors and the matrix spectral norm (i.e., the largest
singular value) when applied to matrices. If \( \{a_n : n \geq 1\} \) and \( \{b_n : n \geq 1\} \) are two sequences of nonnegative numbers, \( a_n \lesssim b_n \) means that there exists a finite positive \( C \) such that \( a_n \leq C b_n \) for all sufficiently large \( n \). Let \( L^\infty(X) \) denote the space of bounded functions under the sup norm, i.e., if \( f : \mathcal{X} \rightarrow \mathbb{R} \) then \( \|f\|_\infty = \sup_{x \in \mathcal{X}} |f(x)| \).

We consider the nonparametric regression model

\[
Y = h_0(X) + \epsilon, \quad E_P[\epsilon|X] = 0,
\]

where \( Y \in \mathbb{R} \) and \( X \in \mathcal{X} \subset \mathbb{R}^d \). Let \( \{(Y_i, X_i)\}_{i=1}^n \) be a random sample. We consider the standard series least-squares estimator of \( h_0 \):

\[
\hat{h}(x) = b^K(x)'(B'B)^{-1}B'Y,
\]

where \( b_{K1}, \ldots, b_{KK} \) are a collection of \( K \) sieve basis functions, and

\[
b^K(x) = (b_{K1}(x), \ldots, b_{KK}(x))', \quad B = (b^K(X_1), \ldots, b^K(X_n))', \quad Y = (Y_1, \ldots, Y_n)'.
\]

We may further impose certain types of restrictions to avoid the curse of dimensionality. For example, we could use a series estimator based on a partially linear additive regression model defined in Carneiro and Lee (2009, Eq. (4.1)).

Define \( \zeta_K = \sup_x \|b^K(x)\| \) and \( \lambda_K = [\lambda_{\min}(E_P[b^K(X_i)b^K(X_i)'])]^{-1/2} \). We introduce some regularity conditions that are standard in the literature on series estimation.

**Assumption C.1** \( \mathcal{X} \) is a Cartesian product of compact connected intervals on which \( X \) has a probability density function that is bounded away from zero.

**Assumption C.2** \( \lambda_{\min}(E_P[b^K(X_i)b^K(X_i)']) > 0 \) for each \( K \).

**Assumption C.3** There exist \( \sigma, \tilde{c} > 0 \) such that \( \sup_x E_P[|\epsilon_i|^q|X_i = x] \leq \frac{q!}{2} \sigma^2 \tilde{c}^{q-2} \) for all \( q \geq 2 \).

**Assumption C.4** \( \inf_{h \in B_K} \|h_0 - h\|_\infty \lesssim K^{-\alpha} \) as \( K \rightarrow \infty \).
Let \( \tilde{b}^K(x) \) denote the orthonormalized vector of basis functions, namely

\[
\tilde{b}^K(x) = E_P[b^K(X_i)b^K(X_i)']^{-1/2}b^K(x)
\]

and let \( \tilde{B} = (\tilde{b}^K(X_1), \ldots, \tilde{b}^K(X_n))' \). The following lemma gives an exponential tail bound for \( \| \tilde{B}'\tilde{B}/n - I_K \| \).

**Lemma C.1.** Under Assumption C.2,

\[
\Pr_{P^n}(\| \tilde{B}'\tilde{B}/n - I_k \| \geq \frac{1}{2}) \leq 2K \exp\left(\frac{-C_1 n}{\zeta^2 K^2 + 1}\right)
\]

for some finite positive constant \( C_1 \).

**Proof.** Follows from Theorem 4.1 of Chen and Christensen (2015) by setting \( \Xi_i = n^{-1}(\tilde{b}^K(X_i)\tilde{b}^K(X_i)' - I_K) \) and noting that \( \max_{1 \leq i \leq n} \| \Xi_i \| \leq n^{-1}(\zeta^2 K^2 + 1) \) and \( \max\{\| \sum_{i=1}^n E_P[\Xi_i \Xi_i'] \|, \| \sum_{i=1}^n E_P[\Xi_i' \Xi_i] \| \} \leq n^{-1}(\zeta^2 K^2 + 1) \). \( \square \)

Let \( B_K = \text{clsp}\{b_{K1}, \ldots, b_{KK}\} \) be a general linear sieve space. Let \( \tilde{h} \) denote the projection of \( h_0 \) onto \( B_K \) under the empirical measure, that is,

\[
\tilde{h}(x) = b^K(x)'(B'B)^{-1}B'H_0 = \tilde{b}^K(x)'(\tilde{B}'\tilde{B})^{-1}\tilde{B}'H_0,
\]

where \( H_0 = (h_0(X_1), \ldots, h_0(X_n))' \). We can bound the sup-norm distance \( \| \tilde{h} - h_0 \|_{\infty} \) using

\[
\| \tilde{h} - h_0 \|_{\infty} \leq \| h_0 - \tilde{h} \|_{\infty} + \| \tilde{h} - \tilde{h} \|_{\infty}
= \text{bias term + variance term}.
\]

The following lemma provides an exponential tail bound on the sup-norm variance term.

**Lemma C.2.** Let \( A_n \) denote the event on which \( \| \tilde{B}'\tilde{B}/n - I_k \| \leq \frac{1}{2} \). Under Assumptions C.1–C.3, for any \( \delta > 0 \),

\[
\Pr_{P^n}(\{ \| \tilde{h} - \tilde{h} \|_{\infty} \geq \delta \} \cap A_n) \lesssim \exp\left\{ -\frac{C_2 n\delta^2}{\zeta^2 K^2 + 1} \right\}
\]

for some finite positive constants \( C_2 \) and \( \zeta^2 \).
for some finite positive constant $C_2$.

**Proof.** By rotational invariance, we have

$$
\hat{h}(x) - \tilde{h}(x) = \tilde{b}^K(x)'(\tilde{B}'\tilde{B}/n)^{-1}\tilde{B}'e/n,
$$

where $e = (\epsilon_1, \ldots, \epsilon_n)'$. Define $G_{i,n}(x) = \tilde{b}^K(x)'(\tilde{B}'\tilde{B}/n)^{-1}\tilde{b}^K(X_i) 1_{A_n}$, where $1_{A_n}$ denotes the indicator function of $A_n$. Then $(\hat{h}(x) - \tilde{h}(x))1_{A_n} = \frac{1}{n} \sum_{i=1}^n G_{i,n}(x)\epsilon_i$. Note that $\|(\tilde{B}'\tilde{B}/n)^{-1}\| \leq 2$ on $A_n$. Then, for $q \geq 2$,

$$
\sum_{i=1}^n E_P^n ([n^{-1}G_{i,n}(x)]^2|\epsilon_i^q|X_i^n]
$$

$$
= \frac{1}{n^2} \sum_{i=1}^n E_P[|\epsilon_i^q|X_i] \tilde{b}^K(x)'(\tilde{B}'\tilde{B}/n)^{-1}\tilde{b}^K(X_i) 1_{A_n}(\tilde{B}'\tilde{B}/n)^{-1}\tilde{b}^K(x)
$$

$$
\leq \sup_x E_P[|\epsilon_i^q|X_i = x] \frac{1}{n^2} \sum_{i=1}^n \tilde{b}^K(x)'(\tilde{B}'\tilde{B}/n)^{-1}\tilde{b}^K(X_i) 1_{A_n}(\tilde{B}'\tilde{B}/n)^{-1}\tilde{b}^K(x)
$$

$$
\leq \frac{q!}{2} \sigma^2 c^q - 2 \frac{1}{n} \tilde{b}^K(x)'(\tilde{B}'\tilde{B}/n)^{-1}(\tilde{B}'\tilde{B}/n)1_{A_n}(\tilde{B}'\tilde{B}/n)^{-1}\tilde{b}^K(x)
$$

$$
\leq \frac{q!}{2} \sigma^2 c^q - 2 \frac{2\zeta^2 K \lambda^2}{n}.
$$

Moreover, for $q \geq 3$,

$$
\sum_{i=1}^n E_P^n [n^{-1}G_{i,n}(x)\epsilon_i^q|X_i^n]
$$

$$
\leq n^{-(q-2)} \sup_x |\tilde{b}^K(x)'(\tilde{B}'\tilde{B}/n)^{-1}1_{A_n}\tilde{b}^K(x)|^{q-2} \sum_{i=1}^n E_P^n [(n^{-1}G_{i,n}(x))^2|\epsilon_i^q|X_i^n]
$$

$$
\leq \left(\frac{2\zeta^2 K \lambda^2}{n}\right) \frac{q-2}{2} \sigma^2 c^q - 2 \frac{2\zeta^2 K \lambda^2}{n}
$$

$$
= \frac{q!}{2} \sigma^2 c^q - 2 \frac{2\zeta^2 K \lambda^2}{n}.
$$

Hence, the conditions of Theorem 2.10 of Boucheron, Lugosi, and Massart (2013) hold with $v = \sigma^2 \frac{2\zeta^2 K \lambda^2}{n}$ and $c = \tilde{c}\frac{2\zeta^2 K \lambda^2}{n}$. By their Corollary 2.11, we have

$$
\Pr_P^n \left( \{ \|\hat{h} - \tilde{h}\|_\infty \geq \delta \} \cap A_n \right)
$$

44
\[= \Pr_{P_n} \left( \left\{ \sup_x \frac{1}{n} \sum_{i=1}^n G_{i,n}(x) \epsilon_i \mid \geq \delta \right\} \cap A_n \right) \]

\[\lesssim \exp \left\{ - \frac{C_2 n \delta^2}{\zeta K^2 \lambda K} \right\} \]

for some finite positive constant \( C_2 \).

Let \( P_{K,n} \) be the empirical projection operator onto \( B_K \), namely

\[P_{K,n} h(x) = b^K(x)'(B'B)^{-1}B'H = \tilde{b}^K(x)'(\tilde{B}'\tilde{B})^{-1}\tilde{B}'H = \tilde{h}(x)\]

where \( H = (h(X_1), \ldots, h(X_n))' \). Let

\[\|P_{K,n}\|_\infty = \sup_{h \in L^\infty(X) : \|h\|_\infty \neq 0} \frac{\|P_{K,n}h\|_\infty}{\|h\|_\infty}\]

denote the sup operator norm of \( P_{K,n} \). The following lemma provides a bound on the sup-norm bias term.

**Lemma C.3.** Under Assumptions C.1 and C.2,

\[\|\tilde{h} - h_0\|_\infty \leq (1 + \|P_{K,n}\|_\infty) \inf_{h \in B_K} \|h_0 - h\|_\infty.\]

and \( \|1_{A_n} P_{K,n}\|_\infty \leq \sqrt{2} \zeta K \lambda K \).

**Proof.** First, for any \( h \in B_K \),

\[\|\tilde{h} - h_0\|_\infty = \|\tilde{h} - h + h - h_0\|_\infty\]

\[= \|P_{K,n}(h_0 - h) + h - h_0\|_\infty\]

\[\leq \|P_{K,n}(h_0 - h)\|_\infty + \|h - h_0\|_\infty\]

\[\leq (1 + \|P_{K,n}\|_\infty) \|h - h_0\|_\infty.\]

Taking the infimum over \( h \in B_K \) yields the first result. Second, take any \( h \in L^\infty(X) \) with \( \|h\|_\infty \neq 0 \).

45
By the Cauchy–Schwarz inequality, we have

\[
|1_{A_n} P_{K,n} h(x)| \leq \|\tilde{b}^K(x)\| \|1_{A_n}(\tilde{B}'\tilde{B}/n)^{-1}\tilde{B}'H/n\|
\]

\[
\leq \zeta_K \lambda_K \|1_{A_n}(\tilde{B}'\tilde{B}/n)^{-1}\tilde{B}'H/n\|
\]

uniformly over \( x \). On \( A_n \), \( \|\tilde{B}'\tilde{B}/n - I_k\| \leq \frac{1}{2} \) and thus \( \lambda_{\text{min}}(\tilde{B}'\tilde{B}/n) \geq \frac{1}{2} \). Then,

\[
\|1_{A_n}(\tilde{B}'\tilde{B}/n)^{-1}\tilde{B}'H/n\|^2 = 1_{A_n}(H'\tilde{B}/n(\tilde{B}'\tilde{B}/n)^{-1}(\tilde{B}'\tilde{B}/n)^{-1}\tilde{B}'H/n
\]

\[
\leq 2(H'\tilde{B}/n(\tilde{B}'\tilde{B}/n)^{-1}\tilde{B}'H/n
\]

\[
\leq 2\|h\|^2 \leq 2\|h\|_\infty^2,
\]

where the second-last inequality follows from \( 1_{A_n} \tilde{B}(\tilde{B}'\tilde{B})^{-1}\tilde{B}' \) being idempotent. It follows that

\[
\|1_{A_n} P_{K,n} h\|_\infty /\|h\|_\infty \leq \sqrt{2}\zeta_K \lambda_K
\]

uniformly in \( h \). Taking the sup over \( h \) yields the second result.

Finally, we conclude the sup-norm convergence rate in expectation for \( \hat{h} \) in the following theorem.

**Theorem C.1.** Under Assumptions C.1–C.4, if \( \lambda_K \lesssim 1 \) and \( \zeta_K \sqrt{\log(K)/n} \rightarrow 0 \), then

\[
E_{P^n}[\|\hat{h} - h_0\|_\infty] = O(K^{-\alpha}) + O\left(\frac{\zeta_K}{\sqrt{n}}\right).
\]

**Proof.** Combining Lemmas C.1–C.3, we obtain

\[
Pr_{P^n}(\|\hat{h} - h_0\|_\infty \geq \delta) \lesssim \exp\left\{-\frac{C_3 n \delta^2}{\zeta_K^2}\right\}
\]

for \( K^{-\alpha} \lesssim \delta \). It follows that

\[
E_{P^n}[\|\hat{h} - h_0\|_\infty] = \int_0^\infty Pr_{P^n}(\|\hat{h} - h_0\|_\infty \geq \delta)d\delta
\]

\[
\leq C_4 K^{-\alpha} + C_5 \int_0^\infty \exp\left\{-\frac{C_3 n \delta^2}{\zeta_K^2}\right\}d\delta
\]

\[
= C_4 K^{-\alpha} + C_6 \frac{\zeta_K}{\sqrt{n}}
\]
When $\mathcal{X}$ is compact and rectangular, $\lambda_K \lesssim 1$, and $\zeta_K \lesssim \sqrt{K}$ for regression splines and $\zeta_K \lesssim K$ for polynomials. If we further assume $h_0(x)$ is continuously differentiable of order $s$ on $\mathcal{X}$, then the approximation rate in Assumption C.4 is $\alpha = s/d_x$. It is straightforward to calculate the convergence rate in Assumption 3.2 as $\psi_n = n^{\frac{1}{2}} \frac{1}{\sqrt{s + 1/d_x}}$ for regression splines and $\psi_n = n^{\frac{1}{2}} \frac{1}{\sqrt{s + 2/d_x}}$ for polynomials.
### Appendix D  Additional Tables and Figures

Table D.1: Sample Averages for the Treatment and Control Groups

|                          | Upper secondary or higher (treatment group) | Less than upper secondary (control group) |
|--------------------------|---------------------------------------------|-------------------------------------------|
|                          | \( N = 841 \)                              | \( N = 1263 \)                            |
| Log hourly wages         | 8.018                                      | 7.209                                     |
| Years of education       | 13.128                                     | 5.585                                     |
| Distance to school (km)  | 1.529                                      | 1.565                                     |
| Distance to health post (km) | 0.331                                      | 0.361                                     |
| Fees per continuing student (1000 Rupiah) | 3.464                                      | 3.992                                     |
| Age                      | 35.668                                     | 36.766                                    |
| Religion                 |                                            |                                           |
| Protestant               | 0.037                                      | 0.008                                     |
| Catholic                 | 0.023                                      | 0.007                                     |
| Other                    | 0.075                                      | 0.039                                     |
| Muslim                   | 0.866                                      | 0.946                                     |
| Father uneducated        |                                            |                                           |
| uneducated               | 0.189                                      | 0.254                                     |
| elementary               | 0.325                                      | 0.251                                     |
| secondary and higher     | 0.275                                      | 0.026                                     |
| missing                  | 0.212                                      | 0.468                                     |
| Mother uneducated        |                                            |                                           |
| uneducated               | 0.182                                      | 0.203                                     |
| elementary               | 0.301                                      | 0.173                                     |
| secondary and higher     | 0.138                                      | 0.011                                     |
| missing                  | 0.379                                      | 0.612                                     |
| Rural household          | 0.483                                      | 0.644                                     |
| North Sumatra            | 0.034                                      | 0.045                                     |
| West Sumatra             | 0.023                                      | 0.023                                     |
| South Sumatra            | 0.069                                      | 0.033                                     |
| Lampung                  | 0.013                                      | 0.028                                     |
| Jakarta                  | 0                                          | 0                                         |
| Central Java             | 0.102                                      | 0.216                                     |
| Yogyakarta               | 0.121                                      | 0.077                                     |
| East Java                | 0.152                                      | 0.201                                     |
| Bali                     | 0.081                                      | 0.035                                     |
| West Nussa Tengara       | 0.084                                      | 0.053                                     |
| South Kalimantan         | 0.040                                      | 0.033                                     |
| South Sulawesi           | 0.040                                      | 0.019                                     |
Figure D.1: Specification Checks for the MTE Estimates

Notes: The (instrument-specific) MTE estimates are evaluated at the mean values of $X$ and $Z_2$. The MTE estimates in the solid line are based on a local quadratic regression with a Gaussian kernel and a bandwidth of 0.27, using a double residual regression procedure described in Appendix B of Heckman, Urzua, and Vytlacil (2006). The bandwidth is picked by Carneiro et al. (2017) using leave-one-out cross-validation. The five other MTE estimates are based on global polynomials in $\hat{p}$ of different orders, with a second-order polynomial (red dashed curve), a third-order polynomial (yellow dotted curve), a fourth-order polynomial (dashes and dots), a fifth-order polynomial (blue dashed curve), and a sixth-order polynomial (green dotted curve).