DISCUSSION PAPER SERIES

DP15092

FILTERED AND UNFILTERED TREATMENT EFFECTS WITH TARGETING INSTRUMENTS

Sokbae Lee and Bernard Salanié

LABOUR ECONOMICS

ISSN 0265-8003

FILTERED AND UNFILTERED TREATMENT EFFECTS WITH TARGETING INSTRUMENTS

Sokbae Lee and Bernard Salanié

Discussion Paper DP15092
Published 23 July 2020
Submitted 20 July 2020

Centre for Economic Policy Research
33 Great Sutton Street, London EC1V 0DX, UK
Tel: +44 (0)20 7183 8801
www.cepr.org

FILTERED AND UNFILTERED TREATMENT EFFECTS WITH TARGETING INSTRUMENTS

Abstract

Multivalued treatments are commonplace in applications. We explore the use of discrete-valued instruments to control for selection bias in this setting. We establish conditions under which counterfactual averages and treatment effects are identified for heterogeneous complier groups. These conditions require a combination of assumptions that restrict both the unobserved heterogeneity in treatment assignment and how the instruments target the treatments. We introduce the concept of filtered treatment, which takes into account limitations in the analyst's information. Finally, we illustrate the usefulness of our framework by applying it to data from the Student Achievement and Retention Project and the Head Start Impact Study.

JEL Classification: N/A

Keywords: identification, selection, multivalued treatments, discrete instruments, unordered monotonicity, factorial design

Sokbae Lee - sl3841@columbia.edu
Columbia University and IFS

Bernard Salanié - bs2237@columbia.edu
Columbia University and CEPR

Acknowledgements
This work is in part supported by the European Research Council (ERC-2014-CoG-646917-ROMIA) and the UK Economic and Social Research Council for research grant (ES/P008909/1) to the CeMMAP.

Filtered and Unfiltered Treatment Effects with Targeting Instruments*

Sokbae Lee† **Bernard Salanié**‡

July 20, 2020

Abstract

Multivalued treatments are commonplace in applications. We explore the use of discrete-valued instruments to control for selection bias in this setting. We establish conditions under which counterfactual averages and treatment effects are identified for heterogeneous complier groups. These conditions require a combination of assumptions that restrict both the unobserved heterogeneity in treatment assignment and how the instruments target the treatments. We introduce the concept of filtered treatment, which takes into account limitations in the analyst's information. Finally, we illustrate the usefulness of our framework by applying it to data from the Student Achievement and Retention Project and the Head Start Impact Study.

KEYWORDS: Identification, selection, multivalued treatments, discrete instruments, unordered monotonicity, factorial design.

*This work is in part supported by the European Research Council (ERC-2014-CoG-646917-ROMIA) and the UK Economic and Social Research Council for research grant (ES/P008909/1) to the CeMMAP.

†Department of Economics, Columbia University and Centre for Microdata Methods and Practice, Institute for Fiscal Studies, sl3841@columbia.edu.

‡Department of Economics, Columbia University, bsalanie@columbia.edu.

Introduction

Much of the literature on the evaluation of treatment effects has concentrated on the paradigmatic "binary/binary" example, in which both treatment and instrument only take two values. Multivalued treatments are common in actual policy implementations, however; and multivalued instruments are just as frequent. Many different programs aim to help train job seekers for instance, and each of them has its own eligibility rules. Tax and benefit regimes distinguish many categories of taxpayers and eligible recipients. The choice of a college and major has many dimensions too, and responds to a variety of financial help programs and other incentives. Randomized experiments in economics resort more and more to factorial designs; they have a long tradition in applied statistics, starting with Fisher in the 1920s[1]. As the training, education choice, and tax-benefit examples illustrate, multivalued treatments are often also subject to selection on unobservables. We explore in this paper the use of discrete-valued instruments in order to control for selection bias when evaluating discrete-valued treatments. We establish conditions under which counterfactual averages and treatment effects are identified for various (sometimes composite) complier groups. These conditions require a combination of assumptions that restrict both the unobserved heterogeneity in treatment assignment and the configuration of the instruments themselves.

Existing work on multivalued treatments under selection on observables includes Imbens (2000), Cattaneo (2010), and Ao, Calonico, and Lee (2019) among others. The literature that uses discrete-valued instruments to evaluate treatment effects under selection on unobservables is more sparse. On the theoretical side, Angrist and Imbens (1995) analyzed two-stage least-squares (TSLS) estimation when the treatment takes a finite number of ordered values. Heckman, Urzua, and Vytlacil (2006, 2008) showed how treatment effects can be identified in discrete choice models for the ordered and unordered cases, respectively. More recently, Heckman and Pinto (2018) focused on unordered treatments and introduced the notion of "unordered monotonicity" under which treatment assignment is formally analogous to an additively separable discrete choice model. Several recent papers have studied the case of binary treatments with multiple instruments, as well as binary instruments with multivalued or continuous treatments. For the former, Mogstad, Torgovitsky, and Walters (2019, 2020) and Goff (2020) analyzed the identifying power of different monotonicity assumptions. For the latter, Torgovitsky (2015), D'Haultfoeuille and Février (2015), Huang, Khalil, and Yildiz (2019), Caetano and Escanciano (2020) and Feng (2020) developed identification results for different models. On the applied side, Kirkeboen, Leuven, and Mogstad (2016) used discrete instruments to obtain TSLS estimates of returns to different fields of study. Kline and

[1] Muralidharan, Romero, and Wüthrich (2019) reviews recent applications of factorial designs.

Walters (2016) revisited the Head Start Impact Study and accounted for the presence of a substitute treatment (alternative preschools in this case). Kamat (2019) developed partial identification results for a selection model with latent choice sets and analyzed the average effects of Head Start preschool access.

Our work is substantially different from any of the aforementioned papers. Rather than focusing on specific cases, we seek a parsimonious framework within which many useful models with multiple treatments and multiple instruments can be analyzed.

Going from a binary treatment to multivalued treatments with discrete-valued instruments raises (at least) two new questions. When treatment can take multiple values, the analyst often only observes a partition of treatment choices. She might for instance only know whether an unemployed individual went through a training program, without knowing exactly which program it was. More generally, the analyst only observes a *filtered treatment* D; underlying it is an *unfiltered treatment* T. Treatment effects are of course harder to identify in the filtered model. The concept of filtering is linked to our earlier work (Lee and Salanié, 2018), which allowed for limited violations of unordered monotonicity and used continuous instruments to identify marginal treatment effects.

Moreover, both filtering and the multiplicity of treatments and instruments may give rise to a bewildering number of cases. In the binary/binary model, the analyst can usually take for granted that switching on the binary instrument makes treatment (weakly) more likely for any observation[2]. With multiple instrument values and multiple treatments, the correspondence is less clear. We start by imposing the unordered monotonicity property of Heckman and Pinto (2018) on the unfiltered treatment model. Under unordered monotonicity, it is natural to speak of an instrument *targeting* an unfiltered treatment by increasing its relative "mean utility". Most of our paper relies on the assumption of *strict targeting*, which obtains when each instrument only promotes the treatments it targets.

To illustrate, consider the effect of various programs T on some outcomes Y. Let each instrument value z stand for a policy regime, under which the access to some programs is made easier or harder than in a control group. Under unordered monotonicity, this translates into a profile of relative mean utilities of any treatment t under the policy regimes $z \in \mathcal{Z}$. We say that an instrument value z targets a treatment t when it maximizes its relative mean utility. Suppose that each policy regime consists of values of subsidies for a subset of the programs, and that these subsidies enter mean utilities additively. Then a policy regime z targets a treatment t if it has the highest subsidy for this program among all policy regimes. Strict targeting requires that all policy regimes z' that do not target t have the same (lower)

[2] This is satisfied under the LATE-monotonicity assumption (e.g., Imbens and Angrist, 1994; Vytlacil, 2002).

subsidy for t. It is easy to translate this property in the other examples cited at the beginning of the introduction.

With complete treatment data (the unfiltered treatment), combining unordered monotonicity and strict targeting allows us to point-identify the size of some complier groups and the corresponding treatment effects, and to partially identify others. When the data on treatments is filtered, unordered monontonicity may not carry over to the filtered treatment D (an observation already in Lee and Salanié (2018)); and strict targeting generally does not. Nevertheless, they confer enough underlying structure to the mapping from instruments to filtered treatments that we can still identify various parameters of interest.

We give numerous examples throughout the paper. We also illustrate the usefulness of our framework by applying it to data from the Student Achievement and Retention (STAR) Project (Angrist, Lang, and Oreopoulos, 2009) and to Kline and Walters's (2016) analysis of the Head Start Impact Study. We find that the large intent-to-treat (ITT) effect of the STAR for female college students results from the aggregation of two very different treatment effects; this highlights the value of unbundling the heterogeneous compliers. We also confirm the importance of taking into consideration alternative preschools when evaluating Head Start; unlike Kline and Walters (2016), we do not rely on parametric selection models.

The remainder of the paper is organized as follows. Section 1 defines our framework and introduces filtered and unfiltered treatments. In Section 2, we study identification in the unfiltered treatment model. We define the concepts of targeting, one-to-one targeting, and strict targeting and their implications for the identification of the probabilities and the treatment effects of various complier groups. Section 3 turns to filtered models. We derive identification results in several leading classes of applications. Finally, we present estimation results for the two aforementioned empirical studies in Section 4. The Appendices contain the proofs of all propositions and lemmata, along with some additional material.

1 Filtered and Unfiltered Treatment

We focus throughout on a treatment that takes discrete values, which we label $d \in \mathcal{D}$. For simplicity, we will call $D = d$ "treatment d" These values are unordered: e.g. $d = 2$, when available, is not "more treatment" than $d = 1$. In most of our examples, there is a well-defined control group, which is denoted by $d = 0$. We assume that discrete-valued instruments $Z_i \in \mathcal{Z}$ are available. We condition on all other exogenous covariates X_i throughout, and we omit them from the notation. We will use the standard counterfactual notation: $D_i(z)$ and $Y_i(d, z)$ denote respectively potential treatments and outcomes.

The validity of the instruments requires the usual exclusion restrictions:

Assumption 1 (Valid Instruments). *(i)* $Y_i(d, z) = Y_i(d)$ *for all* (d, z) *in* $\mathcal{D} \times \mathcal{Z}$.

(ii) $Y_i(d)$ *and* $D_i(z)$ *are independent of* Z_i *for all* (d, z) *in* $\mathcal{D} \times \mathcal{Z}$.

Under Assumption 1, we define $D_i := D_i(Z_i)$ and $Y_i := Y_i(D_i)$. Throughout the paper, we assume that we observe (Y_i, D_i, Z_i) for each i. In addition, the instruments must be relevant. In the usual binary instrument/binary treatment case (hereafter "binary/binary"), this translates into a requirement that the propensity score vary with the instruments. In our more general setting, we impose:

Assumption 2 (Relevant Instruments). *Let* $\mathbf{Z}_i$ *denote a column vector whose elements are 1 and the variables* $\mathbb{1}(Z_i = z)$ *for* $z \in \mathcal{Z}$, *and* $\mathbf{D}_i$ *denote a column vector whose elements are 1 and the variables* $\mathbb{1}(D_i = d)$ *for* $d \in \mathcal{D}$. *Then* $\mathbb{E}[\mathbf{Z}_i \mathbf{D}_i^\top]$ *has full rank.*

1.1 Unobserved Monotonicity

We now move beyond these standard assumptions. First, we need an assumption that restricts the heterogeneity in the counterfactual mappings D_i. In the binary/binary model, this is most often done by imposing LATE-monotonicity.

Assumption 3 (LATE-monotonicity in the binary/binary model). *(i) or (ii) must hold:*

(i) for each observation i, $D_i(1) \geqslant D_i(0)$;

(ii) for each observation i, $D_i(0) \geqslant D_i(1)$.

With more than two treatment values and/or more than two instrument values, there are many ways to restrict the heterogeneity in treatment assignment. Since treatments are not ordered in any meaningful way, we cannot apply the results in Angrist and Imbens (1995) for instance. Mogstad, Torgovitsky, and Walters (2019, 2020) state several versions of monotonicity for a binary treatment model with $|\mathcal{Z}| > 2$. They propose an assumption PM (partial monotonicity) which applies binary LATE-monotonicity component by component. This requires that the instruments be interpretable as vectors, which is not necessarily the case here.

Heckman and Pinto (2018) took another path; they defined an unordered monotonicity property that is motivated by an analogy to revealed preference theory. This can be stated as follows:

Assumption 4 (Unordered Monotonicity at (z, z')). *For any treatment value* $d \in \mathcal{D}$, *(i) or (ii) must hold:*

(i) if $D_i(z) = d$, then $D_i(z') = d$;

(ii) if $D_i(z') = d$, then $D_i(z) = d$.

The easiest way to understand Assumption 4 is to think of treatment assignment as generated by a discrete choice problem. If observation i "chose" treatment value d under z, then a change in instrument value that increases the mean utility of treatment d at least as much as the mean utilities of other treatment values should lead i to still choose d. This is more than illustrative: Heckman and Pinto (2018) show that the treatment assignment models that satisfy unordered monotonicity for each pair of instrument values in a set $\mathcal{Z}$ can be represented by a discrete choice problem with additively separable errors, that is

$$D_i(z) = \arg\max_{d \in \mathcal{D}} (U_z(d) + u_{id})$$

for random vectors $(u_{id})_{d \in \mathcal{D}}$ that are distributed independently of Z_i.[3] Let AS-DCM denote this class of models. Clearly, Asssumption 4 is more restrictive if the set of instrument values $\mathcal{Z}$ is richer. In Example 1 below, we would only want to invoke Assumption 4 on some policy changes.

Example 1. Unemployed individuals can be assigned either to a control group or to three different training programs, with treatment values $d = 1, 2, 3$. Consider three alternative policy changes ($z \to z'$), all of which make more individuals eligible for treatment 1. Policy change A at the same time restricts the eligibility criteria for both treatments 2 and 3 in unspecified ways. Policy change B leaves eligibility criteria unchanged for treatments 2 and 3, and policy change C restricts them for treatment 2 only. Assumption 4 would require that all observations that have treatment 1 under z must also have treatment 1 under z', which seems natural in this context. It would also prevent "two-way moves" between treatment 2 and treatment 3, which seems restrictive. Assumption 4 would be more credible with policy changes B and C. $\square$

1.2 Filtering Treatment Models

Heckman and Pinto (2018) showed how unordered monotonicity could be applied to identify some treatment effects (or weighted averages of treatment effects). In Lee and Salanié (2018), we considered a more general family of models of treatment assignment. We allowed for

[3] As a special case, unordered monotonicity includes ordered treatments in which $D_i(z) = \arg\max_{d \in \mathcal{D}} (U_z(d) + \sigma(d) u_i)$ for some increasing positive function σ.

treatment assignment to be determined by any logical combination of a finite set of threshold-crossing rules of the form $u_{ij} \leqslant Q_j(z)$. All AS-DCM models clearly belong to this class, as $D_i(z)$ is characterized by

$$u_{id} - u_{i,D_i(z)} \leqslant U_z(D_i(z)) - U_z(d) \text{ for all } d \in \mathcal{D}.$$

Our results showed that this class of models can be generated by

1. taking an AS-DCM model of assignment to treatment values $T \in \mathcal{T}$,
2. generating the observed treatment $D \in \mathcal{D}$ from a partition of the set $\mathcal{T}$.

This defines the class of models of assignment to treatment that we analyze in this paper. We call such models "filtered treatment", and we will refer to the (imperfectly observed) model of treatment in 1 above as the "unfiltered treatment". To pursue the discrete choice analogy: in the unfiltered model, each observation chooses a treatment within $\mathcal{D}$ and the analyst observes this choice. In a filtered model, choices are aggregated into groups; the analyst only observes which group the treatment chosen belongs to. The aggregation occurs via a filtering map from $\mathcal{T}$ to $\mathcal{D}$.

Definition 1 (Filtered Treatment)**.** The treatment assignment model is determined by:

1. a finite set $\mathcal{T}$;
2. a partition of $\mathcal{T}$ which we call $\mathcal{D}$; or equivalently, a surjective *filtering map* $M : \mathcal{T} \to \mathcal{D}$;
3. a finite set of instrument values $\mathcal{Z}$; and
4. an AS-DCM model of unfiltered treatment:

$$T_i(z) = \arg\max_{t \in \mathcal{T}} (U_z(t) + u_{it}),$$

where the vector $(u_{it})_{t \in \mathcal{T}}$ is distributed independently of Z_i and has full support on $\mathbb{R}^{|\mathcal{T}|}$.

Our paper focuses on such models.

Assumption 5 (Filtered Treatment)**.** *Treatment is assigned according to Definition 1. We call the model that generates T_i the* underlying unfiltered treatment model.

We will assume throughout, implicitly, that the set of instrument values $\mathcal{Z}$ has been restricted to a subset where unordered monotonicity is a reasonable assumption. We will also take some liberties with language by speaking of individuals "choosing" their "preferred" unfiltered treatments and of "mean utilities" $U_z(t)$. These are only meant to simplify the exposition and do not imply that the individual actually chooses her treatment.

As an example, consider the following double hurdle model; it has $|\mathcal{D}| = 2$, and an underlying unfiltered treatment model with $|\mathcal{T}| = 3$.

Example 2 (Double Hurdle Treatment)**.** The unfiltered treatment has $\mathcal{T} = \{0, 1, 2\}$ and

$$T_i(z) = \arg\max_{t=0,1,2} \left(U_z(t) + u_{it}\right),$$

where the vector (u_{i0}, u_{i1}, u_{i2}) is distributed independently of Z_i and has full support on $\mathbb{R}^3$.

Suppose that the filtered treatment is generated by $D = \mathbb{1}(T = 0)$, which corresponds to the filtering map $M(0) = 1, M(1) = M(2) = 0$; that is,

$$\begin{cases} D_i(z) = 1 & \text{iff } \max(U_z(1) + u_{i1}, U_z(2) + u_{i2}) < U_z(0) + u_{i0} \\ D_i(z) = 0 & \text{otherwise.} \end{cases} \tag{1.1}$$

Here our logical combination of threshold rules is simply an "AND" over the two inequalities: $u_{i0} - u_{i1} > U_z(1) - U_z(0)$ and $u_{i0} - u_{i2} > U_z(2) - U_z(0)$. $\square$

Lee and Salanié (2018) gave a set of assumptions under which the marginal treatment effect can be identified in a filtered treatment model, provided that enough continuous instruments are available. In Example 2, we would need two continuous instruments, and some additional restrictions. The current paper is exploring identification with discrete-valued instruments. In these settings, the combination of Assumptions 1, 2, and 5 is far from sufficient to identify interesting treatment effects in filtered and unfiltered treatment models in general. In order to better understand what is needed, we now resort to the notion of *response-groups* of observations, whose members share the same mapping from instruments z to unfiltered treatments t. We first state a general definition[4].

Definition 2 (Response-vectors and -groups)**.** Let $\tilde{t}$ be an element of $\mathcal{T}^{\mathcal{Z}}$ and $\tilde{t}(z) \in \mathcal{T}$ denote its component for instrument value $z \in \mathcal{Z}$.

- Observation i has (elemental) *response-vector* $R_{\tilde{t}}$ if and only if for all $z \in \mathcal{Z}$, $T_i(z) = \tilde{t}(z)$. The set $C_{\tilde{t}}$ denotes the set of observations with response-vector $R_{\tilde{t}}$ and we call it a *response-group*.

[4]This is analogous to the definitions in Heckman and Pinto (2018).

- We extend the definition in the natural way to incompletely specified mappings, where $\tilde{t}$ is a correspondence from $\mathcal{Z}$ to $\mathcal{T}$. We call the corresponding response-vectors and response-groups *composite*.

2 Identifying the Unfiltered Treatment Model

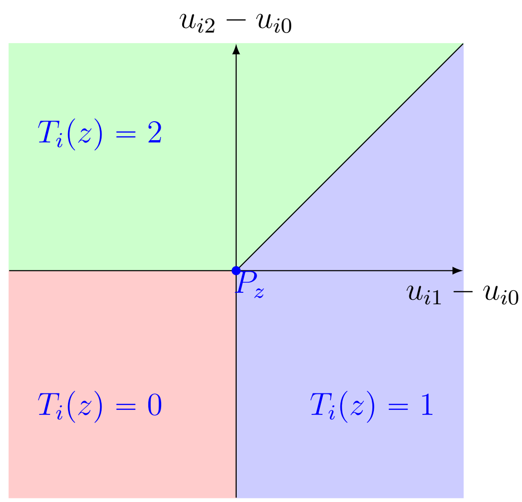
We start by introducing additional assumptions on the underlying unfiltered treatment model. We will illustrate these assumptions in simple graphs; our leading example is the "ternary/ternary" case when $|\mathcal{T}| = |\mathcal{Z}| = 3$.

Example 3 (Ternary/ternary unfiltered model)**.** Assume that $\mathcal{Z} = \{0, 1, 2\}$ and $\mathcal{T} = \{0, 1, 2\}$. In the $(u_{i1} - u_{i0}, u_{i2} - u_{i0})$ plane, the points of coordinates $P_z = (U_z(0) - U_z(1), U_z(0) - U_z(2))$ for $z = 0, 1, 2$ are important; for a given z,

- $T_i(z) = 0$ to the south-west of P_z;
- $T_i(z) = 1$ to the right of P_z and below the diagonal that goes through it;
- $T_i(z) = 2$ above P_z and above the diagonal that goes through it.

This is shown in Figure 1 for a given z, where the origin is in P_z.

Figure 1: Unfiltered treatment assignment in the ternary/ternary model for given z



2.1 Targeted Treatments

"Targeting" will be the common thread in our analysis. Just as in general economic discussions a policy measure may target a particular outcome, we will speak of instruments (in the econometric sense) targeting the assignment to a particular treatment.

Under unordered monotonicity (Assumption 4), assignment to treatment is governed by the differences in mean utilities $(U_z(t) - U_z(\tau))$ and by the differences in unobservables $u_{it} - u_{i\tau}$. Only the former depend on the instrument. Intuitively, an instrument z targets a treatment t if it makes the difference $(U_z(t) - U_z(\tau))$ as large as possible for given τ. Instead of requiring this for any τ, we will choose a *reference treatment* $t_0 \in \mathcal{T}$ and require that z maximize $(U_z(t) - U_z(t_0))$ for this particular t_0. In many applications, the control group is a natural choice for a reference treatment. Since the control group is usually denoted $t = 0$, we will extend the notation and denote the reference treatment $t_0 = 0$.

The following definition makes this more precise.

Definition 3 (Targeted Treatments and Targeting Instruments)**.** Let $t = 0$ denote a reference treatment value. For any $z \in \mathcal{Z}$ and $t \in \mathcal{T}$, we denote

$$\Delta_z(t) \equiv U_z(t) - U_z(0)$$

the *relative mean utility* of treatment t given instrument z.

Let $\bar{\Delta}_t$ be the maximum value of $\Delta_z(t)$ over $z \in \mathcal{Z}$, and $\bar{Z}(t)$ the set of maximizers $z \in \mathcal{Z}$. If $\bar{Z}(t)$ is not all of $\mathcal{Z}$, then for any $z \in \bar{Z}(t)$ we will say that instrument value z *targets* treatment value t; and we write $t \in \bar{T}(z)$. We denote $\mathcal{T}^*$ the set of targeted treatments and $\mathcal{Z}^* = \bigcup_{t \in \mathcal{T}^*} \bar{Z}(t)$ the set of targeting instruments.

Definition 3 calls for several remarks. First, by construction $\Delta_z(0) \equiv 0$ and $\bar{Z}(0) = \mathcal{Z}$. Therefore $t = 0$ is not in $\mathcal{T}^*$. In many of our examples, $\mathcal{T}^* = \mathcal{T} \setminus \{0\}$; the set $\mathcal{T}^*$ may exclude other treatment values, however.

If a treatment value t is not targeted, by definition the function $z \to \Delta_z(t)$ is constant over $z \in \mathcal{Z}$, with value $\bar{\Delta}_t$. While treatment values in $\mathcal{T} \setminus \mathcal{T}^*$ have mean utilities that do not respond to changes in the instruments, these mean utilities may and in general will differ across treatments. The probability that an individual observation takes a treatment $t \in \mathcal{T} \setminus \mathcal{T}^*$ also generally depends on the value of the instrument.

More importantly, the utilities $U_z(t)$ and therefore the targeting maps $\bar{Z}$ and $\bar{T}$ are not observable; any assumption on targeting instruments and targeted treatments must be a priori and will be context-dependent. As we will see, these prior assumptions sometimes have consequences that can be tested.

Let us return to the illustration that we used in the introduction. A policy regime z consists of a set of (possibly zero or negative) subsidies $S_z(t)$ for treatments $t \in \mathcal{T}$. If there is a no-subsidy regime $z = 0$ with $S_0(t) = 0$ for all t, it seems natural to write the mean utility as $U_z(t) = U_0(t) + S_z(t)$. Then relative mean utilities are $\Delta_z(t) = \Delta_0(t) + S_z(t)$ and for any treatment t, the set $\bar{Z}(t)$ consists of the instrument values z that subsidize t most heavily. As this illustration suggests, the sets $\bar{Z}(t)$ may not be singletons, and they may well intersect. We will show this on several examples.

2.1.1 Targeting Examples

Example 4 is an instance of factorial design in which each non-zero treatment value is targeted by two instrument values, and one instrument value targets several treatments.

Example 4 (2 × 2 factorial design)**.** Let $\mathcal{Z} = \{0 \times 0, 0 \times 1, 1 \times 0, 1 \times 1\}$, where the two digits indicate the values of two binary instruments z_1 and z_2. Suppose that $\mathcal{T} = \{0, 1, 2\}$, where $z_1 = 1$ is intended to promote treatment 1 and $z_2 = 1$ is intended to promote treatment 2:

$$\begin{aligned}\Delta_{1\times 0}(1) &= \bar{\Delta}_1 > \max(\Delta_{0\times 0}(1), \Delta_{0\times 1}(1))\\ \Delta_{0\times 1}(2) &= \bar{\Delta}_2 > \max(\Delta_{0\times 0}(2), \Delta_{1\times 0}(2)).\end{aligned}$$

Depending on the context, it may be reasonable to assume that $\Delta_{1\times 1}(1) = \Delta_{1\times 0}(1)$ and $\Delta_{1\times 1}(2) = \Delta_{0\times 1}(2)$: turning on the two instruments increases the appeal of $t = 1$ (resp. $t = 2$) just as much as if only z_1 (resp. z_2) had been turned on. This would be quite natural if $z_1 = 1$ subsidizes treatment 1 and $z_2 = 1$ subsidizes treatment 2: then 1×1 is the policy regime that subsidizes both. Then we have $\bar{Z}(1) = \{1 \times 0, 1 \times 1\}$ and $\bar{Z}(2) = \{0 \times 1, 1 \times 1\}$; instrument $z = 1 \times 1$ targets both $t = 1$ and $t = 2$, so that $\bar{T}(1 \times 1) = \{1, 2\}$. □

Example 5 (Two Instruments Target the Same Treatment)**.** Let us now modify Example 4 slightly: the instrument can only take values $0 \times 0, 1 \times 0$, and 1×1. Then $z = 1 \times 0$ and $z = 1 \times 1$ both target treatment $t = 1$: $\bar{Z}(1) = \{1 \times 0, 1 \times 1\}$. □

Example 6 (An Instrument Targets Two Treatments)**.** In this example, $\mathcal{Z} = \{0, 1\}$ and $\mathcal{T} = \{0, 1, 2\}$. A fraction of individuals in the sample receives a subsidy $z = 1$ that can be used for both treatments $t = 1$ and $t = 2$; under $z = 0$, no treatment is subsidized. We would expect that $\Delta_1(1) > \Delta_0(1)$ and $\Delta_1(2) > \Delta_0(2)$, so that $\bar{Z}(1) = \bar{Z}(2) = \{1\}$; then we have $\mathcal{T}^* = \{1, 2\}$ and $\mathcal{Z}^* = \{1\}$. □

2.1.2 One-to-one Targeting

Sometimes we will impose the much stronger Assumption 6, or only one of its two parts. The first part says that a targeted treatment can only have one targeting instrument; the second part stipulates that a targeting instrument may target only one treatment. Example 4 violates both parts of Assumption 6. Example 5 violates its first part only, and Example 6 only violates its second part.

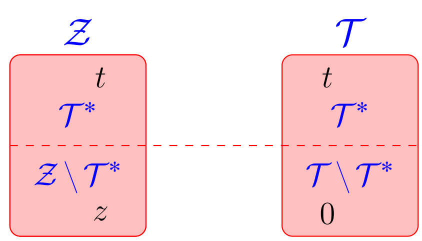
Assumption 6 (One-to-one Targeting). *(i) For any $t \in \mathcal{T}^*$, the set $\bar{Z}(t)$ is a singleton $\{\bar{z}(t)\}$.*

(ii) For any $z \in \mathcal{Z}^*$, $\bar{T}(z)$ is a singleton $\{\bar{t}(z)\}$.

Note that if both parts of Assumption 6 hold, we can identify $\mathcal{Z}^*$ and $\mathcal{T}^*$, and an instrument $z \in \mathcal{Z}^*$ to the treatment it targets.

Definition 4 (Labeling Instruments). Let both parts of Assumption 6 hold. To any $t \in \mathcal{T}^*$ we associate the instrument value $z = \bar{z}(t)$ that targets t, and we denote it by $z = t$. This allows us to define the partition $\mathcal{Z} = (\mathcal{Z} \backslash \mathcal{T}^*) \bigcup \mathcal{T}^*$, which is illustrated in Figure 2.

Figure 2: One-to-one Targeting



Example 7 (Treatment Subsidies). Let $\mathcal{T} = \{0\} \bigcup \mathcal{T}^*$ with $\mathcal{T}^* = \{1, \ldots, |\mathcal{T}| - 1\}$, and $\mathcal{Z} = \mathcal{T}$. Each $z > 0$ instrument can be interpreted as a subsidy that targets the corresponding treatment $t = z$ in the sense that for each $t > 0$, $S_t(t) > S_z(t)$ for all $z \neq t$. $\square$.

Example 8 (Binary Instrument). Let $\mathcal{T} = \{0\} \bigcup \mathcal{T}^*$ with $\mathcal{T}^* = \{1, \ldots, |\mathcal{T}| - 1\}$, and $\mathcal{Z} = \{0, 1\}$. An observation with $z = 1$ receives a subsidy for the treatment $t = 1$, so that $S_1(1) > 0$. Other treatment values are not subsidized: $S_0(t) = S_1(t) = 0$ for all $t \neq 1$. Then $\Delta_1(t) = \Delta_0(t)$ for all $t \neq 1$, so that $\mathcal{Z}^* = \{1\}$. $\square$

Example 9 (No Control). Let $\mathcal{Z} = \mathcal{T}^*$, so that there is one fewer instrument value than treatment values. The simplest example in this class is the *ternary/binary* model, with $\mathcal{T} = \{0, 1, 2\}$ and $\mathcal{Z} = \mathcal{T}^* = \{1, 2\}$. There are only two classes of observations: those with $z = 1$ were offered a subsidy for $t = 1$, and those with $z = 2$ were offered a subsidy for $t = 2$.

2.2 Strict Targeting

Assumption 4, conjoined with Assumption 6, imposes some useful restrictions on response groups.

Proposition 1 (Unfiltered response groups (1))**.** *Under Assumptions 4 and 6, for any $t \in \mathcal{T}^*$:*

- *if $T_i(t) = 0$, then $T_i(z) \neq t$ for all $z \in \mathcal{Z}$;*
- *as a consequence, all response-groups $C_{\tilde{t}}$ with $\tilde{t}(t) = 0$ and $\tilde{t}(z) = t$ for some $z \neq t$ are empty.*

Example 3 (continued) Return to the ternary/ternary model and assume that the targeted set of treatments $\mathcal{T}^* = \{1, 2\}$ and that Assumptions 4 and 6 hold. This imposes

$$\Delta_1(1) > \max(\Delta_2(1), \Delta_0(1)) \text{ and } \Delta_2(2) > \max(\Delta_1(2), \Delta_0(2)).$$

A possible interpretation is that policy regime $z = 1$ (resp. $z = 2$) subsidizes treatment $t = 1$ (resp. $t = 2$) more that policy regimes $z = 0$ and $z = 2$ (resp. $z = 1$) do.

Since P_z has coordinates $(-\Delta_z(1), -\Delta_z(2))$,

- P_1 must lie to the left of P_0 and P_2,
- P_2 must lie below P_0 and P_1.

This is easily rephrased in terms of the response-vectors of definition 2. First note that in the ternary/ternary case, there are $3^3 = 27$ response-vectors, R_{000} to R_{222}, with corresponding response-groups C_{000} to C_{222}. Groups C_{ddd} are "always-takers"[5] of treatment value d. All other groups are "compliers" of some kind, in that their treatment changes under some changes in the instrument. We will also pay special attention to some non-elemental groups. For instance, R_{0*2} will denote the group who is assigned treatment 0 under $z = 0$ and treatment 2 under $z = 2$, and any treatment under $z = 1$. That is,

$$C_{0*2} = C_{002} \bigcup C_{012} \bigcup C_{022}.$$

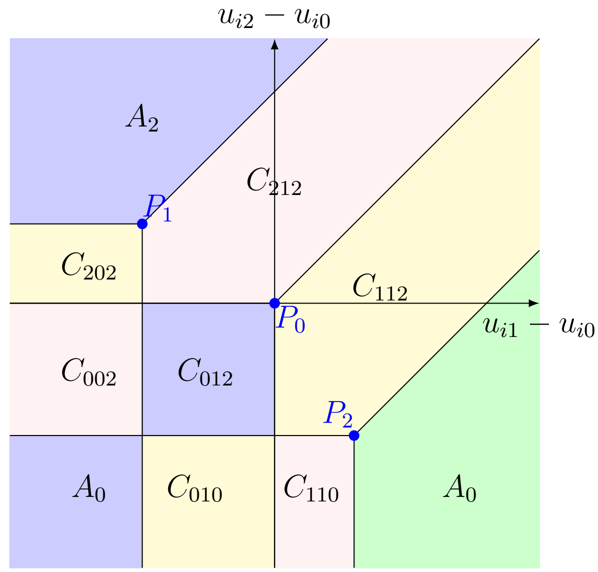
Assumption 4 asserts the emptiness of four composite groups out of the 27 possible: C_{10*}, C_{*01}, C_{*20}, and C_{2*0} by Proposition 1. They correspond to 10 elemental groups.[6]

[5]Observations in group C_{000} are usually called the "never-takers". We prefer not to break the symmetry in our notation. We hope this will not cause confusion.

[6]Specifically, they are: $C_{100}, C_{101}, C_{102}, C_{001}, C_{201}, C_{020}, C_{120}, C_{220}, C_{200}$, and C_{210}.

This still leaves us with 17 elemental groups, and potentially complex assignment patterns. Consider for instance Figure 3. It shows one possible configuration for the ternary/ternary model; the positions for P_0, P_1 and P_2 are consistent with Assumptions 4 and 6.

Figure 3: Unordered monotonic ternary/ternary models: an example



The number of distinct response-groups (ten) and the contorted shape of the C_{212} and C_{112} groups in Figure 3 point to the difficulties we face in identifying response-groups without further assumptions. Moreover, this is only one possible configuration: other cases exist, which would bring up other response-groups.

Figure 3 also suggests that if we could make sure that P_1 is directly to the left of P_0, the shape of C_{212} would become nicer—and group C_{202} would be empty. Bringing P_2 directly under P_0 would have a similar effect. But these are assumptions on the dependence of the $U_z(d)$ on instruments. The first one imposes $\Delta_1(2) = \Delta_0(2)$ and the second one imposes $\Delta_2(1) = \Delta_0(1)$. To put it differently, we are now requiring that instrument $z = t$, which maximizes $\Delta_z(t) = U_z(t) - U_z(0)$, should not shift assignment between the other values of the treatment. This can be interpreted as policy regime $z = 1$ (resp. $z = 2$) subsidizing treatment $t = 1$ (resp. $z = 2$) only. $\square$

The following assumption is a direct extension of the discussion above to our general discrete model.

Assumption 7 (Strict Targeting)**.** *Take any targeted treatment value $t \in \mathcal{T}^*$. Then the function $z \in \mathcal{Z} \to \Delta_z(t)$ takes the same value for all $z \notin \bar{Z}(t)$. We denote this common value by $\underline{\Delta}_t$, and we will say of the instrument values $z \in \bar{Z}(t)$ that they* strictly target *t.*

Under Assumption 7, turning on instrument $z \in \bar{Z}(t)$ promotes treatment t without affecting the relative mean utilities of other treatment values. This explains our use of the term "strict targeting". To return to the analogy with a discrete choice model, an instrument in $\bar{Z}(t)$ plays the role of a price discount on good t in a model of demand for goods whose mean utilities only depend on their own prices. In the language of program subsidies, all $z \in \bar{Z}(t)$ subsidize t at the same high rate, and all other instrument values offer the same, lower subsidy (which could be zero or negative).

Note that while we only state the assumption for $t \in \mathcal{T}^*$, it holds by definition for all $t \in \mathcal{T} \backslash \mathcal{T}^*$. Since $\bar{Z}(t) = \mathcal{Z}$ for these treatment values, $\underline{\Delta}_t = \bar{\Delta}_t$ is the common value of $\Delta_z(t)$ over all of $\mathcal{Z}$.

Moreover, Assumption 7 only bites for a given $t \in \mathcal{T}^*$ if $\mathcal{Z} \backslash \bar{Z}(t)$ has at least two values. Since $\bar{Z}(t)$ is never empty, this shows that Assumption 7 automatically holds if $|\mathcal{Z}| = 2$ (one binary instrument). Therefore it is satisfied in our Examples 6 and 9. Strict targeting also holds in our Example 7, and in our factorial design of Example 4 if $\Delta_{0\times1}(1) = \Delta_{0\times0}(1)$ and $\Delta_{1\times0}(2) = \Delta_{0\times0}(2)$ (so that $z = 0 \times 1$ does not subsidize $t = 1$ and $z = 1 \times 0$ does not subsidize $t = 2$).

Note that one-to-one targeting and strict targeting are logically independent assumptions: neither one implies the other. As we just saw, the factorial design in Example 4 may exhibit strict targeting; but it never satisfies one-to-one targeting. The converse may also hold, for instance if $z = 1$ and $z = 2$ both subsidize $t = 1$, and $z = 2$ is a more generous subsidy. Then we would expect $\bar{Z}(1) = \{2\}$ yet $\Delta_0(1) < \Delta_1(1)$.

Example 10 (Tuition Subsidies)**.** To shed light on Assumption 7, consider two types of policies aimed at making education more affordable. Our first policy consists of field-specific tuition subsidies. Each individual i is offered randomly a choice of $m_i \geqslant 0$ vouchers for a subset Z_i of fields. If $m_i \geqslant 1$, the individual may choose to use a voucher to study in a field in Z_i, to study in another field, or not pursue education. Let T_i denote this choice, with $T_i = 0$ for no education. For fixed $t \neq 0$, the value of $\Delta_z(t)$ is highest when $t \in z$ as a voucher can be used. Therefore $\bar{Z}(t)$ is the set of menus of vouchers that include field t; and $\mathcal{T}^*$ is the set of fields for which a voucher is sometimes, but not always offered. Whether $\bar{Z}(t)$ is a single menu or not, all other menus of vouchers yield the same $\Delta_z(t)$: the field t is strictly targeted[7].

Another possible policy consists in subsidizing tuition for every year of study in the hope of increasing the number of years of education. Now z is a subsidy rate, and t the number

[7]Note that iff $m_i \leqslant 1$ for each individual, targeting is one-to-one. If not, either part of Assumption 6 could fail. If the fields are French (F), Greek (G), Korean (K), and Latin (L), a set of two menus $z_1 = \{F, K\}$ and $z_2 = \{G, L\}$ fails the second part of Assumption 6; a set $z_1 = \{F, K\}$ and $z_2 = \{G, K\}$ fails both parts.

of years of education. Since a higher subsidy rate reduces the cost of education, for any t the function $\Delta_z(t)$ achieves its maximum $\bar{\Delta}_t$ for the highest subsidy $\bar{z}$ on offer: for each t, $\bar{Z}(t) = \{\bar{z}\}$ and Assumption 7 fails. More importantly, if $|\mathcal{Z}| > 2$ then for any $t > 0$, the value of $\Delta_z(t)$ increases with $z \neq \bar{z}$. Strict targeting would clearly not be an appropriate assumption in this setting. $\square$

Extending our geometric illustration of Example 3, let P_z be the point in $\mathbb{R}^{|\mathcal{T}^*|}$ with coordinates $(-\Delta_z(t))_{t\in\mathcal{T}^*}$. Under Assumption 7, the point P_z has its t coordinate equal to $-\overline{\Delta}_t$ on any axis t which it targets ($t \in \bar{T}(z)$), and $-\underline{\Delta}_t$ on any other axis. Since $-\underline{\Delta}_t > -\overline{\Delta}_t$, two points P_z and $P_{z'}$ have the same coordinate on any axis $t \notin \bar{T}(z) \bigcup \bar{T}(z')$; and P_z is below $P_{z'}$ on axis t if $t \in \bar{T}(z)\backslash\bar{T}(z')$.

Now suppose that in addition to $\mathcal{Z}^*$, the set of instruments contains at least two values z_0 and z_1. Since neither targets any treatment, under Assumption 7 $\Delta_{z_1}(t) = \Delta_{z_0}(t) = 0$ for any $t \in \mathcal{T}^*$. Moreover, $\Delta_z(t)$ equals $\underline{\Delta}_t$ for all $z \in \mathcal{Z}$ if $t \notin \mathcal{T}^*$. This implies that the counterfactual treatments $T_i(z_0)$ and $T_i(z_1)$ must be equal for any observation i. In that sense, z_1 is superfluous and we can aggregate it with z_0 in a category that we will call $z = 0$. By the previous paragraph, if $z \neq 0$ then the point P_0 is above the point P_z on any axis $t \in \bar{T}(z)$.

We summarize this in Lemma 1.

Lemma 1 (Some consequences of strict targeting)**.** *Under Assumption 4 and Assumption 7,*

(i) The coordinates of two points P_z and $P_{z'}$ in $\mathbb{R}^{|\mathcal{T}^*|}$ coincide on any axis t' that is not in the symmetric difference $\bar{T}(z) \triangle \bar{T}(z')$.

(ii) If $z \in \bar{Z}(t)$ and $z' \in \bar{Z}(t)$, the point $P_{z'}$ is above the point P_z on the axis t.

(iii) The set of instrument values $\mathcal{Z}$ is either $\mathcal{Z}^*$, or the union of $\mathcal{Z}^*$ and of a single instrument value that we denote $z_0 \notin \mathcal{Z}^*$. In the latter case, for any $z \neq z_0$ the point P_z is below the point P_{z_0} on any axis $t \in \bar{T}(z)$, and it has the same coordinates on all other axes. For simplicity, if such a z_0 exists we denote $z_0 = 0$.

Just as we chose to denote our reference treatment as $t_0 = 0$, our choice of $z_0 = 0$ is a convention. It may be most natural when for instance a subpopulation receives no program subsidy.

Strict targeting imposes a lot of structure on the mapping from instruments to treatments. To make this clear, we first state a definition.

Definition 5 (Preferred targeted and alternative treatments)**.** Take any observation i in the population.

(i) For $z \in \mathcal{Z}^*$, let

$$V_i^*(z) = \max_{t \in \bar{T}(z)} (\bar{\Delta}_t + u_{it})$$

and $T_i^*(z) \subset \bar{T}(z)$ denote the set of maximizers. We call the elements of T_i^* the *preferred targeted treatments.*

(ii) Also define

$$\Delta_i^* = \max_{t \in \mathcal{T}} (\underline{\Delta}_t + u_{it})$$

and let $\tau_i^* \subset \mathcal{T}$ denote the set of maximizers. We call the elements of τ_i^* the *preferred alternative treatments.*

Under strict targeting, an observation i can react to being assigned an instrument z in two ways. If z is in $\mathcal{Z}^*$, then i can choose among the treatments that z targets. Alternatively, it may choose as if no treatment was targeted (as it must if z is not in $\mathcal{Z}^*$). We now make this more rigorous by proving that observations can only opt for one of their preferred targeted treatments, if any, or for one of their preferred alternative treatments.

By Lemma 1, $\mathcal{Z}$ is either $\mathcal{Z}^*$ or $\mathcal{Z}^* \bigcup \{0\}$. We now state our main result on response-groups.

Proposition 2 (Unfiltered response groups under strict targeting)**.** *Let Assumptions 4 and 7 hold. Then for every observation i,*

(i) if $z \in \mathcal{Z}^*$, then $T_i(z)$ can only be in $T_i^*(z)$ or in τ_i^*.

(ii) if $\mathcal{Z} \neq \mathcal{Z}^*$, then $T_i(0) \in \tau_i^*$.

For simplicity, we work from now on under the assumption that the distribution of the error terms in the AS-DCM has no mass points. Then the sets τ_i^* and $T_i^*(z)$ are singletons with probability 1; with a minor abuse of notation, we let τ_i^* and $T_i^*(z)$ denote their elements[8].

Assumption 8 (Absolutely continuous errors)**.** *The distribution of the random vector $(u_{it})_{t \in \mathcal{T}}$ is absolutely continuous.*

Proposition 3 (Unfiltered classes under strict targeting)**.** *Under Assumptions 4, 7, and 8, the population contains at most two subpopulations denoted by $\mathcal{P}_1$ and $\mathcal{P}_2$.*

(i) Subpopulation $\mathcal{P}_1$ can only exist if $\mathcal{Z} = \mathcal{Z}^*$. If $i \in \mathcal{P}_1$, then $T_i(z) = T_i^*(z)$ for all $z \in \mathcal{Z}$.

[8] Note that this does not extend to the sets $\bar{Z}(t)$ and $\bar{T}(z)$, which can still have several elements.

(ii) Subpopulation $\mathcal{P}_2$ consists of classes denoted by $c(A, \tau)$, where A is a possibly empty subset of $\mathcal{Z}^*$ and τ is a treatment value. If observation i is in $c(A, \tau)$, then the following holds.

- *$T_i(z) = T_i^*(z)$ for all $z \in A$.*
- *If $A \neq \mathcal{Z}$, then $\tau_i^* = \tau$; and for all $z \in \mathcal{Z} \setminus A$, $T_i(z) = \tau$.*
- *If $A \neq \mathcal{Z}$ and $\tau \in \mathcal{T}^*$, then $\bar{Z}(\tau) \subset A$.*

(iii) If $\mathcal{Z} = \mathcal{Z}^*$, then there is no class in $\mathcal{P}_2$ with $A = \mathcal{Z}^*$.

2.2.1 Strict one-to-one targeting

Proposition 3 has a straightforward corollary under one-to-one targeting (Assumption 6). Recall that under one-to-one targeting, the sets $\bar{Z}(t)$ and $\bar{T}(z)$ are singletons and we can identify each targeting instrument with the treatment it targets. As a consequence, $T_i^*(z) = z$ for each z in $\mathcal{Z}$, and if $\tau \in \mathcal{T}^*$ then $\bar{Z}(\tau) = \{\tau\}$. This simplifies the statement of our characterization result.

Corollary 1 (Unfiltered classes under strict, one-to-one targeting)**.** *Under Assumptions 4, 6, 7, and 8, the population contains at most two subpopulations denoted by $\mathcal{P}_1$ and $\mathcal{P}_2$.*

(i) Subpopulation $\mathcal{P}_1$ can only exist if $\mathcal{Z} = \mathcal{Z}^*$. If $i \in \mathcal{P}_1$, then $T_i(z) = z$ for all $z \in \mathcal{Z}$.

(ii) Subpopulation $\mathcal{P}_2$ consists of classes denoted by $c(A, \tau)$, where A is a possibly empty subset of $\mathcal{Z}^*$ and τ is a treatment value. If observation i is in $c(A, \tau)$, then the following holds.

- *$T_i(z) = z$ for all $z \in A$.*
- *If $A \neq \mathcal{Z}$, then $\tau_i^* = \tau$; and for all $z \in \mathcal{Z} \setminus A$, $T_i(z) = \tau$.*
- *If $A \neq \mathcal{Z}$ and $\tau \in \mathcal{T}^*$, then $\tau \in A$.*

(iii) If $\mathcal{Z} = \mathcal{Z}^*$, then there is no class in $\mathcal{P}_2$ with $A = \mathcal{Z}^*$.

The subpopulation $\mathcal{P}_1$, when it exists, regroups “super-compliers”: they always take the treatment that is targeted by the instrument value they were assigned. E.g. if $\mathcal{Z} = \mathcal{Z}^* = \{1, 2, 3\}$, under strict one-to-one targeting this subpopulation would be the response group C_{123}. It is easy to see from the proof that an observation i belongs to $\mathcal{P}_1$ if and only if for all $z \in \mathcal{Z} = \mathcal{Z}^*$, $V_i^*(z) > \Delta_i^*$.

Given any (possibly empty) subset A of $\mathcal{T}^*$ and a treatment value τ, an observation i belongs to $c(A, \tau)$ if and only if

- for all $z \in A$, $V_i^*(z) > \Delta_i^*$;
- for all $z \in \mathcal{Z}^* \setminus A$, $V_i^*(z) < \Delta_i^*$;
- $\Delta_i^* = \underline{\Delta}_\tau + u_{i\tau}$.

First consider the case when A is empty. Whatever the value of the instrument z is, an observation i in $c(\varnothing, \tau)$ will take up the treatment τ that maximizes u_{it} over $\mathcal{T}$. Such observations are *always-takers* of τ. In the polar case $A = \mathcal{Z}^*$, when it is assigned a targeting instrument value ($z \in \mathcal{Z}^*$), the observation complies by picking one of the treatments it targets ($T_i(z) = T_i^*(z)$, which is z under one-to-one targeting). When both A and $\mathcal{Z} \setminus A$ are non-empty, the observation complies when the instrument z is in A, and it does not respond to changes in the value of z when it is in $\mathcal{Z} \setminus A$.

Figure 4: An unfiltered class $c(A, \tau)$ under strict one-to-one targeting

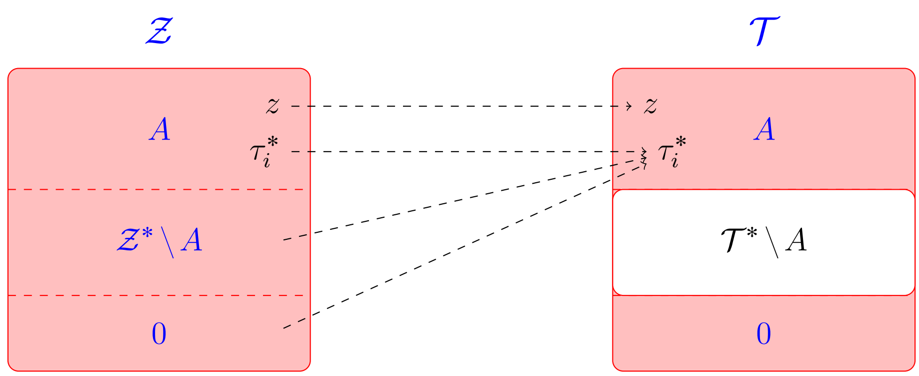


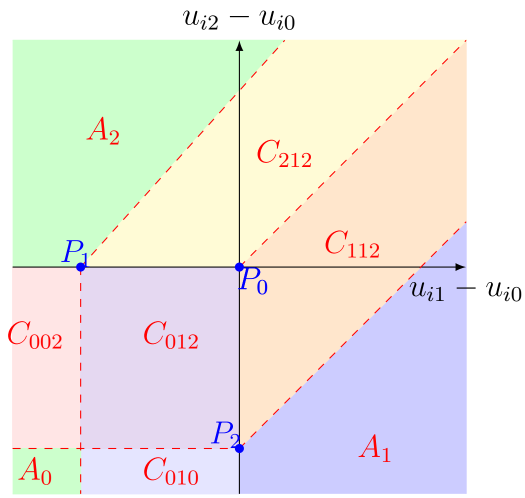
Figure 4 represents the mapping of instruments to treatments for an observation i in population $\mathcal{P}_2$ under strict one-to-one targeting. We illustrate a case for which $\mathcal{Z}^* = \mathcal{Z} \setminus \{0\}$, $\mathcal{Z}^* \setminus A$ is not empty, and $\tau \in A$. The white area shows that treatment values in $\mathcal{T}^* \setminus A$ are not assigned.

2.2.2 Applications

Example 3 (continued) To illustrate Corollary 1, we return to the ternary/ternary model of Example 3, where $\mathcal{Z}^* = \mathcal{T}^* = \{1, 2\}$ and $\mathcal{Z} = \mathcal{T} = \{0, 1, 2\}$.

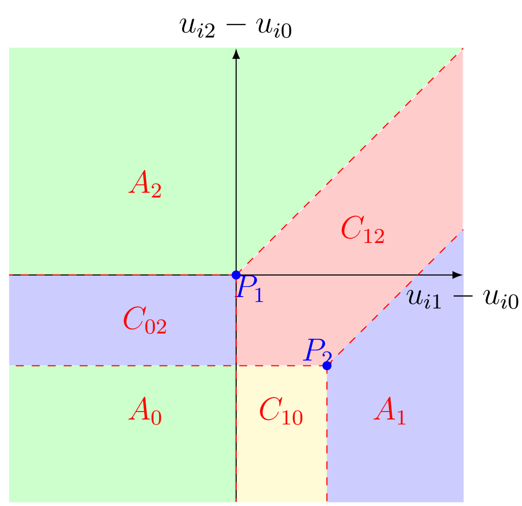
- $\mathcal{P}_1$ does not exist.
- A can be $\varnothing, \{1\}, \{2\}$, or $\{1, 2\}$, with corresponding values of τ in $\{0\}, \{0, 1\}, \{0, 2\}$ or $\{0, 1, 2\}$ respectively. The class $c(\varnothing, 0)$ corresponds to the always-takers of 0, $A_0 = C_{000}$. For $A = \{1\}$ we get C_{010} and A_1, and for $A = \{2\}$ we get C_{002} and A_2. Finally, with $A = \{1, 2\}$ we obtain the composite response group $C_{*12} = C_{012} \bigcup C_{112} \bigcup C_{212}$.

Figure 5: Unfiltered, strictly one-to-one targeted treatment: ternary/ternary model



The eight elemental response groups are illustrated in Figure 5, again with the origin in P_0. Comparing Figure 5 with Figure 3 shows the identifying power of Assumption 7. Kirkeboen, Leuven, and Mogstad (2016) used a ternary-ternary model in their investigation of field of study and earnings. We show in Appendix C.1 that our combination of Assumption 4 and Assumption 7 yields exactly the same identifying restrictions as in Kirkeboen, Leuven, and Mogstad (2016), by a quite different path. $\square$

Figure 6: Unfiltered, strict one-to-one targeting: ternary/binary model with no control



Our next example has $\mathcal{Z} = \mathcal{Z}^*$: all individuals are assigned a targeting instrument.

Example 11 (Ternary/binary model with no control)**.** Let us return to Example 9, consider $\mathcal{T} = \{0, 1, 2\}$, and $\mathcal{Z} = \mathcal{Z}^* = \{1, 2\}$: $z = 1$ strictly targets $t = 1$ and $z = 2$ strictly targets $t = 2$.

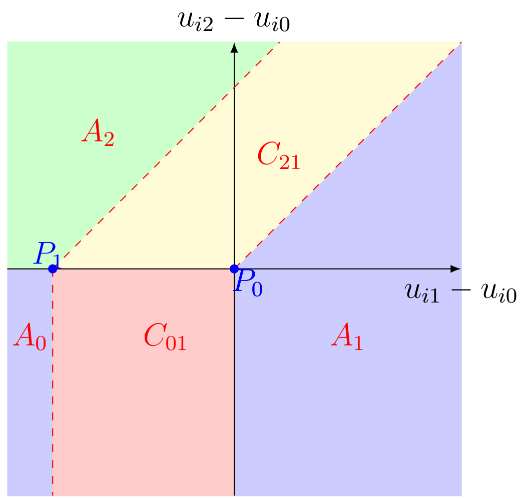
Now the subpopulation $\mathcal{P}_1$ exists; it corresponds to the response group C_{12} of super-compliers. A can be $\varnothing$, with $\tau = 0$; it can be $\{1\}$, with $\tau \in \{0,1\}$; or it can be $\{2\}$, with $\tau \in \{0,2\}$. This generates response groups A_0; C_{10} and A_1; and C_{02} and A_2. These six elemental response-groups are represented in Figure 6, where we put the origin at $u_{i0} = u_{i1} = u_{i2}$ since there is no P_0 point any more. $\square$

Sometimes one can obtain the characterization in Corollary 1 with a weaker assumption than Assumption 6. To see this, consider the following variant of Example 8.

Example 12 (Only one type of subsidy)**.** Assume that $\mathcal{T} = \{0,1,2\}$ and $\mathcal{Z} = \{0,1\}$. We interpret $z = 1$ as offering a subsidy for $t = 1$, and $z = 0$ as the absence of subsidy; treatment $t = 2$ is never subsidized. Therefore $\Delta_1(1) > \Delta_0(1)$ and $\Delta_1(2) = \Delta_0(2)$; we have $\bar{Z}(1) = \{1\}, \bar{Z}(2) = \{0,1\} = \mathcal{Z}$, and $\mathcal{T}^* = \mathcal{Z}^* = \{1\}$. Since we only have a binary instrument, strict targeting holds in this example.

The subpopulation $\mathcal{P}_1$ cannot exist here since $z = 0$ is not in $\mathcal{Z}^*$. In subpopulation $\mathcal{P}_2$, we can have classes $A = \varnothing$ with $\tau \in \{0,2\}$, and $A = \{1\}$ with $\tau \in \mathcal{T}$. The former generates the always-takers groups $A_0 = C_{00}$ and $A_2 = C_{22}$, and the latter has the two groups of compliers C_{01} and C_{21} and the always-taker group $A_1 = C_{11}$. These five elemental response-groups are illustrated in Figure 7.

Figure 7: Unfiltered, targeted treatment: ternary/binary model with only one type of subsidy



If we had not imposed $\Delta_0(2) = \Delta_1(2)$, Assumption 7 would still hold but $t = 2$ would belong to $\mathcal{T}^*$. If for instance $t = 1$ and $t = 2$ are both training programs, being offered a subsidy for $t = 1$ may also make the recipient more aware of the value of training in general. In that case we would have $\Delta_1(2) > \Delta_0(2)$ and $\bar{Z}(2) = \{1\}$, so that $\mathcal{T}^* = \{1,2\}$. We would not have one-to-one targeting anymore since $z = 1$ would target both $t = 1$ and $t = 2$.

Therefore, Corollary 1 would not apply. Proposition 3 would apply, however, allowing for a sixth response group C_{12}, with $A = \{1\}$ and $\tau = 2$.

Still, it seems likely that a subsidy for $t = 1$ would increase the appeal of $t = 1$ more than that of $t = 2$:

$$(U_1(1) - U_0(1)) - (U_1(2) - U_0(2)) = (\Delta_1(1) - \Delta_0(1)) - (\Delta_1(2) - \Delta_0(2)) > 0.$$

This is enough to rule out the possibility of the response group C_{12}. To see this, assume that $T_i(0) = 1$. This implies $U_0(1) + u_{i1} > U_0(2) + u_{i2}$, so that

$$U_1(1) + u_{i1} > U_0(2) + (U_1(1) - U_0(1)) + u_{i2} > U_1(2) + u_{i2}$$

and $T_i(1)$ cannot be 2. $\square$

2.3 Identifying Group Probabilities

Now that we have characterized response-groups, we seek to identify the probabilities of the corresponding response-groups in the unfiltered treatment model.

Definition 6 (Genralized propensity scores)**.** We write $P(t|z)$ for the generalized propensity score $\Pr(T_i = t|Z_i = z)$.

2.3.1 Strict, one-to-one Targeting

Under Assumptions 6 and 7, the response-groups are easily enumerated.

Proposition 4 (Counting response-groups under strict one-to-one targeting)**.** *Under Assumptions 4, 6, 7. and 8, the number of response-groups is*

$$N = (2|\mathcal{T}| - |\mathcal{Z}^*|) \times 2^{|\mathcal{Z}^*|-1} - (|\mathcal{T}| - 1)\mathbb{1}(\mathcal{Z} = \mathcal{Z}^*).$$

The data gives us the generalized propensity scores $P(t|z) = \Pr(T_i = t|Z_i = z)$ for $(t, z) \in \mathcal{T} \times \mathcal{Z}$. The adding-up constraints $\sum_{t\in\mathcal{T}} P(t|z) = 1$ for each $k \in \mathcal{Z}$ reduce the count of independent data points to $|\mathcal{T}^*| \times |\mathcal{Z}|$. As the probabilities of the response-groups must sum to one, we have $(N - 1)$ unknowns.

Table 1 shows some values of the number of equations $|\mathcal{T}^*| \times |\mathcal{Z}|$ and the number of unknowns $(N - 1)$ for a number of examples. The first row of $|\mathcal{T}| = 2$ and $|\mathcal{Z}| = 2$ is the standard LATE case: the response group consists of never-takers (A_0), compliers (C_{01}), and

Table 1: Number of required identifying restrictions: unfiltered treatment under strict, one-to-one targeting

Row	$\mathcal{T}$	$\mathcal{Z}$	$\mathcal{Z}^*$	$N-1$	$\lvert\mathcal{Z}\rvert\ \lvert\mathcal{T}^*\rvert$	Required	Example
(1)	{0,1}	{0,1}	{1}	2	2	0	*LATE*
(2)	{0,1,2}	{0,1}	{1}	4	4	0	Example 12
(3)	$\{0,1,\ldots,\lvert\mathcal{T}\rvert-1\}$	{0,1}	{1}	$2(\lvert\mathcal{T}\rvert-1)$	$2(\lvert\mathcal{T}\rvert-1)$	0	Example 8
(4)	{0,1,2}	{1,2}	{1,2}	5	4	1	*Example 11*
(5)	{0,1,2}	{0,1,2}	{1,2}	7	6	1	*Example 3*
(6)	{0,1,2,3}	{0,1,2}	{1,2}	11	9	2	
(7)	{0,1,2,3}	{1,2}	{1,2}	8	6	2	
(8)	{0,1,2,3}	{1,2,3}	{1,2,3}	16	9	7	
(9)	{0,1,2,3}	{0,1,2,3}	{1,2,3}	19	12	7	

always-takers (A_1). Rows (2) and its extension (3) show another case of exact identification. In other rows, as $|\mathcal{T}|$ gets larger, the degree of underidentification tends to increase.

It is not difficult to write down the equations that link observed propensity scores and group probabilities.

Proposition 5 (Identifying equations for response-groups: unfiltered treatment under strict one-to-one targeting)**.** *For any subset A of $\mathcal{Z}^*$, let A^+ denote the set $A \bigcup (\mathcal{T} \setminus \mathcal{T}^*)$. Under Assumptions 1, 2, 4, 6, 7, and 8, the empirical content of the generalized propensity scores of the unfiltered treatment model is the following system of equations:*

(i) If $\mathcal{Z} \neq \mathcal{Z}^*$:

- *for $z \in \mathcal{Z}^*$ and $t \in \mathcal{T}$:*

$$P(t|z) = \sum_{A \subset \mathcal{Z}^* \setminus \{z\}} \mathbb{1}(t \in A^+) \Pr(c(A,t)) + \mathbb{1}(t \in \mathcal{Z}^*, t = z) \sum_{\substack{A \subset \mathcal{Z}^* \\ z \in A}} \sum_{\tau \in A^+} \Pr(c(A,\tau)). \tag{2.1}$$

- *for $z \notin \mathcal{Z}^*$ and $t \in \mathcal{T}$:*

$$P(t|z) = \sum_{A \subset \mathcal{Z}^*} \mathbb{1}(t \in A^+) \Pr(c(A,t)). \tag{2.2}$$

(ii) If $\mathcal{Z} = \mathcal{Z}^*$, for $t \in \mathcal{T}$:

$$P(t|z) = \sum_{A \subset \mathcal{Z} \setminus \{z\}} \mathbb{1}(t \in A^+) \Pr(c(A,t)) + \mathbb{1}(t = z) \left(\Pr(\mathcal{P}_1) + \sum_{\substack{A \subset \mathcal{Z}, A \neq \mathcal{Z} \\ z \in A}} \sum_{\tau \in A^+} \Pr(c(A,\tau)) \right). \tag{2.3}$$

2.3.2 Applications

Proposition 5 can be applied directly to some of the rows of Table 1. According to the table, our Example 8 is just identified under strict, one-to-one targeting. Proposition 6 confirms it and gives explicit formulæ, along with simple testable predictions. To avoid repetitions, in the remainder of Section 2, we assume that Assumptions 1, 2, 4, 6, 7, and 8 hold with $\mathcal{D} = \mathcal{T}$.

Proposition 6 (Response-group probabilities in Example 8)**.** *The following probabilities are identified:*

$$\begin{aligned} \Pr(A_1) &= P(1|0), \\ \Pr(A_t) &= P(t|1) \text{ for } t \neq 1, \\ \Pr(C_{t1}) &= P(t|0) - P(t|1) \text{ for } t \neq 1. \end{aligned} \tag{2.4}$$

The model has $(|T| - 1)$ testable predictions:

$$P(t|0) \geqslant P(t|1) \text{ for } t \neq 1.$$

Row (5) of Table 1 is the ternary/ternary model of Example 3, in which eight elemental groups are non-empty. One restriction is missing to point-identify the probabilities of all eight response-groups. The following proposition shows that the probabilities of four of the eight elemental groups are point-identified: two groups of always-takers, and two groups of compliers. In addition, the probabilities of two composite groups of compliers are point-identified. The other four probabilities are constrained by three adding-up constraints.

Proposition 7 (Response-group probabilities in the ternary/ternary model of Example 3)**.**

The following probabilities are identified:

$$
\begin{aligned}
\Pr(A_1) &= P(1|2),\\
\Pr(A_2) &= P(2|1),\\
\Pr(C_{112}) &= P(1|0) - P(1|2),\\
\Pr(C_{212}) &= P(2|0) - P(2|1),\\
\Pr(C_{010} \bigcup C_{012}) &= P(0|0) - P(0|1),\\
\Pr(C_{002} \bigcup C_{012}) &= P(0|0) - P(0|2),\\
\Pr(C_{002} \bigcup C_{012} \bigcup C_{010} \bigcup A_0) &= P(0|0).
\end{aligned}
\tag{2.5}
$$

The model has the following testable implications:

$$P(1|0) \geqslant P(1|2) \tag{2.6}$$

$$P(2|0) \geqslant P(2|1) \tag{2.7}$$

$$P(0|0) \geqslant \max(P(0|1), P(0|2)). \tag{2.8}$$

The model of Example 11 is equally easy to analyze. The probabilities of two groups of always-takers are point-identified, and two equations link the probabilities of the other three elemental groups.

Proposition 8 (Response-group probabilities in the ternary/binary model of Example 11)**.** *The following probabilities are identified:*

$$
\begin{aligned}
\Pr(A_1) &= P(1|2),\\
\Pr(A_2) &= P(2|1),\\
\Pr(C_{10} \bigcup C_{12}) &= P(1|1) - P(1|2),\\
\Pr(C_{02} \bigcup A_0) &= P(0|1),\\
\Pr(C_{10} \bigcup A_0) &= P(0|2).
\end{aligned}
\tag{2.9}
$$

The model has the following testable implication:

$$P(1|1) \geqslant P(1|2). \tag{2.10}$$

2.4 Identifying Effects of Unfiltered Treatment

We now establish identification of treatment effects for the complier groups whose probabilities are identified. To simplify the exposition, we introduce one more element of notation.

Definition 7 (Conditional average group outcomes)**.** For any $z \in \mathcal{Z}, t \in \mathcal{T}$, and for any response group C with nonzero probability, we define

$$E_z(t|C) = \mathbb{E}(Y_i \mathbb{1}(T_i = t)|Z_i = z, i \in C)$$

and we call it the *conditional average group outcome.* We define the conditional average outcome by

$$\bar{E}_z(t) = \mathbb{E}(Y_i \mathbb{1}(T_i = t)|Z_i = z).$$

To give a trivial example, the LATE formula (row (1) of Table 1) is

$$\mathbb{E}(Y_i(1)|i \in C_{01}) = \frac{\bar{E}_1(1) - \bar{E}_0(1)}{P(1|1) - P(1|0)} \quad \text{and} \quad \mathbb{E}(Y_i(0)|i \in C_{01}) = \frac{\bar{E}_0(0) - \bar{E}_1(0)}{P(1|1) - P(1|0)},$$

yielding the familiar form:

$$\mathbb{E}(Y_i(1) - Y_i(0)|i \in C_{01}) = \frac{\mathbb{E}(Y_i|Z_i = 1) - \mathbb{E}(Y_i|Z_i = 0)}{\Pr(T_i = 1|Z_i = 1) - \Pr(T_i = 1|Z_i = 0)}.$$

While the $\bar{E}_z(t)$ are directly identified from the data, the conditional average group outcomes of course are not. We do know that some of them are zero; and that they combine with the group probabilities to form the observed conditional average outcomes. We will use the following identity repeatedly:

Lemma 2 (Decomposing conditional average outcomes)**.** *Let $z \in \mathcal{Z}$ and $t \in \mathcal{T}$. Then*

$$\bar{E}_z(t) = \sum_{C_{(z)}=t} \mathbb{E}(Y_i(t)|i \in C) \Pr(i \in C),$$

where $C_{(z)} = t$ means that response group C has treatment t when assigned instrument z. In addition,

$$\mathbb{E}(Y_i|Z_i = z) = \sum_{t \in \mathcal{T}} \bar{E}_z(t).$$

First consider Example 8, where the probabilities of all $(2|\mathcal{T}| - 1)$ response groups are identified (Proposition 6).

Proposition 9 (Identification in the ternary/binary model under strict one-to-one targeting). *The following quantities are point-identified:*

$$
\begin{aligned}
\mathbb{E}\left[Y_i(1)|i \in A_1\right] &= \frac{\bar{E}_0(1)}{P(1|0)}, \\
\mathbb{E}\left[Y_i(t)|i \in A_t\right] &= \frac{\bar{E}_1(t)}{P(t|1)} \text{ for } t \neq 1, \\
\mathbb{E}\left[Y_i(t)|i \in C_{t1}\right] &= \frac{\bar{E}_0(t) - \bar{E}_1(t)}{P(t|0) - P(t|1)} \text{ for } t \neq 1.
\end{aligned}
$$

However, the standard Wald estimator only partially identifies the average treatment effects on the complier groups C_{t1}:

$$
\begin{aligned}
\frac{\mathbb{E}(Y_i|Z_i = 1) - \mathbb{E}(Y_i|Z_i = 0)}{\Pr(D_i = 1|Z_i = 1) - \Pr(D_i = 1|Z_i = 0)} &= \frac{(\bar{E}_1(1) - \bar{E}_0(1)) - \sum_{t\neq 1}(\bar{E}_0(t) - \bar{E}_1(t))}{P(1|1) - P(1|0)} \\
&= \sum_{t\neq 1} \alpha_t \mathbb{E}\left[Y_i(1) - Y_i(t)|i \in C_{t1}\right],
\end{aligned}
\tag{2.11}
$$

where the weights $\alpha_t = \Pr(i \in C_{t1}|i \in \bigcup_{\tau\neq 1} C_{\tau 1}) = (P(t|0) - P(t|1))/(P(1|1) - P(1|0))$ *are positive and sum to 1.*

Proposition 9 shows that we only identify a convex combination (with point-identified weights) of the ATEs on the $|\mathcal{T}^*|$ complier groups. It is possible to bound the average treatment effects in a straightforward manner if we assume that the support of Y_i is known and finite. Alternatively, we may add conditions to achieve point identification of average treatment effects for the compliers. Assuming that the ATEs are all equal is one obvious solution. Another one is to assume the homogeneity of the average outcomes under treatment.

Corollary 2 (Treatment effects in the one-subsidy model). *Suppose that the average counterfactual outcomes under treatment* 1 *are identical for all complier groups:*

$$
\mathbb{E}\left[Y_i(1)|i \in C_{t1}\right] \text{ does not depend on } t \neq 1. \tag{2.12}
$$

Then the average treatment effects for all complier groups C_{t1} are point-identified:

$$
\begin{aligned}
&\mathbb{E}\left[Y_i(1) - Y_i(t)|i \in C_{t1}\right] \\
&= \frac{\bar{E}_1(1) - \bar{E}_0(1)}{P(1|1) - P(1|0)} - \frac{\bar{E}_0(t) - \bar{E}_1(t)}{P(t|0) - P(t|1)}.
\end{aligned}
$$

To interpret the homogeneity condition in (2.12), suppose that we are concerned with the effect of one subsidized program ($t = 1$) when other, unsubsidized programs ($t > 1$) are also available. Then (2.12) imposes that outcomes for compliers (who switch to the subsidized program when offered a subsidy) are on average the same regardless where the compliers switched from.

We now move on the ternary/ternary model in Example 3. As we mentioned earlier, in this example our assumptions allow us to use the results of Kirkeboen, Leuven, and Mogstad (2016). Their Proposition 2 tells us that

$$\beta_1 = \mathbb{E}\left[Y_i(1) - Y_i(0) \middle| i \in C_{010} \bigcup C_{012}\right],$$
$$\beta_2 = \mathbb{E}\left[Y_i(2) - Y_i(0) \middle| i \in C_{002} \bigcup C_{012}\right],$$

where β_1 and β_2 are the probability limits of the instrumental variable estimators in

$$Y_i = \beta_0 + \beta_1 \mathbb{1}(T_i = 1) + \beta_2 \mathbb{1}(T_i = 2) + \varepsilon_i. \tag{2.13}$$

We now show that we can also identify the average treatment effects for the response groups C_{112} and C_{212}, whose probabilities are point-identified.

Proposition 10 (Identification of treatment effects for Example 3)**.** *The average treatment effects of C_{112} and C_{212} are identified:*

$$\begin{aligned}&\mathbb{E}\left[Y_i(1) - Y_i(2) | i \in C_{212}\right] \\ &= \frac{\left(\mathbb{E}\left[Y_i | Z_i = 1\right] - \mathbb{E}\left[Y_i | Z_i = 0\right]\right) - \beta_1 \left(P(0|0) - P(0|1)\right)}{P(2|0) - P(2|1)}\end{aligned}$$

and

$$\begin{aligned}&\mathbb{E}\left[Y_i(2) - Y_i(1) | i \in C_{112}\right] \\ &= \frac{\left(\mathbb{E}\left[Y_i | Z_i = 2\right] - \mathbb{E}\left[Y_i | Z_i = 0\right]\right) - \beta_2 \left(P(0|0) - P(0|2)\right)}{P(1|0) - P(1|2)}.\end{aligned}$$

The average treatment effect $\mathbb{E}\left[Y_i(1) - Y_i(2) | i \in C_{212}\right]$ brings interesting information of a different nature than $\beta_1 = \mathbb{E}\left[Y_i(1) - Y_i(0) \middle| i \in C_{010} \bigcup C_{012}\right]$, which Kirkeboen, Leuven, and Mogstad (2016) focus on. We can illustrate this on the choice of college education, using a special case of Example 10. Let $z = 1$ be a tuition subsidy for a college STEM curriculum, and $z = 2$ a tuition subsidy that is available to all college students. The treatments are: not going to college ($t = 0$), studying STEM in college ($t = 1$), and opting for a non-STEM college curriculum ($t = 2$); the outcome Y is later earnings.

Both response groups C_{010}, C_{012}, and C_{212} are all comprised of individuals who will study

STEM if and only if they receive a STEM subsidy. On the other hand, individuals in $C_{010} \bigcup C_{012}$ will not go to college unless they receive a subsidy, while those in C_{212} are "college always-takers". These are quite different populations and there is no reason to expect that the effect of a STEM major on their future earnings should be the same, even on average.

3 The Filtered Treatment Model

We now turn to filtered versions of the treatment model we analyzed in the previous section. That is, we consider a model with a treatment variable $D_i \in \mathcal{D}$, where the set of filtered treatment values $\mathcal{D}$ is a non-trivial partition of the set of unfiltered treatment values $\mathcal{T} = 0, \ldots, |\mathcal{T}| - 1$. By definition, $2 \leqslant |\mathcal{D}| < |\mathcal{T}|$. We impose unordered monotonicity (Assumption 4) on the unfiltered treatment model.

Let $M : \mathcal{T} \to \mathcal{D}$ denote the "filtering map": for any $d \in \mathcal{D}$, the set of unfiltered t's that generate the observation $D = d$ is $M^{-1}(d)$. The statistics that can be identified from the data are obtained by summing their unfiltered equivalent over $t \in M^{-1}(d)$.

To make this more precise, we add superscripts T or D to response groups, conditional probabilities and expectations to indicate whether they pertain to the unfiltered treatment model or to the filtered treatment model. For instance, C^T refers to a response group in the unfiltered treatment model (a "T-response group"). The filtering map transforms C^T into a "D-response group" C^D straightforwardly: if $C^T_{(z)} = t$, then $C^D_{(z)} = M(t)$. Define $\bar{M}$ to be the component-by-component extension of M, so that $\bar{M}(C^T) \equiv (M(t_1), \ldots, M(t_{|\mathcal{Z}|}))$ for $(t_1, \ldots, t_{|\mathcal{Z}|}) \in C^T$. Then the D-response groups are

$$C^D = \bigcup_{C^T \mid \bar{M}(C^T) = C^D} C^T,$$

with probabilities

$$\Pr(i \in C^D) = \sum_{C^T \mid \bar{M}(C^T) = C^D} \Pr(i \in C^T).$$

We let $P^T(t|z)$ denote the generalized propensity scores, and $E^T_z(t|C^T)$ and $\bar{E}^T_z(t)$ the conditional average group outcomes and conditional average outcomes of Definition 7. Their equivalents in the filtered treatment model are

$$P^D(d|z) \equiv \Pr(D_i = d|Z = z) = \sum_{t \in M^{-1}(d)} P^T(t|z) \tag{3.1}$$

and for any response group C,

$$E_z^D(d|C) \equiv \mathbb{E}(Y_i \mathbb{1}(D_i = d)|Z_i = z, i \in C) = \sum_{t \in M^{-1}(d)} E_z^T(t|C).$$

Finally,

$$\bar{E}_z^D(d) \equiv \mathbb{E}(Y_i \mathbb{1}(D_i = d)|Z_i = z) = \sum_{t \in M^{-1}(d)} \bar{E}_z^T(t). \tag{3.2}$$

Since we do not observe T_i, only the left-hand sides in Equation (3.1) and Equation (3.2) are identified from the data. Finally, we let $T_i(z)$ and $D_i(z)$ denote the counterfactual treatments, and $Y_i^T(t)$ and $Y_i^D(d)$ the counterfactual outcomes.

3.1 Applications

It would be easy, but perhaps not that useful, to translate the general results of Section 2.3 and Section 2.4 to the filtered treatment model. We choose to focus here on two useful classes of examples in which the unfiltered treatment model satisfies strict, one-to-one targeting.

3.1.1 Binary filtered treatment

Let us first return to the binary instrument/multiple unfiltered treatment model (Example 8). Since $z = 1$ targets unfiltered treatment $t = 1$, it seems natural to start with a binary filtered treatment: $D_i = \mathbb{1}(T_i = 1)$. This corresponds to a filtering map M_1 defined by

- $M_1(1) = 1$
- $M_1(t) = 0$ for $t \neq 1$.

In this case, the analyst can observe whether an observation i took the targeted treatment; if not, then i could be in any other treatment cell.

The mapping of T-response groups to D-response groups is straightforward. The groups of always takers of treatment $t = d = 1$ coincide: $A_1^D = A_1^T$. The other always-takers map into the single group $A_0^D = \bigcup_{t \neq 1} A_t^T$; and the compliers C_{t1}^T combine into $C_{01}^D = \bigcup_{t \neq 1} C_{t1}^T$. Under M_1, we have $P^D(1|z) = P^T(1|z)$ for $z = 0, 1$. That is the sum of our information on group probabilities. Moving to treatment effects, we observe $\bar{E}_z^D(1) = \bar{E}_z^T(1)$ and

$$\bar{E}_z^D(0) = \sum_{t \neq 1} \bar{E}_z^T(t) \tag{3.3}$$

for $z = 0, 1$.

This allows us to identify the probabilities of D-response group and a weighted LATE, with unknown weights this time.

Proposition 11 (Identification in the filtered binary instrument model (1))**.** *(i) The probabilities of the three D-response groups are point-identified:*

$$\begin{aligned}
\Pr(A_1^D) &= \Pr(A_1^T) = P^D(1|0)\\
\Pr(A_0^D) &= \sum_{t\neq 1}\Pr(A_t^T) = 1 - P^D(1|1)\\
\Pr(C_{01}^D) &= \sum_{t\neq 1}\Pr(C_{t1}^T) = P^D(1|1) - P^D(1|0).
\end{aligned}$$

with the testable implication $P^D(1|1) \geqslant P^D(1|0)$.

(ii) The following counterfactual expectations are identified:

$$\begin{aligned}
\mathbb{E}(Y_i^D(0)|i \in A_0^D) &= \frac{\bar{E}_1^D(0)}{1 - P^D(1|1)},\\
\mathbb{E}(Y_i^D(1)|i \in A_1^D) &= \frac{\bar{E}_0^D(1)}{P^D(1|0)}.
\end{aligned}$$

(iii) The standard Wald estimator identifies the following combination of LATEs:

$$\begin{aligned}
\frac{\mathbb{E}(Y_i|Z_i = 1) - \mathbb{E}(Y_i|Z_i = 0)}{\Pr(D_i = 1|Z_i = 1) - \Pr(D_i = 0|Z_i = 0)} &= \frac{(\bar{E}_1^D(1) - \bar{E}_0^D(1)) - (\bar{E}_0^D(0) - \bar{E}_1^D(0))}{P^D(1|1) - P^D(1|0)}\\
&= \mathbb{E}(Y_i^D(1)|i \in C_{01}^D) - \sum_{t\neq 1}\alpha_t^T \mathbb{E}(Y_i^T(t)|i \in C_{t1}^T),
\end{aligned} \tag{3.4}$$

where the numbers $\alpha_t^T = \Pr(i \in C_{t1}^T | i \in C_{01}^D)$ are unidentified positive weights that sum to one.

The LHS of Equation (3.4) is a particular form of weighted LATE: the substitution of $\mathbb{E}(Y_i^D(0)|i \in C_{01}^D)$ by the weighted average in its second term reflects the lack of information of the analyst on the respective sizes of the groups C_{t1}^T within C_{01}^D, and on the dispersion of the average counterfactual outcomes when $z = 0$ across these groups. If these outcomes are homogeneous, then we get a stronger (if somewhat obvious) result.

Corollary 3 (Identification in the filtered binary instrument model (2))**.** *Assume that*

$\mathbb{E}(Y_i^T(t)|i \in C_{t1}^T)$ *is the same for all* $t \neq 1$*. Then*

$$\mathbb{E}(Y_i^D(0)|i \in C_{01}^D) = \sum_{t \neq 1} \alpha_t^T \mathbb{E}(Y_i^T(t)|i \in C_{t1}^T)$$

and the standard Wald estimator identifies the LATE on D-compliers:

$$\mathbb{E}(Y_i^D(1) - Y_i^D(0)|i \in C_{01}^D) = \frac{\mathbb{E}(Y_i|Z_i = 1) - \mathbb{E}(Y_i|Z_i = 0)}{\Pr(D_i = 1|Z_i = 1) - \Pr(D_i = 1|Z_i = 0)}.$$

3.1.2 Ternary Filtered Treatment

If we interpret $t = 0$ as a control group and all other values (including $t = 1$) as alternative treatments, then the analyst may only know whether observation i received some kind of treatment. The corresponding filtering map would be

- $M_2(0) = 0$
- $M_2(t) = 1$ for $t > 0$.

Given the structure of the problem, this is very limited information. It becomes more useful if we combine it with M_1. Let M_3 be the join of M_1 and M_2:

- $M_3(0) = 0$
- $M_3(1) = 1$
- $M_3(t) = 2$ for $t > 1$.

It allows the analyst to know whether an observation was treated, and if treated, whether it received the targeted treatment. The D-response groups consist of the always-takers $A_0^D = A_0^T$, $A_1^D = A_1^T$, $A_2^D = \bigcup_{t>1} A_t^T$; and the complier groups $C_{01}^D = C_{01}^T$ and $C_{21}^D = \bigcup_{t>1} C_{t1}^T$.

Proposition 12 (Identification in the filtered binary instrument model (3))**.** *(i) The probabilities of the five D-response groups are point-identified:*

$$\begin{aligned}
\Pr(i \in A_0^D) &= P^D(0|1),\\
\Pr(i \in A_1^D) &= P^D(1|0),\\
\Pr(i \in A_2^D) &= P^D(2|1),\\
\Pr(i \in C_{01}^D) &= P^D(0|0) - P^D(0|1),\\
\Pr(i \in C_{21}^D) &= P^D(2|0) - P^D(2|1)
\end{aligned}$$

with the testable implications $P^D(0|0) \geqslant P^D(0|1)$ *and* $P^D(2|0) \geqslant P^D(2|1)$*.*

(ii) The standard Wald estimator identifies the following combination of LATEs:

$$(3.5)\qquad \mathbb{E}(Y_i^D(1)|i \in C_{01}^D) - \alpha_0^D \mathbb{E}(Y_i^T(0)|i \in C_{01}^T) - (1-\alpha_0^D)\sum_{t>1} \beta_t^T \mathbb{E}(Y_i^T(t)|i \in C_{t1}^T)$$
$$= \frac{\mathbb{E}(Y_i|Z_i = 1) - \mathbb{E}(Y_i|Z_i = 0)}{\Pr(D_i = 1|Z_i = 1) - \Pr(D_i = 1|Z_i = 0)},$$

where

- $\alpha_0^D = \Pr(i \in C_{01}^D | i \in C_{01}^D \bigcup C_{21}^D)$ *is a positive weight, smaller than 1, identified as* $(P^D(0|0) - P^D(0|1))/(P^D(1|1) - P^D(1|0))$*;*
- *the numbers* $\beta_t^T = \Pr(i \in C_{t1}^T | i \in C_{21}^D)$ *are unidentified positive weights that sum to one.*

The extension to more general filters is trivial: any finer partition will identify more α_d^D parameters and allow the analyst to gain more information on the sizes of D-complier groups and to refine the interpretation of the Wald estimator.

Let us now turn to the ternary/ternary unfiltered treatment model of Example 3. Remember that $z = 1$ subsidizes $t = 1$ and $z = 2$ subsidizes $t = 2$. Suppose now that the analyst only observes whether an individual took one of the subsidized treatments ($d = 1$ iff $t > 0$) or not ($d = t = 0$). Then $M^{-1}(0) = 0$ and $M^{-1}(1) = \{1, 2\}$. The ternary/ternary unfiltered treatment model becomes a ternary/binary filtered treatment model. The eight T-response groups of Proposition 7 combine into five D-response groups:

$$\begin{aligned}
A_0^D &= A_0^T,\\
A_1^D &= A_1^T \bigcup A_2^T \bigcup C_{112}^T \bigcup C_{212}^T,\\
C_{001}^D &= C_{002}^T,\\
C_{010}^D &= C_{010}^T,\\
C_{011}^D &= C_{012}^T.
\end{aligned}$$

We observe the conditional probabilities $P^D(1|z)$ and the average outcomes $\bar{E}_z^D(0)$ and $\bar{E}_z^D(1)$ for $z = 0, 1, 2$.

Proposition 13 (Identification in the ternary/binary filtered model (1)). *(i) The probability of the always-taker group* A_1^D *is point-identified as* $P^D(1|0)$*. The other four*

D-response groups probabilities are connected by three equations:

$$\Pr(C^D_{01*}) = \Pr(C^D_{010}) + \Pr(C^D_{011}) = P^D(0|0) - P^D(0|1),$$
$$\Pr(C^D_{0*1}) = \Pr(C^D_{001}) + \Pr(C^D_{011}) = P^D(0|0) - P^D(0|2),$$
$$\Pr(C^D_{00*}) = \Pr(C^D_{001}) + \Pr(A^D_0) = P^D(0|1).$$

with the testable implications $P^D(0|0) \geqslant P^D(0|1)$ *and* $P^D(0|0) \geqslant P^D(0|2)$.

The four partially-identified probabilities can be parameterized as

$$\begin{aligned}
\Pr(C^D_{011}) &= p, \\
\Pr(C^D_{010}) &= P^D(0|0) - P^D(0|1) - p, \\
\Pr(C^D_{001}) &= P^D(0|0) - P^D(0|2) - p, \\
\Pr(A^D_0) &= P^D(0|2) + P^D(0|1) - P^D(0|0) + p,
\end{aligned}$$

where

$$\max(0, P^D(0|0) - P^D(0|1) - P^D(0|2)) \leqslant p \leqslant P^D(0|0) - \max(P^D(0|1), P^D(0|2)).$$

(ii) The following average conditional counterfactual outcomes are point-identified:

$$\begin{aligned}
\mathbb{E}(Y^D_i(0)|i \in C^D_{00*}) &= \frac{\bar{E}^D_1(0)}{P^D(0|1)}, \\
\mathbb{E}(Y^D_i(0)|i \in C^D_{01*}) &= \frac{\bar{E}^D_0(0) - \bar{E}^D_1(0)}{P^D(0|0) - P^D(0|1)}, \\
\mathbb{E}(Y^D_i(1)|i \in C^D_{01*}) &= \frac{\bar{E}^D_1(1) - \bar{E}^D_0(1)}{P^D(0|0) - P^D(0|1)}, \\
\mathbb{E}(Y^D_i(1)|i \in A^D_1) &= \frac{\bar{E}^D_1(0)}{P^D(1|0)}.
\end{aligned}$$

(iii) The standard Wald estimators identify the LATE on C^D_{01*} *and on* C^D_{0*1}*:*

$$\mathbb{E}(Y^D_i(1) - Y_i(0)|i \in C^D_{01*}) = \frac{\mathbb{E}(Y_i|Z_i = 1) - \mathbb{E}(Y_i|Z_i = 0)}{\Pr(D_i = 1|Z_i = 1) - \Pr(D_i = 1|Z_i = 0)}, \tag{3.6}$$

$$\mathbb{E}(Y^D_i(1) - Y_i(0)|i \in C^D_{0*1}) = \frac{\mathbb{E}(Y_i|Z_i = 2) - \mathbb{E}(Y_i|Z_i = 0)}{\Pr(D_i = 1|Z_i = 2) - \Pr(D_i = 1|Z_i = 0)}. \tag{3.7}$$

Note that the width of the interval on the unknown p cannot be larger than $\min(P^D(0|1), P^D(0|2))$: if either instrument $z = 1, 2$ is very effective at getting people to adopt public transportation, the sizes of all D-response groups will be almost point-identified. Since the average counterfactual outcomes on elemental D-response groups are connected by equations like

$$\mathbb{E}(Y_i(d)|i \in C^D_{01*}) = q\mathbb{E}(Y_i(d)|i \in C^D_{011}) + (1-q)\mathbb{E}(Y_i(d)|i \in C^D_{010})$$

with $q = p/(P^D(0|0) - P^D(0|1))$, one could go further and impose homogeneity assumptions to improve the identification of elemental LATEs.

3.2 Filtered Factorial Design

We now return to Example 4, which featured a factorial experimental design. Recall that we had $z = 0 \times 0, 0 \times 1, 1 \times 0, 1 \times 1$, and $\mathcal{T} = \{0, 1, 2\}$. Each instrument combines two binary instruments: the first one is meant to promote treatment $t = 1$ and the second one promotes $t = 2$. We focus here on the case when there is no complentarity (positive or negative) between the two binary instruments[9]: $\Delta^T_{1\times 1}(1) = \Delta^T_{1\times 0}(1)$ and $\Delta^T_{1\times 1}(2) = \Delta^T_{0\times 1}(2)$. This would hold for instance if each binary instrument is a price subsidy and prices enter mean utilities additively—a common asssumption in discrete choice models. As we saw in Section 2.1, we have $\bar{Z}(1) = \{1 \times 0, 1 \times 1\}$ and $\bar{Z}(2) = \{0 \times 1, 1 \times 1\}$, so that this treatment model does not statisfy one-to-one targeting. On the other hand, we also saw that strict targeting holds if each binary instrument only has an effect on the treatment value that it targets. We will impose the corresponding assumptions $\Delta^T_{0\times 1}(1) = \Delta^T_{0\times 0}(1)$ and $\Delta^T_{1\times 0}(2) = \Delta^T_{0\times 0}(2)$.

3.2.1 Identification

Let us now introduce a filter, so that the analyst only observes $D_i = \mathbb{1}(T_i > 0)$.[10] This yields a ternary/binary filtered treatment model, much as in the previous subsection. There are two important differences—the instrument takes four values rather than three, and we imposed several constraints on the mean utilities:

$$\begin{aligned} \Delta^T_{1\times 1}(1) &= \Delta^T_{1\times 0}(1) \\ \Delta^T_{1\times 1}(2) &= \Delta^T_{0\times 1}(2) \\ \Delta^T_{0\times 1}(1) &= \Delta^T_{0\times 0}(1) \\ \Delta^T_{1\times 0}(2) &= \Delta^T_{0\times 0}(2). \end{aligned} \tag{3.8}$$

[9]We use a superscript T to remind the reader that the argument in parentheses is an unfiltered treatment value in $\mathcal{T}$.

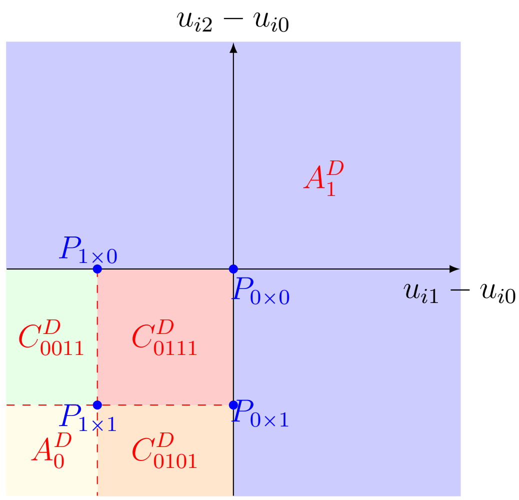
[10]Appendix C.2 provides a variant of filtered factorial design.

In spite of the filtering, they will allow us to point-identify the relevant LATEs. To see this, first note that for any given observation i,

$$D_i(z) = 0 \text{ iff } u_{i0} > \max(\Delta_z^T(1) + u_{i1}, \Delta_z^T(2) + u_{i2}), \tag{3.9}$$

so that the filtered treatment model has the structure of a double hurdle model (Example 2).

Figure 8: Filtered Factorial Design



First note that under our assumptions, the right hand side is largest when $z = 1 \times 1$. Therefore if $D_i(1\times1) = 0$, observation i always takes $d = 0$. If on the other hand $D_i(0\times0) = 1$, then i is in A_1^D since the right-hand side can only be larger for the other instrument values. Denote indices in response-groups in the order $0\times0, 1\times0, 0\times1, 1\times1$. The preceding arguments leave only the D-response groups C_{0**1}^D. The group C_{0001}^D cannot exist since in the absence of complementarity between the binary instruments,

$$\max(\Delta_{1\times1}^T(1) + u_{i1}, \Delta_{1\times1}^T(2) + u_{i2}) = \max(\Delta_{1\times0}^T(1) + u_{i1}, \Delta_{0\times1}^T(2) + u_{i2}).$$

The three other groups are[11]:

- the eager compliers C_{0111}^D: any instrument except 0×0 causes them to adopt $d = 1$
- the reluctant compliers C_{0011}^D and C_{0101}^D: they only adopt $d = 1$ if the right binary instrument is switched on.

The resulting five D-response groups are shown in Figure 8. Table 2 shows which groups take $D_i = d$ when $Z_i = z$.

[11] We borrow here the terminology of Mogstad, Torgovitsky, and Walters (2019), which they apply to a rather different model.

Table 2: D-response Groups

	$D_i(z)=0$	$D_i(z)=1$
$z=0$	$C^D_{0**} = A^D_0 \bigcup C^D_{001} \bigcup C^D_{010} \bigcup C^D_{011}$	A^D_1
$z=1$	$C^D_{00*} = A^D_0 \bigcup C^D_{001}$	$C^D_{*1*} = A^D_1 \bigcup C^D_{010} \bigcup C^D_{011}$
$z=2$	$C^D_{0*0} = A^D_0 \bigcup C^D_{010}$	$C^D_{**1} = A^D_1 \bigcup C^D_{001} \bigcup C^D_{011}$

Proposition 14 (Identifying the Filtered Factorial Design Model). *(i) the probabilities of the D-response groups are point-identified by*

$$
\begin{aligned}
\Pr(A^D_0) &= P^D(1|0\times 0)\\
\Pr(A^D_1) &= P^D(0|1\times 1)\\
\Pr(C^D_{0011}) &= P^D(0|1\times 0) - P^D(0|1\times 1)\\
\Pr(C^D_{0101}) &= P^D(0|0\times 1) - P^D(0|1\times 1)\\
\Pr(C^D_{0111}) &= P^D(0|0\times 0) + P^D(0|1\times 1) - P^D(0|1\times 0) - P^D(0|0\times 1),
\end{aligned}
$$

and the model has three testable implications:

$$
\begin{aligned}
P^D(0|1\times 0) &\geqslant P^D(0|1\times 1),\\
P^D(0|0\times 1) &\geqslant P^D(0|1\times 1),\\
P^D(0|0\times 0) + P^D(0|1\times 1) &\geqslant P^D(0|1\times 0) + P^D(0|0\times 1).
\end{aligned}
$$

(ii) The LATEs on the three groups of compliers are point-identified by

$$
\begin{aligned}
&\mathbb{E}(Y^D_i(1) - Y^D_i(0)|i \in C^D_{0101}) = \frac{\mathbb{E}(Y|Z=1\times 1) - \mathbb{E}(Y|Z=0\times 1)}{P^D(1|1\times 1) - P^D(1|0\times 1)}\\
&\mathbb{E}(Y^D_i(1) - Y^D_i(0)|i \in C^D_{0011}) = \frac{\mathbb{E}(Y|Z=1\times 1) - \mathbb{E}(Y|Z=1\times 0)}{P^D(1|1\times 1) - P^D(1|1\times 0)}\\
&\mathbb{E}(Y^D_i(1) - Y^D_i(0)|i \in C^D_{0111}) =\\
&\frac{\mathbb{E}(Y|Z=1\times 0) + \mathbb{E}(Y|Z=0\times 1) - \mathbb{E}(Y|Z=1\times 1) - \mathbb{E}(Y|Z=0\times 0)}{P^D(1|1\times 0) + P^D(1|1\times 0) - P^D(1|1\times 1) - P^D(1|0\times 0)}.
\end{aligned}
$$

Proposition 14 states that (i) the average treatment effects for reluctant compliers are identified by suitable Wald statistics and that (ii) the average treatment effect for eager compliers is identified by a ratio between difference-in-differences (DiD) population quantities. The latter estimand can be viewed as a two-dimensional version of Wald statistics.

3.2.2 Estimation of $\mathbb{E}(Y_i^D(1) - Y_i^D(0)|i \in C_{0111}^D)$ with covariates

Most estimands in the paper are expressed in terms of simple Wald estimators, which can be easily estimated with covariates (e.g. Frölich, 2007). Two exceptional cases are $\mathbb{E}(Y_i^D(1) - Y_i^D(0)|i \in C_{0111}^D)$ in Proposition 14 and $\mathbb{E}(Y_i^D(1) - Y_i^D(0)|i \in C_{0001}^D)$ in Proposition 15 in Appendix C.2.

We here discuss how to estimate $\mathbb{E}(Y_i^D(1) - Y_i^D(0)|i \in C_{0111}^D)$ with covariates. Estimation of $\mathbb{E}(Y_i^D(1) - Y_i^D(0)|i \in C_{0001}^D)$ is similar. Introduce covariates X_i explicitly and define:

$$\mathbb{E}[Y_i(1) - Y_i(0)|i \in C_{0111}^D, X_i = x] = \frac{\mathrm{DiD}_Y(x)}{\mathrm{DiD}_D(x)},$$

where

$$\begin{aligned}
\mathrm{DiD}_Y(x) &:= \mathbb{E}[Y_i|Z_i = 1 \times 0, X_i = x] + \mathbb{E}[Y_i|Z_i = 0 \times 1, X_i = x] \\
&\quad - \mathbb{E}[Y_i|Z_i = 1 \times 1, X_i = x] - \mathbb{E}[Y_i|Z_i = 0 \times 0, X_i = x], \\
\mathrm{DiD}_D(x) &:= \Pr[T_i = 1|Z_i = 1 \times 0, X_i = x] + \Pr[T_i = 1|Z_i = 0 \times 1, X_i = x] \\
&\quad - \Pr[T_i = 1|Z_i = 1 \times 1, X_i = x] - \Pr[T_i = 1|Z_i = 0 \times 0, X_i = x].
\end{aligned}$$

Then,

$$\mathbb{E}(Y_i^D(1) - Y_i^D(0)|i \in C_{0111}^D) = \mathbb{E}\left[\frac{\mathrm{DiD}_Y(X)}{\mathrm{DiD}_D(X)}\middle| i \in C_{0111}^D\right].$$

Lemma 3. *Assume that $DiD_D(X) \neq 0$ almost surely. Then,*

$$\mathbb{E}(Y_i^D(1) - Y_i^D(0)|i \in C_{0111}^D) = \frac{\mathbb{E}[DiD_Y(X)]}{\mathbb{E}[DiD_D(X)]}.$$

Lemma 3 suggests the following two-step estimation strategy: first, estimate $\mathbb{E}[Y_i|Z_i = k, X_i = x]$ and $\Pr[T_i = 1|Z_i = k, X_i = x]$ for each $k \in \{0 \times 0, 1 \times 0, 1 \times 0, 1 \times 1\}$ and $x \in \{X_1, \ldots, X_n\}$; second, evaluate $\mathrm{DiD}_Y(X_i)$ and $\mathrm{DiD}_D(X_i)$, construct their averages and take the ratio.

For example, the first step can be implemented using sieve estimators. In view of Ackerberg, Chen, and Hahn (2012); Ackerberg, Chen, Hahn, and Liao (2014), the resulting two-step sieve estimator is semiparametrically efficient, and furthermore, conventional normal inference, pretending that we have a two-step parametric model, is valid for semiparametric inference. For brevity of the paper, we omit details.

4 Empirical Examples

4.1 The Student Achievement and Retention Project

In this section, we revisit Angrist, Lang, and Oreopoulos (2009), who analyzed the Student Achievement and Retention Project. STAR was a randomized evaluation of academic services and incentives for college freshmen at a Canadian university. It was a factorial design, with two binary instruments. The Student Fellowship Program (SFP) offered students the chance to win merit scholarships for good grades in the first year; the Student Support Program (SSP) offered students access to both a peer-advising service and a supplemental instruction service. Entering first-year undergraduates were randomly assigned to one of four groups: a control group ($z = 0 \times 0$), SFP only ($z = 0 \times 1$), SSP only ($z = 1 \times 0$), and an intervention offering both (SFSP, $z = 1 \times 1$).

The data indicates whether a student who was offered a program signed up, and whether those who were offered SSP or SFSP and signed up made use of the services of SSP. Angrist, Lang, and Oreopoulos (2009) used the sign-up as the treatment variable. They comment that "in the SSP and SFSP, a further distinction can be made between compliance via sign up and compliance via service use" (p. 147). Many students who sign up did not in fact use the services; this suggests defining an unfiltered treatment variable as a pair $T_i = (A_i, S_i)$, where $A_i = 1$ (for "accept") denotes that student i signed up and $S_i = 1$ (for "services") that (s)he used the services provided by SSP.

Obviously, $S_i = 0$ if $A_i = 0$. Hence T_i can only take three values: $(0,0), (1,0)$, and $(1,1)$. With a slight change in notation, we model the choice as

$$T_i(z) = \arg\max \left(u_i(0,0), \Delta_z^T(1,0) + u_i(1,0), \Delta_z^T(1,1) + u_i(1,1)\right).$$

While there are four instrument values and three treatment values, this is in fact a ternary/ternary model, with some specific features. First note that $T_i = 0$ for all observations in the control group; this allows us to set $\Delta_{0\times 0}^T(1,0)$ and $\Delta_{0\times 0}^T(1,1)$ at minus infinity. In addition, S_i can only be zero if $z = 0 \times 1$, so that we can set $\Delta_{0\times 1}^T(1,0)$ and $\Delta_{0\times 1}^T(1,1)$ at minus infinity too. As a consequence, we do not lose any information by redefining the control group to be $0 \equiv \{0 \times 0, 0 \times 1\}$.

In addition, students should be more likely to sign up under $z = 1 \times 1$ than under $z = 1 \times 0$, as the former adds the lure of a fellowship. We will also assume that it makes them more likely to use the services—an assumption that we will test below. Then both treatment values $(1,0)$ and $(1,1)$ are targeted by 1×1, but they cannot be strictly targeted. Take for instance $\bar{Z}(1,1)=\{1\times 1\}$; strict targeting would require $\Delta^T_{1\times 0}(1,1)=\Delta^T_0(1,1)$, which is minus infinity.

Rather than to pursue with the unfiltered treatment model, let us move on to filtered models. In our terminology, Angrist, Lang, and Oreopoulos (2009) chose to use a particular filter $M(A,S)=A$, which is close to intent-to-treat as they point out. Here we take $M(A,S)=S$ instead: we define

$$D_i(z)=S_i(z)=\mathbb{1}\left(\Delta^T_z(1,1)+u_i(1,1)>\max(u_i(0,0),\Delta^T_z(1,0)+u_i(1,0))\right). \tag{4.1}$$

Since the SFP incentives applied to the first year grades only, we take the grades in the second year as our outcome variable Y_i.

Equation (4.1) has a similar structure to the double hurdle model of Equation (3.9). The models are quite different, however. This new filtered model has $D_i(0)=0$ with probability one; and we are assuming that an offer of a fellowship makes students more likely to use the services. Of the a priori four possible response-groups $C^D_{0,d,d'}$ for $d,d'=0,1$, this assumption eliminates one: if $D_i(1\times 1)=0$ then a fortiori $D_i(1\times 0)=0$. This leaves three groups: the never-takers A^D_0, and two groups of compliers C^D_{001} and C^D_{011}. The group C^D_{001} consists of reluctant compliers, who only use SSP if it is offered along with SFP. Those in C^D_{011} are eager compliers: they use SSP whenever it is offered to them with or without a fellowship.

Remember that $P^D(1|z):=\Pr(D_i=1|Z_i=z)$ for $z=0,1\times 0,1\times 1$. Then $P^D(1|0)=0$ and the proportions of the three response-groups are given by

$$\begin{aligned}
\Pr(A^D_0) &= 1-P^D(1|1\times 1)\\
\Pr(C^D_{001}) &= P^D(1|1\times 1)-P^D(1|1\times 0)\\
\Pr(C^D_{011}) &= P^D(1|1\times 0).
\end{aligned}$$

Note that given Equation (4.1),

$$\begin{aligned}
P^D(1|1\times 0) = \Pr\big(&u_i(1,1)-u_i(0,0)>-\Delta^T_{1\times 0}(1,1)\\
&\text{and }\ u_i(1,1)-u_i(1,0)>\Delta^T_{1\times 0}(1,0)-\Delta^T_{1\times 0}(1,1)\big)\\
P^D(1|1\times 1) = \Pr\big(&u_i(1,1)-u_i(0,0)>-\Delta^T_{1\times 1}(1,1)\\
&\text{and }\ u_i(1,1)-u_i(1,0)>\Delta^T_{1\times 1}(1,0)-\Delta^T_{1\times 1}(1,1)\big).
\end{aligned}$$

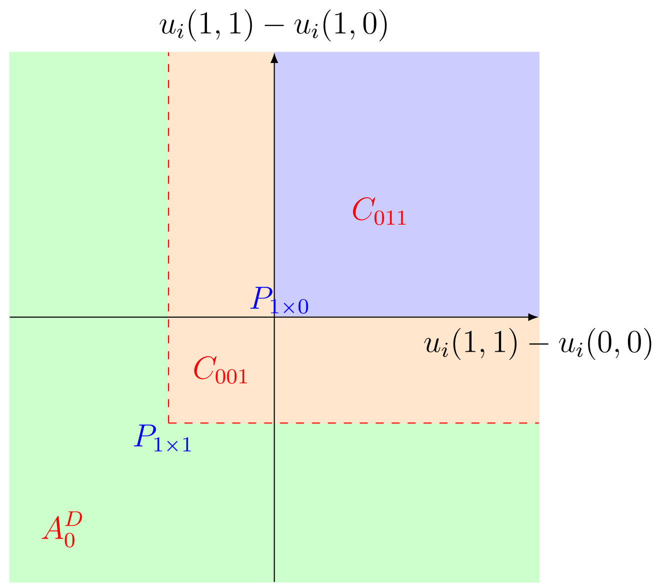
Our assumption that students are more likely to use the services under SFSP translates into

$$\Delta^T_{1\times 1}(1,1) > \Delta^T_{1\times 0}(1,1) \text{ and } \Delta^T_{1\times 1}(1,0) - \Delta^T_{1\times 1}(1,1) < \Delta^T_{1\times 0}(1,0) - \Delta^T_{1\times 0}(1,1).$$

Figure 9 illustrates a configuration in which these inequalities hold, where

$P_{1\times 0} = (-\Delta^T_{1\times 0}(1,1), \Delta^T_{1\times 0}(1,0) - \Delta^T_{1\times 0}(1,1))$ and $P_{1\times 1} = (-\Delta^T_{1\times 1}(1,1), \Delta^T_{1\times 1}(1,0) - \Delta^T_{1\times 1}(1,1))$.

Figure 9: STAR example



Under our assumptions, it is straightforward to show that

$$\mathbb{E}[Y_i^D|Z_i = 1\times 0] - \mathbb{E}[Y_i^D|Z_i = 0] = \mathbb{E}[Y_i^D(1) - Y_i^D(0)|i \in C^D_{011}]\Pr(i \in C^D_{011}),$$
$$\mathbb{E}[Y_i^D|Z_i = 1\times 1] - \mathbb{E}[Y_i^D|Z_i = 1\times 0] = \mathbb{E}[Y_i^D(1) - Y_i^D(0)|i \in C^D_{001}]\Pr(i \in C^D_{001}).$$

Therefore, we have

$$\mathbb{E}[Y_i^D(1) - Y_i^D(0)|i \in C^D_{011}] = \frac{\mathbb{E}[Y_i^D|Z_i = 1\times 0] - \mathbb{E}[Y_i^D|Z_i = 0]}{P^D(1|1\times 0)},$$
$$\mathbb{E}[Y_i^D(1) - Y_i^D(0)|i \in C^D_{001}] = \frac{\mathbb{E}[Y_i|Z_i = 1\times 1] - \mathbb{E}[Y_i^D|Z_i = 1\times 0]}{P^D(1|1\times 1) - P^D(1|1\times 0)}.$$

Since $\Pr(D_i = 1|Z_i = 0) = 0$, the first estimand is the IV formula of Bloom (1984); the second estimand is the LATE formula of Imbens and Angrist (1994).

Table 3 reports estimation results. We only focus on the subsample of women since the STAR program had no effect on men. Panel A of Table 3 shows the estimated proportions

Table 3: Empirical Results from STAR

Panel A.	Proportion of Compliers
$\Pr(i \in C_{011}^D)$	0.288
	(0.040)
$\Pr(i \in C_{001}^D)$	0.245
	(0.069)

Panel B.	GPA	On probation or withdrawal	Good standing	Credits earned
$\mathbb{E}[Y_i \mid Z_i = 1 \times 0] - \mathbb{E}[Y_i \mid Z_i = 0]$	0.084	0.045	-0.039	-0.065
	(0.088)	(0.043)	(0.046)	(0.147)
$\mathbb{E}[Y_i(1) - Y_i(0) \mid i \in C_{011}^D]$	0.291	0.156	-0.137	-0.225
	(0.303)	(0.152)	(0.161)	(0.516)
$\mathbb{E}[Y_i \mid Z_i = 1 \times 1] - \mathbb{E}[Y_i \mid Z_i = 1 \times 0]$	0.186	-0.141	0.163	0.350
	(0.127)	(0.058)	(0.065)	(0.208)
$\mathbb{E}[Y_i(1) - Y_i(0) \mid i \in C_{001}^D]$	0.758	-0.576	0.664	1.427
	(0.532)	(0.265)	(0.305)	(0.887)

Notes. Standard errors are in the parentheses. The estimation sample consists of women in the control, SSP and SFSP groups. The sample size is $n = 837$. The outcome variables are second year GPA, an indicator of probation or withdrawal in the second year, a "good standing" variable that indicates whether students attempted at least four credits and were not on probation, and the credits earned. Estimates of treatment effects are computed based on linear regression models using the full set of controls used in Angrist, Lang, and Oreopoulos (2009).

of the two complier groups: 0.288 for C^D_{011} and 0.245 for C^D_{001}. The majority group is the never-takers whose share is 0.467. This is because the usage of SSP was low. Panel B reveals remarkable heterogeneity between the two complier groups. We do not find any significant treatment effect for C^D_{011}, whereas we do find sizeable and significant impact on probation/withdrawal and good standing for C^D_{001}.[12] As can be seen in Figure 9, C^D_{001} is closer to the group of never-takers: they have higher unobserved disutilities of using academic support services than those in C^D_{011}. However, those in C^D_{001} reaped greater benefits of using the SSP by avoiding probation or withdrawal in the second year.

The main parameter of interest in Angrist, Lang, and Oreopoulos (2009) was the intent-to-treat (ITT) effect of the SFSP program: $\mathbb{E}[Y_i|Z_i = 1 \times 1] - \mathbb{E}[Y_i|Z_i = 0 \times 0]$ in our notation. Our analysis suggests that the ITT effect of the SFSP program is a mix of two very different treatment effects. This highlights the importance of unbundling heterogeneous complier groups.

4.2 Head Start

Let us now reexamine Kline and Walters's (2016) analysis of the Head Start Impact Study (HSIS) using our framework. The structure of HSIS is identical to that of Example 12. The treatments consist of no preschool (n), Head Start (h), and other preschool centers (c): $\mathcal{T} = \{n, h, c\}$. We will take $t_0 = n$ as our reference treatment. The instrument is binary, with a control group ($z = 0$) and a group that is offered admission to Head Start ($z = 1$). This gives five response groups: $A_n = C_{nn}$, $A_c = C_{cc}$, $A_h = C_{hh}$, C_{nh}, and C_{ch}. The first three groups are always-takers and the last two groups are compliers.

4.2.1 Group proportions and counterfactual means

Their proportions in the sample are given by (2.4) in Proposition 6; they are shown in Panel A of Table 4. As expected, they coincide with those in Kline and Walters (2016).

Panel B of Table 4 shows the counterfactual means of test scores as per Proposition 9. Among those that are point-identified, the average test scores are the highest for the groups who always choose other preschool centers (about 0.3 standard deviation). There is a noticeable difference between the two complier groups: $\mathbb{E}[Y_i(n)|i \in C_{nh}]$ is negative, but $\mathbb{E}[Y_i(c)|i \in C_{ch}]$ is above 0.1 standard deviation. This indicates that among compliers, the children who used other centers had higher scores than those who stayed at home. Head

[12] The point estimates for probation/withdrawal and good standing are very large in absolute value; however, the standard errors are large as well, resulting in wide confidence intervals. This is partially because the sample size is relatively small and partially because the estimand is the ratio of two population quantities with the small denominator.

Start may therefore attract some children who already were at a good preschool. Kline and Walters (2016) call this pattern the "substitution effect" of Head Start. However, the way we achieve the identification of $\mathbb{E}[Y_i(n)|i \in C_{nh}]$ and $\mathbb{E}[Y_i(c)|i \in C_{ch}]$ is new.

Table 4: Proportions, Counterfactual Means and Treatment Effects by Response Groups

	3-year-olds	4-year-olds	Pooled
Panel A. Proportions of Response Groups via Proposition 6			
Always – no preschool (A_n)	0.092	0.099	0.095
Always – Head Start (A_h)	0.147	0.122	0.136
Always – other centers (A_c)	0.058	0.114	0.083
Compliers from n to h (C_{nh})	0.505	0.393	0.454
Compliers from c to h (C_{ch})	0.198	0.272	0.232
Panel B. Counterfactual Means of Test Scores via Proposition 9			
$\mathbb{E}[Y_i(n)\| \in A_n]$	-0.050	-0.017	-0.035
$\mathbb{E}[Y_i(h)\| \in A_h]$	0.007	-0.080	-0.028
$\mathbb{E}[Y_i(c)\| \in A_c]$	0.293	0.330	0.316
$\mathbb{E}[Y_i(n)\|i \in C_{nh}]$	-0.027	-0.116	-0.062
$\mathbb{E}[Y_i(c)\|i \in C_{ch}]$	0.112	0.144	0.129
Panel C. Counterfactual Means of Test Scores via Corollary 2			
$\mathbb{E}[Y_i(h)\|i \in C_{nh}] = \mathbb{E}[Y_i(h)\|i \in C_{ch}]$	0.252	0.169	0.216
Panel D. Treatment Effects via Corollary 2			
$\mathbb{E}[Y_i(h) - Y_i(n)\|i \in C_{nh}]$ for compliers from 'n' to 'h'	0.279	0.285	0.278
	(0.063)	(0.076)	(0.050)
$\mathbb{E}[Y_i(h) - Y_i(c)\|i \in C_{ch}]$ for compliers from 'c' to 'h'	0.140	0.025	0.087
	(0.089)	(0.097)	(0.063)
$\mathbb{E}[Y_i(h) - Y_i(n)\|i \in C_{nh}] - \mathbb{E}[Y_i(h) - Y_i(c)\|i \in C_{ch}]$	0.139	0.260	0.191
	(0.098)	(0.115)	(0.071)

Notes: Head Start (h), other centers (c), no preschool (n). Standard errors in parentheses are clustered at the Head Start center level.

4.2.2 Treatment Effects

To fully measure the substitution effect, one needs to identify $\mathbb{E}[Y_i(h)|i \in C_{nh}]$ and $\mathbb{E}[Y_i(h)|i \in C_{ch}]$. However, under Proposition 9, they are only partially identified by

$$\begin{aligned}
&\mathbb{E}[Y_i(h)|i \in C_{nh}]\{\Pr(T_i = n|Z_i = 0) - \Pr(T_i = n|Z_i = 1)\}\\
&+ \mathbb{E}[Y_i(h)|i \in C_{ch}]\{\Pr(T_i = c|Z_i = 0) - \Pr(T_i = c|Z_i = 1)\}\\
&= \mathbb{E}[Y_i \mathbb{1}(T_i = h)|Z_i = 1] - \mathbb{E}[Y_i \mathbb{1}(T_h = 1)|Z_i = 0].
\end{aligned}$$

This is exactly the formula on Kline and Walters (2016, pp.1811), where they point out that the LATE for Head Start is a weighted average of "subLATEs" with weights S_c and $(1 - S_c)$ with

$$S_c := \frac{\Pr(C_{ch})}{\Pr(C_{nh}) + \Pr(C_{ch})} = \frac{\Pr(T_i = c|Z_i = 0) - \Pr(T_i = c|Z_i = 1)}{\Pr(T_i \neq h|Z_i = 0) - \Pr(T_i \neq h|Z_i = 1)}.$$

Kline and Walters (2016) first tried to estimate $\mathbb{E}[Y_i(h) - Y_i(c)|i \in C_{ch}]$ and $\mathbb{E}[Y_i(h) - Y_i(n)|i \in C_{nh}]$ separately using two-stage least squares (2SLS), using interaction of the instrument with covariates or experimental sites in an attempt to generate enough variation. They acknowledged the limitations of this interacted 2SLS approach and developed a parametric selection model à la Heckman (1979). Using a parametric selection model and pooled cohorts, Kline and Walters (2016, Table VIII, column (4) full model) obtain estimates of the treatment effect of 0.370 (0.088) for C_{nh} and -0.093 (0.154) for C_{ch} respectively (standard errors in parentheses).

Our Corollary 2 provides an alternative approach to separating the two treatment effects. If we assume that $\mathbb{E}[Y_i(h)|i \in C_{nh}] = \mathbb{E}[Y_i(h)|i \in C_{ch}]$, we can point-identify the average treatment effects for both groups of compliers. The resulting estimates are shown in Panels C and D of Table 4. The average impact on test scores of participating in Head Start is around 0.28 for C_{nh}, whereas it is smaller and insignificant for C_{ch}. Their difference is significantly different when the two cohorts are pooled together.

We obtained these estimates of the treatment effects by a completely different route than Kline and Walters (2016). While the two sets of estimates are similar, our estimate of the difference between the treatment effects on the two groups of compliers is twice smaller. Our homogeneity assumption $\mathbb{E}[Y_i(h)|i \in C_{nh}] = \mathbb{E}[Y_i(h)|i \in C_{ch}]$ may be too strong. It might be more plausible to assume that

$$\mathbb{E}[Y_i(h)|i \in C_{nh}] \leqslant \mathbb{E}[Y_i(h)|i \in C_{ch}]$$

as children who would not attend preschool in the absence of offer to Head Start are likely to be less well-prepared than children who would attend other preschools. Then our estimated difference between the two complier groups will be a lower bound of the true difference.

Concluding Remarks

We have shown that our targeting and filtering concepts are a useful way to analyze models with multivalued treatments and multivalued instruments. While our characterization is sharpest under strict, one-to-one targeting (Corollary 1), our framework remains useful even without strict targeting. In addition to the examples we discussed in the text and to the two applications we revisited, we give an example in Appendix C.2, with a ternary/ternary model where the analyst only observed the *least*-preferred treatment in a factorial design.

Our paper only analyzed discrete-valued instruments and treatments. Some of the notions we used would extend naturally to continuous instruments and treatments: the definitions of targeting, one-to-one targeting, and filtering would translate directly. Strict targeting, on the other hand, is less appealing in a context in which continuous values may denote intensities. Our earlier paper (Lee and Salanié, 2018) can be seen as analyzing continuous-instruments/discrete-treatments filtered models; so does Mountjoy's (2019)'s study of 2-year colleges. Extending our analysis to models with continuous treatments is an interesting topic for further research.

Appendices

A Proofs for Section 2

Proof of Proposition 1. Let $T_i(t) = 0$ for some $t \in \mathcal{T}^*$. Then $u_{i0} > \bar{\Delta}_t + u_{it}$. However, $\bar{\Delta}_t > \Delta_z(t)$ if $z \notin \bar{Z}(t)$. Therefore $u_{i0} > \Delta_z(t) + u_{it}$, and $T_i(z)$ cannot be t. □

Proof of Lemma 1. The lemma is proved in the main text. □

Proof of Proposition 2. Take any observation i and an instrument value $z \in \mathcal{Z}$. The treatment $T_i(z)$ must maximize $(U_z(t) + u_{it})$ over $t \in \mathcal{T}$. Under Assumption 7, for any t we have

- $U_z(t) = U_z(0) + \bar{\Delta}_t$ if $t \in \bar{T}(z)$
- $U_z(t) = U_z(0) + \underline{\Delta}_t$ otherwise.

Therefore, eliminating $U_z(0)$,

$$T_i(z) \in \arg\max \left(\max_{t \notin \bar{T}(z)} (\underline{\Delta}_t + u_{it}), \max_{t \in \bar{T}(z)} (\bar{\Delta}_t + u_{it}) \right). \tag{A.1}$$

Since $\bar{\Delta}_t \geqslant \underline{\Delta}_t$ for all $t \in \mathcal{T}$, a fortiori $\bar{\Delta}_t + u_{it} \geqslant \underline{\Delta}_t + u_{it}$ when $t \in \bar{T}(z)$. As a consequence, we can rewrite Equation (A.1) as

$$T_i(z) \in \arg\max \left(\Delta_i^*,\ V_i^*(z)\right).$$

(i) If $z \in \mathcal{Z}^*$, then $\bar{T}(z)$ is not empty and the maximizer can be either in τ_i^* or in $T_i^*(z)$.

(ii) If $z \in \mathcal{Z} \setminus \mathcal{Z}^*$, then z can only be 0. $\bar{T}(0) = \varnothing$ and $T_i(0)$ can only be in τ_i^*.

□

Proof of Proposition 3. Take an observation i and define $A_i = \{z \in \mathcal{Z}^* \mid T_i(z) = T_i^*(z)\}$.

(i) By definition, $A_i \subset \mathcal{Z}^*$; therefore $A_i = \mathcal{Z}$ (which defines the subpopulation $\mathcal{P}_1$) requires $\mathcal{Z} = \mathcal{Z}^*$.

(ii) Now suppose that $A_i \neq \mathcal{Z}$. If $z \in \mathcal{Z}^* \setminus A_i$, then by construction $T_i(z) \neq T_i^*(z)$. By Proposition 2(i), $T_i(z)$ can only be τ_i^*. If $z \notin \mathcal{Z}^*$, then $z = 0$ and we know from Proposition 2(ii) that $T_i(0) = \tau_i^*$.

(iii) Assume that $\tau_i^* = \tau \in \mathcal{T}^*$. Then $\bar{Z}(\tau) \neq \varnothing$. For any z in $\bar{Z}(\tau)$,

$$V_i^*(z) \geqslant \bar{\Delta}_\tau + u_{i\tau} > \underline{\Delta}_\tau + u_{i\tau} = \Delta_i^*;$$

therefore $z \in A_i$. This proves that $\bar{Z}(\tau) \subset A_i$.

□

Proof of Corollary 1. It follows directly from Proposition 3. □

Proof of Proposition 4. First assume that $\mathcal{Z}^* \neq \mathcal{Z}$, so that only the classes in $\mathcal{P}_2$ exist. The set A of Corollary 1 must be a subset of $\mathcal{Z}^*$. For each such subset, τ can take any value in $\mathcal{T} \setminus \mathcal{T}^*$; and if $\tau \in \mathcal{T}^*$ then τ must be in A. Each subset A of $\mathcal{Z}^*$ with a elements therefore allows for $(a + |\mathcal{T}| - |\mathcal{T}^*|)$ values of τ. This gives a total of

$$\sum_{a=0}^{|\mathcal{Z}^*|} \binom{|\mathcal{Z}^*|}{a} (a + |\mathcal{T}| - |\mathcal{T}^*|)$$

response-types in subpopulation $\mathcal{P}_2$. Moreover, we know that $|\mathcal{T}^*| = |\mathcal{Z}^*|$ under one-to-one targeting. Using the identities

$$\sum_{a=0}^{b} \binom{b}{a} = (1+1)^b = 2^b$$
$$\sum_{a=0}^{b} a\binom{b}{a} = b \times \sum_{a=0}^{b-1} \binom{b-1}{a} = b \times 2^{b-1},$$

we obtain a total of $(2|\mathcal{T}| - |\mathcal{Z}^*|) \times 2^{|\mathcal{Z}^*|-1}$ types.

If $\mathcal{Z} = \mathcal{Z}^*$, we must add the one type in $\mathcal{P}_1$. On the other hand, we must subtract the $|\mathcal{T}|$ classes $c(\mathcal{Z}^*, \tau)$ that are ruled out by Corollary 1(iii). □

Proof of Proposition 5. (i) First assume that $\mathcal{Z}^* \neq \mathcal{Z}$, so that the subpopulation $\mathcal{P}_1$ does not exist. There are two ways to obtain $T_i(z) = t$.

- The first one is for i to belong to in any $c(A, t)$ element, with $A \subset \mathcal{Z}^*$ and t constrained to be in A^+. This requires that $z \notin A$. If z is in $\mathcal{Z}^*$, this implies $A \subset \mathcal{Z}^* \backslash \{z\}$. If not, then A can be any subset of $\mathcal{Z}^*$. This gives the first term in (2.1), and (2.2).
- The second way to get $T_i(z) = t$ is if $t = z$, which can only happen if $z \in \mathcal{Z}^*$. Then if $i \in c(A, \tau)$ for any A that contains z and any $\tau \in A^+$, we have $T_i(z) = z$. This gives the second term in (2.1).

(ii) If $\mathcal{Z}^* = \mathcal{Z}$, we only need to add in the subpopulation $\mathcal{P}_1$ if $z = t$, and to delete from the summations the case $A = \mathcal{Z}^* = \mathcal{Z}$. Introducing these changes in (2.1) gives (2.3). Since $\mathcal{Z}^* = \mathcal{Z}$ there is obviously no subcase $z \notin \mathcal{Z}^*$.

□

Proof of Proposition 6. Since $\mathcal{Z}^* = \{1\} \neq \mathcal{Z}$ in Example 8, we apply equations (2.1) and (2.2). With $\mathcal{Z}^* = \{1\}$, we can only have $A = \varnothing$, with $A^+ = \mathcal{T} \backslash \{1\}$, or $A = \{1\}$, with $A^+ = \mathcal{T}$. Equation (2.1) gives

$$P(t|1) = \mathbb{1}(t \neq 1) \Pr(c(\varnothing, t)) + \mathbb{1}(t = 1) \sum_{\tau \in \mathcal{T}} \Pr(c(\{1\}, \tau));$$

and equation (2.2) gives

$$P(t|0) = \mathbb{1}(t \neq 1) \Pr(c(\varnothing, t)) + \Pr(c(\{1\}, 1)).$$

We already know that $c(\varnothing, t)$ is A_t and $c(\{1\}, \tau)$ is A_1 if $\tau = 1$ and the complier group $C_{\tau 1}$ otherwise. Therefore for $t = 1$

$$P(1|1) = \Pr(A_1) + \sum_{\tau \neq 1} \Pr(C_{\tau 1})$$

and $P(1|0) = \Pr(A_1)$; while for $t \neq 1$ we have $P(t|1) = \Pr(A_t)$ and $P(t|0) = \Pr(C_{t1}) + \Pr(A_t)$. $\square$

Proof of Proposition 8. It is straightforward from Figure 6. $\square$

Proof of Proposition 7. It is straightforward from Figure 5. $\square$

Proof of Lemma 2. We start from the sum over all response groups:

$$\bar{E}_z(t) = \sum_C E_z(t|C) \Pr(i \in C).$$

First note that if group C does not have treatment t under instrument z, it should not figure in the sum. Now if $C_{(z)} = t$, we have

$$\begin{aligned} E_z(t|C) &= \mathbb{E}(Y_i \mathbb{1}(T_i = t) | Z_i = z, i \in C) \\ &= \mathbb{E}(Y_i(t) | Z_i = z, i \in C) \\ &= \mathbb{E}(Y_i(t) | i \in C). \end{aligned}$$

The second part of the Lemma is just adding up. $\square$

Proof of Proposition 9. By Lemma 2, we get

$$\begin{aligned} \bar{E}_0(1) &= \mathbb{E}\left[Y_i(1) | i \in A_1\right] \Pr(i \in A_1) \\ \bar{E}_0(t) &= \mathbb{E}\left[Y_i(t) | i \in A_t\right] \Pr(i \in A_t) \\ &\quad + \mathbb{E}\left[Y_i(t) | i \in C_{t1}\right] \Pr(i \in C_{t1}) \text{ for } t \neq 1, \\ \bar{E}_1(1) &= \mathbb{E}\left[Y_i(1) | i \in A_1\right] \Pr(i \in A_1) \\ &\quad + \sum_{t \neq 1} \mathbb{E}\left[Y_i(1) | i \in C_{t1}\right] \Pr(i \in C_{t1}), \\ \bar{E}_1(t) &= \mathbb{E}\left[Y_i(t) | i \in A_t\right] \Pr(i \in A_t) \text{ for } t \neq 1. \end{aligned}$$

Since Proposition 6 identifies all type probabilities, the first and fourth equations give directly $\mathbb{E}(Y_i(t) | i \in A_t)$ for all t. Then the second equation identifies $\mathbb{E}(Y_i(t) | i \in C_{t1})$ for $t \neq 1$.

However, the values $\mathbb{E}(Y_i(1)|i \in C_{t1})$ for $t \neq 1$ are only linked by

$$\bar{E}_1(1) - \bar{E}_0(1) = \sum_{t \neq 1} \mathbb{E}\left[Y_i(1)|i \in C_{t1}\right] \Pr(i \in C_{t1}).$$

By subtraction, we obtain

$$\begin{aligned}
&(\bar{E}_1(1) - \bar{E}_0(1)) - \sum_{t \neq 1}(\bar{E}_0(t) - \bar{E}_1(t)) \\
&= \sum_{t \neq 1} \mathbb{E}\left[Y_i(1) - Y_i(t)|i \in C_{t1}\right] \Pr(i \in C_{t1}).
\end{aligned}$$

Combining these results with Proposition 6 and Lemma 2 yields the formula in the Proposition. The denominator

$$\sum_{t \neq 1}(P(t|0) - P(t|1)) = P(1|1) - P(1|0)$$

is positive, since all terms in the sum are positive. It follows that all α_t weights are positive and sum to 1. $\square$

Proof of Corollary 2. The corollary follows directly from the proof of Proposition 9, as

$$\sum_{t \neq 1} \Pr(i \in C_{t1}) = \sum_{t \neq 1}(P(t|0) - P(t|1)) = P(1|1) - P(1|0)$$

gives $\mathbb{E}(Y_i(1)|i \in C_{t1}) = (\bar{E}_1(1) - \bar{E}_0(1))/(P(1|1) - P(1|0))$. $\square$

Proof of Proposition 10. For $z = 0, 1, 2$,

$$\begin{aligned}
&\mathbb{E}\left[Y_i|Z_i = z\right] \\
&= \sum_{t=0}^{2} \mathbb{E}\left[Y_i|Z_i = z, i \in A_t\right] \Pr(i \in A_t) + \mathbb{E}\left[Y_i|Z_i = z, i \in C_{112}\right] \Pr(i \in C_{112}) \\
&+ \mathbb{E}\left[Y_i|Z_i = z, i \in C_{212}\right] \Pr(i \in C_{212}) + \mathbb{E}\left[Y_i|Z_i = z, i \in C_{010}\right] \Pr(i \in C_{010}) \\
&+ \mathbb{E}\left[Y_i|Z_i = z, i \in C_{002}\right] \Pr(i \in C_{002}) + \mathbb{E}\left[Y_i|Z_i = z, i \in C_{012}\right] \Pr(i \in C_{012}).
\end{aligned}$$

Note that the first term is also

$$\sum_{t=0}^{2} \mathbb{E}\left[Y_i(t)|i \in A_t\right] \Pr(i \in A_t),$$

which does not depend on z. It follows that

$$
\begin{aligned}
&\mathbb{E}\left[Y_i|Z_i = 0\right] - \mathbb{E}\left[Y_i|Z_i = 1\right] \\
&= \mathbb{E}\left[Y_i(2) - Y_i(1)|i \in C_{212}\right] \Pr(i \in C_{212}) + \mathbb{E}\left[Y_i(0) - Y_i(1)|i \in C_{01*}\right] \Pr(i \in C_{01*}), \\
&\mathbb{E}\left[Y_i|Z_i = 0\right] - \mathbb{E}\left[Y_i|Z_i = 2\right] \\
&= \mathbb{E}\left[Y_i(1) - Y_i(2)|i \in C_{112}\right] \Pr(i \in C_{112}) + \mathbb{E}\left[Y_i(0) - Y_i(2)|i \in C_{0*2}\right] \Pr(i \in C_{0*2}).
\end{aligned}
$$

Combining these formulæ with Proposition 7 yields the result. □

B Proofs for Section 3

Proof of Proposition 11. (i) It follows directly from Proposition 6 and from the mapping of types.

(ii) From Proposition 9, we have

$$\mathbb{E}(Y_i^D(1)|i \in A_1^D) = \mathbb{E}(Y_i^T(1)|i \in A_1^T) = \frac{\bar{E}_0^T(1)}{P^T(1|0)} = \frac{\bar{E}_0^D(1)}{P^D(1|0)}.$$

Moreover,

$$
\begin{aligned}
\mathbb{E}(Y_i^D(0)|i \in A_0^D) &= \mathbb{E}(Y_i^D(0)|i \in \bigcup_{t\neq 1} A_t^T) \\
&= \sum_{t\neq 1} \mathbb{E}(Y_i^T(t)|i \in A_t^T) \frac{P^T(t|1)}{1 - P^T(1|1)} \\
&= \sum_{t\neq 1} \frac{\bar{E}_1^T(t)}{1 - P^T(1|1)} \\
&= \frac{\bar{E}_1^D(0)}{1 - P^D(1|1)}.
\end{aligned}
$$

(iii) Now consider the weighted LATE $\sum_{t\neq 1} \alpha_t^T \mathbb{E}(Y_i^T(1) - Y_i^T(t)|i \in C_{t1}^T)$, which is identified in the unfiltered treatment model (equation 2.11). The weights $\alpha_t^T = (P^T(t|0) - P^T(t|1))/(P^T(1|1) - P^T(1|0))$ are not identified any more. Note however that for any variable W_i,

$$\sum_{t\neq 1} \alpha_t^T \mathbb{E}(W_i|i \in C_{t1}) = \mathbb{E}(W_i|i \in C_{01}^D);$$

therefore $\sum_{t\neq 1} \alpha_t^T \mathbb{E}(Y_i^D(1)|i \in C_{t1}^T) = \mathbb{E}(Y_i^D(1)|i \in C_{01}^D)$. The LHS of Equation (2.11) becomes

$$\mathbb{E}(Y_i^D(1)|i \in C_{01}^D) - \sum_{t \neq 1} \alpha_t^T \mathbb{E}(Y_i^T(t)|i \in C_{t1}).$$

On the RHS we had

$$\frac{(\bar{E}_1^T(1) - \bar{E}_0^T(1)) - \sum_{t \neq 1}(\bar{E}_0^T(t) - \bar{E}_1^T(t))}{P^T(1|1) - P^T(1|0)}.$$

The denominator is still identified as $P^D(1|1) - P^D(1|0)$, as is the first term of the numerator, which equals $\bar{E}_1^D(1) - \bar{E}_0^D(1)$. From equation 3.3,

$$\sum_{t \neq 1}(\bar{E}_0^T(t) - \bar{E}_1^T(t)) = \bar{E}_1^D(0).$$

Therefore we identify

$$\mathbb{E}(Y_i^D(1)|i \in C_{01}^D) - \sum_{t \neq 1} \alpha_t^T \mathbb{E}(Y_i^T(t)|i \in C_{t1}^T) = \frac{(\bar{E}_1^D(1) - \bar{E}_0^D(1)) - (\bar{E}_0^D(0) - \bar{E}_1^D(0))}{P^D(1|1) - P^D(1|0)},$$

which is the standard Wald estimator.

□

Proof of Corollary 3. It is obvious by direct substitution into Equation (3.4). □

Proof of Proposition 12. (i) It follows directly from the mapping of groups.

(ii) Part (i) identifies the weight $\alpha_0^T = (P^D(0|0) - P^D(0|1))/(P^D(1|1) - P^D(1|0))$, which we denote α_0^D in the Proposition. The other terms obtain by simple factorization, with $1 - \alpha_D^0 = \Pr(i \in C_{t1}^T | t > 1)$.

□

Proof of Proposition 13. Recall that Table 2 shows which groups take $D_i = d$ when $Z_i = z$.

(i) We have $P^D(0|z) = P^T(0|z)$ for $z = 0, 1$. Given Proposition 7(i), this gives us $\Pr(C_{010}^T) + \Pr(C_{012}^T) = P^D(0|0) - P^D(0|1)$ and $\Pr(C_{002}^T) + \Pr(C_{012}^T) = P^D(0|0) - P^D(0|2)$, which map into

$$\Pr(C_{010}^D) + \Pr(C_{011}^D) = P^D(0|0) - P^D(0|1)$$
$$\Pr(C_{001}^D) + \Pr(C_{011}^D) = P^D(0|0) - P^D(0|2);$$

and the last equation in Proposition 7(i) maps into

$$\Pr(C^D_{001}) + \Pr(C^D_{010}) + \Pr(C^D_{011}) + \Pr(A^D_0) = P^D(0|0).$$

Finally, $\Pr(A^D_1) = P^D(1|0)$ from the table. Defining $p = \Pr(C^D_{011})$ gives the equations in the proposition, along with the constraints on p. Note also that $\Pr(C^D_{0*0}) = \Pr(A^D_0) + \Pr(C^D_{010}) = P^D(0|2)$.

(ii) First note that $\bar{E}^D_0(1) = \mathbb{E}(Y_i \mathbb{1}(i \in A^D_1)) = \mathbb{E}(Y^D_i(1)|i \in A^D_1)\Pr(A^D_1)$. The other equations can be read from the table:

$$\begin{aligned}
\bar{E}^D_1(0) &= \mathbb{E}(Y^D(0)\mathbb{1}(C^D_{00*})) \\
\bar{E}^D_2(0) &= \mathbb{E}(Y^D(0)\mathbb{1}(C^D_{0*0})) \\
\bar{E}^D_0(0) &= \mathbb{E}(Y^D(0)\mathbb{1}(C^D_{0**})) \\
\bar{E}^D_1(1) &= \mathbb{E}(Y^D(1)\mathbb{1}(C^D_{*1*})) \\
\bar{E}^D_2(1) &= \mathbb{E}(Y^D(1)\mathbb{1}(C^D_{**1})).
\end{aligned}$$

Part (i) showed that we point-identify $\Pr(A^D_1)$, $\Pr(C^D_{01*})$, $\Pr(C^D_{0*1})$, and $\Pr(C^D_{00*})$. This allows us to rewrite the last three lines as

$$\begin{aligned}
\bar{E}^D_0(0) &= P^D(0|1)\,\mathbb{E}(Y^D(0)|C^D_{00*}) + (P^D(0|0) - P^D(0|1))\,\mathbb{E}(Y^D(0)|C^D_{01*}) \\
&= P^D(0|2)\,\mathbb{E}(Y^D(0)|C^D_{0*0}) + (P^D(0|0) - P^D(0|2))\,\mathbb{E}(Y^D(0)|C^D_{0*1}) \\
\bar{E}^D_1(1) &= (P^D(0|0) - P^D(0|1))\,\mathbb{E}(Y^D(1)|C^D_{01*}) + P^D(1|0)\,\mathbb{E}(Y^D(1)|A^D_1) \\
\bar{E}^D_2(1) &= (P^D(0|0) - P^D(0|2))\,\mathbb{E}(Y^D(1)|C^D_{0*1}) + P^D(1|0)\,\mathbb{E}(Y^D(1)|A^D_1),
\end{aligned}$$

where we used the fact that $C^D_{1*1} = C^D_{11*} = A^D_1$.

Simple calculations give

$$\begin{aligned}
\mathbb{E}(Y^D(0)|C^D_{00*}) &= \frac{\bar{E}^D_1(0)}{\Pr(C^D_{00*})} = \frac{\bar{E}^D_1(0)}{P^D(0|1)} \\
\mathbb{E}(Y^D(0)|C^D_{0*0}) &= \frac{\bar{E}^D_2(0)}{\Pr(C^D_{0*0})} = \frac{\bar{E}^D_2(0)}{P^D(0|2)} \\
\mathbb{E}(Y^D(1)|C^D_{01*}) &= \frac{\bar{E}^D_1(1) - \bar{E}^D_0(1)}{P^D(0|0) - P^D(0|1)} \\
\mathbb{E}(Y^D(1)|C^D_{0*1}) &= \frac{\bar{E}^D_2(1) - \bar{E}^D_0(1)}{P^D(0|0) - P^D(0|2)}.
\end{aligned}$$

Finally,

$$
\begin{aligned}
\mathbb{E}(Y^D(0)|C^D_{01*}) &= \frac{\bar{E}^D_1(1) - \mathbb{E}(Y^D(0)\mathbb{1}(C^D_{00*}))}{P^D(0|0) - P^D(0|1)} = \frac{\bar{E}^D_1(1) - \bar{E}^D_1(0)}{P^D(0|0) - P^D(0|1)} \\
\mathbb{E}(Y^D(0)|C^D_{0*1}) &= \frac{\bar{E}^D_2(1) - \mathbb{E}(Y^D(0)\mathbb{1}(C^D_{00*}))}{P^D(0|0) - P^D(0|1)} = \frac{\bar{E}^D_0(0) - \bar{E}^D_1(0)}{P^D(0|0) - P^D(0|1)}
\end{aligned}
$$

(iii) From (ii) we obtain directly, using Lemma 2,

$$
\begin{aligned}
\mathbb{E}(Y^D(1) - Y^D(0)|C^D_{01*}) &= \frac{\bar{E}^D_1(1) + \bar{E}^D_1(0) - \bar{E}^D_0(1) - \bar{E}^D_0(0)}{P^D(0|0) - P^D(0|1)} = \frac{\mathbb{E}(Y|Z=1) - \mathbb{E}(Y|Z=0)}{P^D(0|0) - P^D(0|1)} \\
\mathbb{E}(Y^D(1) - Y^D(0)|C^D_{0*1}) &= \frac{\bar{E}^D_2(1) + \bar{E}^D_2(0) - \bar{E}^D_0(1) - \bar{E}^D_0(0)}{P^D(0|0) - P^D(0|2)} = \frac{\mathbb{E}(Y|Z=2) - \mathbb{E}(Y|Z=0)}{P^D(0|0) - P^D(0|2)}.
\end{aligned}
$$

□

Proof of Proposition 14. Recall that Table 5 shows how response groups map instrument values into filtered treatment values. The proof follows directly. □

Proof of Proposition 15. The proof is omitted since it is similar to those of Propositions 13 and 14. □

Proof of Lemma 3. Using the fact that $\Pr(i \in C^D_{0001}|X_i = x) = \text{DiD}_D(x)$, we have that

$$
\begin{aligned}
\theta_0 &= \int \mathbb{E}[Y_i(1) - Y_i(0)|i \in C^D_{0001}, X_i = x] f(x|i \in C^D_{0001}) dx \\
&= \int \mathbb{E}[Y_i(1) - Y_i(0)|i \in C^D_{0001}, X_i = x] \frac{\Pr(i \in C^D_{0001}|X_i = x) f(x)}{\Pr(i \in C^D_{0001})} dx \\
&= \int \frac{\text{DiD}_Y(x)}{\text{DiD}_D(x)} \frac{\Pr(i \in C^D_{0001}|X_i = x) f(x)}{\Pr(i \in C^D_{0001})} dx \\
&= \frac{\int \text{DiD}_Y(x) f(x) dx}{\Pr(i \in C^D_{0001})} \\
&= \frac{\int \text{DiD}_Y(x) f(x) dx}{\int \text{DiD}_D(x) f(x) dx},
\end{aligned}
$$

which proves the lemma. □

C Additional Material

C.1 Strict Targeting in the Ternary/ternary Model

Just like ours, Kirkeboen, Leuven, and Mogstad (2016)'s approach to identification relies on a monotonicity assumption and a restriction on the mapping from instruments to treatments. We translate them here in our notation to show that in this model, their assumptions are equivalent to ours.

Kirkeboen, Leuven, and Mogstad (2016) impose the following in their Assumption 4:

- if $T_i(0) = 1$ then $T_i(1) = 1$
- if $T_i(0) = 2$ then $T_i(2) = 2$.

This can be viewed as a monotonicity assumption. It excludes the twelve response groups $C_{10*}, C_{12*}, C_{2*0}$, and C_{2*1}.

Their Proposition 2 proves point-identification of response-groups when one of three alternative assumptions is added to their Assumption 4. We focus here on their assumption (iii), which is the weakest of the three and the one their application relies on. In our notation, it states that:

- if $(T_i(0) \neq 1$ and $T_i(1) \neq 1)$, then $(T_i(0) = 2$ iff $T_i(1) = 2)$
- if $(T_i(0) \neq 2$ and $T_i(2) \neq 2)$, then $(T_i(0) = 1$ iff $T_i(2) = 1)$.

These complicated statements can be simplified. Take the first part. If both $T_i(0)$ and $T_i(1)$ are not 1, then they can only be 0 or 2. Therefore we are requiring $T_i(0) = T_i(1)$. Applying the same argument to the second part, Assumption (iii) becomes:

- if $(T_i(0) \neq 1$ and $T_i(1) \neq 1)$, then $T_i(0) = T_i(1)$
- if $(T_i(0) \neq 2$ and $T_i(2) \neq 2)$, then $T_i(0) = T_i(2)$.

It therefore excludes the response-groups $C_{02*}, C_{20*}, C_{0*1}$, and C_{1*0}. The response-group C_{021} appears twice in this list; and four other response-groups were already ruled out by Assumption 4. The reader can easily check that the $3^3 - 12 - (11 - 4) = 8$ response-groups left are exactly the same as in our Figure 5.

C.2 A Variant of Filtered Factorial Design

Let us return to the factorial design of Example 4, with a twist: the unfiltered treatment consists of the full ranking of the three alternatives. The instrument values are still $(0 \times 0, 0 \times 1, 1 \times 0, 1 \times 1)$; now T_i is a pair that consists of the most-preferred alternative

$$\overline{T}_i(z) = \arg\max_{t=0,1,2} (U_z(t) + u_{it})$$

and of the least-preferred alternative

$$\underline{T}_i(z) = \arg\min_{t=0,1,2} (U_z(t) + u_{it}).$$

In Section 3.2, we considered the case when T_i is only $\overline{T}_i$; and we added a filter $D_i = \mathbb{1}(\overline{T}_i > 0)$. Let us now take $T_i = (\overline{T}_i, \underline{T}_i)$, with the filter $D_i = \mathbb{1}(\underline{T}_i = 0)$.

The model of Section 3.2, where we only observed whether the most-preferred alternative was 0, led to a double hurdle model. In this variant, we only observe whether the *least*-preferred alternative is 0, which leads to a different filtered treatment model:

$$D_i(z) = \mathbb{1}(u_{i0} < \min(\Delta_z^T(1) + u_{i1}, \Delta_z^T(2) + u_{i2})). \tag{C.1}$$

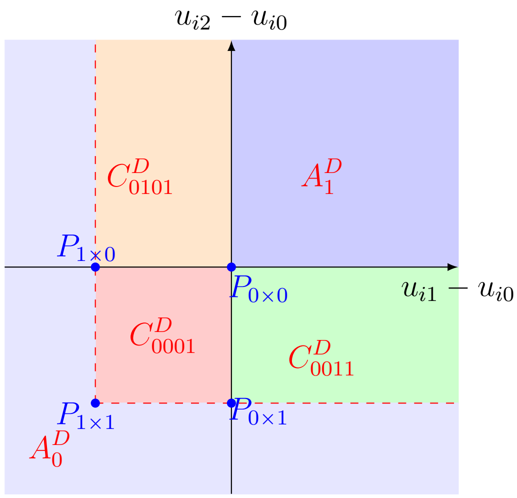
We keep the same constraints on the mean utilities as in (3.8). Under Equation (C.1), we have five response groups, as shown in Figure 10. Table 5 shows how response groups map instrument values into filtered treatment values.

Table 5: D-response Groups for the Alternative Factorial Design Model

	$D_i(z) = 0$	$D_i(z) = 1$
$z = 0 \times 0$	$A_0^D \bigcup C_{0011}^D \bigcup C_{0101}^D \bigcup C_{0111}^D$	A_1^D
$z = 1 \times 0$	$A_0^D \bigcup C_{0011}^D$	$A_1^D \bigcup C_{0101}^D \bigcup C_{0111}^D$
$z = 0 \times 1$	$A_0^D \bigcup C_{0101}^D$	$A_1^D \bigcup C_{0011}^D \bigcup C_{0111}^D$
$z = 1 \times 1$	A_0^D	$A_1^D \bigcup C_{0011}^D \bigcup C_{0101}^D \bigcup C_{0111}^D$

Proposition 15 (Identifying the Model with Equation (C.1)). *(i) The probabilities of the*

Figure 10: Response Groups under Equation (C.1)



D-response groups are point-identified by

$$\begin{aligned}
\Pr(A^D_0) &= P^D(1|0\times 0)\\
\Pr(A^D_1) &= P^D(0|1\times 1)\\
\Pr(C^D_{0011}) &= P^D(0|0\times 0) - P^D(0|0\times 1)\\
\Pr(C^D_{0101}) &= P^D(0|0\times 0) - P^D(0|1\times 0)\\
\Pr(C^D_{0111}) &= P^D(0|1\times 0) + P^D(0|1\times 0) - P^D(0|0\times 0) - P^D(0|1\times 1),
\end{aligned}$$

and the model has three testable implications:

$$\begin{aligned}
P^D(0|0\times 0) &\geqslant P^D(0|0\times 1),\\
P^D(0|0\times 0) &\geqslant P^D(0|1\times 0),\\
P^D(0|1\times 0) + P^D(0|0\times 1) &\geqslant P^D(0|0\times 0) - P^D(0|1\times 1).
\end{aligned}$$

(ii) The LATEs on the three groups of compliers are point-identified by

$$\begin{aligned}
&\mathbb{E}(Y^D_i(1) - Y^D_i(0)|i \in C^D_{0101}) = \frac{\mathbb{E}(Y|Z = 0\times 1) - \mathbb{E}(Y|Z = 0\times 0)}{P^D(1|0\times 1) - P^D(1|0\times 0)}\\
&\mathbb{E}(Y^D_i(1) - Y^D_i(0)|i \in C^D_{0011}) = \frac{\mathbb{E}(Y|Z = 1\times 0) - \mathbb{E}(Y|Z = 0\times 0)}{P^D(1|1\times 0) - P^D(1|0\times 0)}\\
&\mathbb{E}(Y^D_i(1) - Y^D_i(0)|i \in C^D_{0001}) =\\
&\frac{\mathbb{E}(Y|Z = 0\times 0) + \mathbb{E}(Y|Z = 1\times 1) - \mathbb{E}(Y|Z = 1\times 0) - \mathbb{E}(Y|Z = 0\times 1)}{P^D(1|0\times 0) + P^D(1|1\times 1) - P^D(1|1\times 0) - P^D(1|0\times 1)}.
\end{aligned}$$

It is worth comparing Proposition 14 with Proposition 15. One interesting observation is that the share of C^D_{0111} in Proposition 14 is identical up to sign to that of C^D_{0001} in Proposition 15. Namely, $\Pr(C^D_{0111}) = \text{DiD}(T)$ and $\Pr(C^D_{0001}) = 0$ in Proposition 14; $\Pr(C^D_{0111}) = 0$ and $\Pr(C^D_{0001}) = -\text{DiD}(T)$ in Proposition 15, where $\text{DiD}(T)$ is the difference-in-differences of the propensity score defined by:

$$\begin{aligned}\text{DiD}(T) = &\{\Pr[D_i = 1|Z_i = 1 \times 0] - \Pr[D_i = 1|Z_i = 0 \times 0]\} \\ &- \{\Pr[D_i = 1|Z_i = 1 \times 1] - \Pr[D_i = 1|Z_i = 0 \times 1]\}.\end{aligned}$$

In terms of economic interpretation, one may think of the selection mechanism in Proposition 14 as the scenario when instruments $\mathbb{1}(Z_i = 1 \times 0)$ and $\mathbb{1}(Z_i = 0 \times 1)$ are substitutes to encourage agents to take treatment. On the contrary, the selection mechanism in Proposition 15 corresponds to the case that instruments $\mathbb{1}(Z_i = 1 \times 0)$ and $\mathbb{1}(Z_i = 0 \times 1)$ are complements. The same estimands identify the average treatment effects for conceptually distinct subpopulations, depending on the details of the selection mechanism. This suggests that it is important to learn about the nature of selection into treatment before interpreting the causal parameters of different compliers. To do this, one can estimate the difference-in-differences of the propensity score $\text{DiD}(T)$ and use its sign to determine whether equation (3.9) or equation (C.1) is more plausible in any particular application.

References

Ackerberg, D., X. Chen, and J. Hahn (2012): "A practical asymptotic variance estimator for two-step semiparametric estimators," *Review of Economics and Statistics*, 94(2), 481–498.

Ackerberg, D., X. Chen, J. Hahn, and Z. Liao (2014): "Asymptotic efficiency of semiparametric two-step GMM," *Review of Economic Studies*, 81(3), 919–943.

Angrist, J., D. Lang, and P. Oreopoulos (2009): "Incentives and Services for College Achievement: Evidence from a Randomized Trial," *American Economic Journal: Applied Economics*, 1(1), 136–63.

Angrist, J. D., and G. W. Imbens (1995): "Two-stage least squares estimation of average causal effects in models with variable treatment intensity," *Journal of the American Statistical Association*, 90(430), 431–442.

Ao, W., S. Calonico, and Y.-Y. Lee (2019): "Multivalued Treatments and Decomposition Analysis: An Application to the WIA Program," *Journal of Business & Economic Statistics*, in press.

Bloom, H. S. (1984): "Accounting for no-shows in experimental evaluation designs," *Evaluation Review*, 8(2), 225–246.

Caetano, C., and J. C. Escanciano (2020): "Identifying Multiple Marginal Effects with a Single Instrument," *Econometric Theory*, forthcoming.

Cattaneo, M. D. (2010): "Efficient semiparametric estimation of multi-valued treatment effects under ignorability," *Journal of Econometrics*, 155(2), 138–154.

D'Haultfoeuille, X., and P. Février (2015): "Identification of Nonseparable Triangular Models With Discrete Instruments," *Econometrica*, 83(3), 1199–1210.

Feng, J. (2020): "Matching Points: Supplementing Instruments with Covariates in Triangular Models," Job market paper, available at `https://econ.columbia.edu/e/jfeng/`.

Frölich, M. (2007): "Nonparametric IV estimation of local average treatment effects with covariates," *Journal of Econometrics*, 139(1), 35–75.

Goff, L. (2020): "A Vector Monotonicity Assumption for Multiple Instruments," available at `http://www.columbia.edu/~ltg2111/`.

Heckman, J., and R. Pinto (2018): "Unordered Monotonicity," *Econometrica*, 86(1), 1–35.

Heckman, J. J. (1979): "Sample Selection Bias as a Specification Error," *Econometrica*, 47(1), 153–161.

Heckman, J. J., S. Urzua, and E. Vytlacil (2006): "Understanding instrumental variables in models with essential heterogeneity," *Review of Economics and Statistics*, 88(3), 389–432.

——— (2008): "Instrumental variables in models with multiple outcomes: The general unordered case," *Annales d'économie et de statistique*, 91/92, 151–174.

Huang, L., U. Khalil, and N. Yildiz (2019): "Identification and estimation of a triangular model with multiple endogenous variables and insufficiently many instrumental variables," *Journal of Econometrics*, 208(2), 346–366.

Imbens, G. W. (2000): "The role of the propensity score in estimating dose-response functions," *Biometrika*, 87(3), 706–710.

Imbens, G. W., and J. D. Angrist (1994): "Identification and Estimation of Local Average Treatment Effects," *Econometrica*, 62(2), 467–475.

Kamat, V. (2019): "Identification with Latent Choice Sets," arXiv:1711.02048, `https://arxiv.org/abs/1711.02048`.

Kirkeboen, L. J., E. Leuven, and M. Mogstad (2016): "Field of study, earnings, and self-selection," *Quarterly Journal of Economics*, 131(3), 1057–1111.

Kline, P., and C. R. Walters (2016): "Evaluating public programs with close substitutes: The case of Head Start," *Quarterly Journal of Economics*, 131(4), 1795–1848.

Lee, S., and B. Salanié (2018): "Identifying effects of multivalued treatments," *Econometrica*, 86(6), 1939–1963.

Mogstad, M., A. Torgovitsky, and C. R. Walters (2019): "Identification of Causal Effects with Multiple Instruments: Problems and Some Solutions," Working Paper 25691, National Bureau of Economic Research.

——— (2020): "Policy Evaluation With Multiple Instrumental Variables," Working Paper 27546, National Bureau of Economic Research.

Mountjoy, J. (2019): "Community Colleges and Upward Mobility," Chicago Booth mimeo.

Muralidharan, K., M. Romero, and K. Wüthrich (2019): "Factorial Designs, Model Selection, and (Incorrect) Inference in Randomized Experiments," Working Paper 26562, National Bureau of Economic Research.

Torgovitsky, A. (2015): "Identification of Nonseparable Models Using Instruments With Small Support," *Econometrica*, 83(3), 1185–1197.

Vytlacil, E. (2002): "Independence, monotonicity, and latent index models: An equivalence result," *Econometrica*, 70(1), 331–341.