WORKING PAPERS

N° TSE –1031

August 2019

“Identification with Latent Choice Sets”

Vishal Kamat

Toulouse School of Economics

Identification with Latent Choice Sets*

Vishal Kamat
Toulouse School of Economics
University of Toulouse Capitole
vishal.kamat@tse-fr.eu

August 8, 2019

Abstract

In a common experimental format, individuals are randomly assigned to either a treatment group with access to a program or a control group without access. In such experiments, analyzing the average effects of the treatment of program access may be hindered by the problem that some control individuals do not comply with their assigned status and receive program access from outside the experiment. Available tools to account for such a problem typically require the researcher to observe the receipt of program access for every individual. However, in many experiments, this is not the case as data is not collected on where any individual received access. In this paper, I develop a framework to show how data on only each individual's treatment assignment status, program participation decision and outcome can be exploited to learn about the average effects of program access. I propose a nonparametric selection model with latent choice sets to relate where access was received to the treatment assignment status, participation decision and outcome, and a linear programming procedure to compute the identified set for parameters evaluating the average effects of program access in this model. I illustrate the framework by analyzing the average effects of Head Start preschool access using the Head Start Impact Study. I find that the provision of Head Start access induces parents to enroll their child into Head Start and also positively impacts test scores, and that these effects heterogeneously depend on the availability of access to an alternative preschool.

KEYWORDS: Program evaluation, latent choice sets, unobserved treatment, program access, multiple treatments, average treatment effect, noncompliance, discrete choice, partial identification, social experiments, head start impact study.

JEL classification codes: C14, C31.

*I am extremely grateful to Ivan Canay, Chuck Manski and Alex Torgovitsky for their extensive guidance and feedback. I also thank Eric Auerbach, Gideon Bornstein, Pedro Carneiro, Joel Horowitz, Tom Meling, Magne Mogstad, Francesca Molinari, Sam Norris, Matt Notowidigdo, Nicolas Inostroza, Rob Porter, Pedro Sant'Anna, Lola Segura, Azeem Shaikh, Max Tabord-Meehan and conference and seminar participants at several institutions for useful comments, and Research Connections for providing data on the Head Start Impact Study. Funding from ANR under grant ANR-17-EURE-0010 (Investissements d'Avenir program) and the Robert Eisner Memorial Fellowship is gratefully acknowledged.

1 Introduction

Social experiments are a prominent method to evaluate the impact of a program. In one common experimental format, a treatment group is provided access to the program, whereas a control group is not provided access. If assignment to groups is random and no control individual received the treatment of program access, a researcher can evaluate the average effect of program access on a given response by estimating the intention-to-treat (ITT) estimand—the difference in mean responses between treatment and control groups. In addition, the researcher can evaluate the average effect of program access on a given outcome for the subgroup of individuals who in fact participate when access is received by estimating the instrumental variable (IV) estimand—the ratio of the ITT estimand on the outcome to that on program participation.

However, in some experiments, control individuals may not comply with their assigned status and receive program access from outside the experiment. In such cases, as formally illustrated in Section S.1 of the Supplement Appendix, the ITT and IV estimates allow the researcher to only evaluate the so-called local average effects of program access for the subgroup of compliers, i.e. those who comply when assigned to the control group. To move beyond such local effects, several treatment effect tools have been developed—see, for example, Mogstad and Torgovitsky (2018, Section 6) for a recent overview of such tools. While these tools could potentially be applied to evaluate the average effects of program access, their application typically requires the researcher to observe the receipt of program access for every individual. Unfortunately, in many experiments, this is not the case as data is not collected on where any individual received access.

In this paper, I develop a nonparametric framework to show how data on only the treatment assignment status, the program participation decision and the outcome in such experiments can be exploited to learn about the average effects of program access. The framework is based on the observation that while where access was received is not directly observed, it can be partially inferred through the treatment assignment status and participation decision. Specifically, assignment to the treatment group reveals that access to the program was received, whereas participation in the program also reveals that access to the program was received.

To formally exploit such information, I propose a nonparametric model that leverages the insight that receiving access to a given alternative can be framed as receiving that alternative in the choice set from which participation decisions are made. In particular, the model treats the individual's participation decision to be equal to a utility maximization decision given their latent preferences over the various alternatives and their latent choice set of available alternatives. In this model, the information that the treatment assignment status provides can directly be imposed as a restriction on the feasible values that the choice set can take, whereas the information that the observed participation decision provides is indirectly captured through the relationship between the choice set and the decision.

In the context of the model, I study what we can learn about parameters that evaluate the average effect of program access on the decision to participate in the program and on the subsequent outcome using the information provided by the experiment. Formally, these parameters correspond to comparing mean decisions and outcomes under counterfactual choice sets with and without the program. In addition, I study the ratio of these two parameters, which evaluates the average effect of program access on the outcome conditional on the subgroup who in fact participate in the program when access is received.

Given the partial nature of the information provided by the experiment, I find that the parameters are generally partially or set identified. By leveraging results from Charnes and Cooper (1962), I illustrate how a linear programming procedure can be used to tractably characterize these identified sets. The procedure allows additional assumptions to be flexibly imposed on the baseline model, which can potentially shrink the identified set. I discuss several examples of such assumptions that impose nonparametric restrictions on how the choice sets, preferences and outcomes may be related to each other.

Several experiments in economics are implemented using the aforementioned format and face noncompliance and observational problems with respect to access. For example, in the Head Start Impact Study (HSIS), children were randomly provided access to a specific Head Start preschool, but control children could also receive Head Start access by finding Head Start preschools with available slots outside the experiment. In the Oregan Health Insurance Experiment (OHIE), eligible individuals were randomly provided Medicaid access through a subprogram of Oregan's Medicaid program, but control individuals could also receive Medicaid access by being eligible for an alternative subprogram. In a microfinance experiment studied in Angelucci et al. (2015), individuals living in certain villages were randomly provided microfinance access, but individuals living in the control villages could also receive access by having a viable address in a treated village. In each of these experiments, data was collected on each individual's treatment assignment status and participation decision, but not on where they received access.

I illustrate the framework using the HSIS and, for ease of exposition, also develop the framework in the context of this experiment.[1] Applying the analysis to the HSIS data, I find that the identified sets are informative and imply that the effect of receiving Head Start access has more average benefits when access to an alternative non-Head Start preschool is absent. In particular, under the most informative model specification, I find that the provision of Head Start access induces between 85.5% and 91.1% of parents to enroll their three-year-old child into Head Start and improves their child's test score on average between 0.17 and 0.89 standard deviations, when alternative preschool access is absent. In contrast, I find that it only induces between 20.5% and 85.5% of parents to enroll their child into Head Start and only improves their child's test score on average between 0

[1] In Section S.5 of the Supplement Appendix, I present the generalized version of the framework and also show how it applies to the OHIE and the microfinance experiment studied in Angelucci et al. (2015).

and 0.43 standard deviations, when alternative preschool access is present.

The analysis in this paper contributes to several literatures. It is related to a large literature that studies the partial identification of the average treatment effect—see Ho and Rosen (2017, Section 5) for a recent survey. Papers in this literature typically use the observed distribution of the response and the treatment along with additional data and assumptions to partially identify the average treatment effect. In contrast, in the setting studied in this paper, the treatment of interest corresponding to the receipt of program access is not directly observed for any individual. I illustrate how a selection model that relates this treatment to the observed treatment assignment status, participation decision and outcome can be exploited in the analysis. The use of the selection model also makes the analysis distinct from one in alternative settings with unobserved treatments where the researcher directly observes an indicator for whether the treatment is observed or not (Horowitz and Manski, 1998, 2000; Molinari, 2010).

The analysis is also related to a recent literature that analyzes various treatment effect parameters in settings where individuals select into multiple unordered alternatives—see, for example, Heckman et al. (2006, 2008), Heckman and Pinto (2018), Kirkeboen et al. (2016), Kline and Walters (2016) and Lee and Salanié (2018). These papers aim to show how selection models with weak restrictions on unobserved heterogeneity in choice behaviour can be exploited in their respective analyses. However, unlike the model in this paper, the selection models in these papers do not separately model the unobserved heterogeneity that may potentially be present in both preferences and choice sets. Indeed, as we will observe in the analysis, this feature is key to leverage the insight that receiving access to an alternative can be viewed as receiving it in the choice set and, in turn, use a selection model to define and analyze the parameters studied in this paper.

The selection model used in the analysis draws on the literature on nonparametric discrete choice analysis. Specifically, as further elaborated in Section 3.1 below, it exploits elements from the discrete choice model of Manski (2007), and the related model from Marschak (1960), with novel modifications. The model of Manski (2007) was designed to use nonparametric revealed preference arguments to learn about latent preferences from observed choices and exogenous variation in observed choice sets, and then predict choices under counterfactual choice sets. But, as where an individual receives access is unobserved and may depend on their preferences, exogenous variation in observed choice sets is absent in the setting studied in this paper. As a result, I build on this model to allow the choice sets to be unobserved and correlated with preferences.

Since the selection model takes choice sets to be unobserved, the analysis is also related to the literature on analyzing preferences in discrete choice settings with unobserved heterogeneity in choice sets. In such settings, a common approach is to specify a parametric model on the mechanism determining the unobserved choice sets and on how these choice sets are related to the unobserved preferences (Ben-Akiva and Boccara, 1995; Gaynor et al., 2016; Goeree, 2008). In contrast, the analysis in this paper provides an approach to evaluate preferences using only nonparametric revealed preference arguments, which leaves the choice set mechanism and the relation between choice sets and preferences completely unrestricted. The approach complements recent proposals such as, for example, Abaluck and Adams (2018), Barseghyan et al. (2018), Cattaneo et al. (2018) and Crawford et al. (2019), which, under alternative models and settings, aim to develop tools that analyze preferences under weak restrictions on the unobserved heterogeneity in choice sets.

The remainder of the paper is organized as follows. Section 2 briefly describes the setting of the HSIS and motivates the identification problem. Section 3 develops the framework on how to identify the average effects of Head Start access. Section 4 applies the framework to the HSIS data. Section 5 concludes. In the Supplementary Appendix to the paper, I provide auxiliary results pertinent to the analysis along with proofs and additional mathematical derivations for all results. Specifically, Section S.1 expands on the motivation by illustrating what standard approaches can evaluate in the HSIS with respect to the effects of Head Start access and by briefly discussing why we may be interested in such effects. Section S.2 provides additional results and details for the empirical analysis in Section 4. Section S.3 and Section S.4 provide details on how to perform statistical inference and measure the sensitivity of the estimates to the imposed assumptions, respectively. Section S.5 presents the generalized version of the framework and shows how it can be applied to two additional experiments. Section S.6 provides proofs for all results.

2 The Head Start Impact Study

Head Start is the largest early childhood education program in the United States that provides access to free preschool education to three- and four-year-old children from disadvantaged households. Program eligibility for households is primarily determined by the federal poverty line, yet certain exceptions qualify additional low-income-households. In Fall 2002, the Head Start Impact Study (HSIS) was a publicly-funded randomized experiment implemented to evaluate Head Start. Puma et al. (2010) provides further details on the program and experiment. Below, I briefly summarize the details of the experiment that are relevant for the analysis developed in this paper.

The experiment acquired a sample of Head Start preschools and participating children using a multistage stratified sampling scheme. Preschools were sampled from the subpopulation of so-called non-saturated Head Start preschools, which referred to preschools where the number of available slots was strictly smaller than the number of applicants. From each sampled preschool, children were then sampled separately from the subpopulation of three- and four-year-old applicants in Fall 2002, who were not previously enrolled in any Head Start services.

Once these samples were acquired, the randomization was performed. To better understand certain details, it is useful to describe the remainder of the experiment in stages:

Stage 1: At each sampled Head Start preschool, the experiment randomly assigned sampled children to either a treatment group with access to that preschool or a control group without access to that preschool. The experiment did not control the child's chances of receiving access from other preschools, both Head Start preschools from the one the child was not sampled and alternative non-Head Start preschools. In particular, parents could potentially receive access to these preschools depending on whether they could find such preschools that had available slots for their child.

Stage 2: The experiment collected data on the type of the care setting where the parents enrolled their child, which consisted of Head Start and alternative preschools and a home care or non-preschool setting where their child was taken care of either by themselves, a relative or some known individual.

Stage 3: The experiment collected data on a number of outcomes, such as various test scores, at the end of the school year in Spring 2003.

For the purposes of the analysis in this paper, it is important to highlight two relevant features in Stage 1 that complicate evaluating the average effects of Head Start access—see Section S.1 of the Supplement Appendix for a further elaboration and formal illustration on these complications. First, note that children assigned to the control group could choose to not comply with their assigned status and receive access to Head Start from outside the experiment. As a result, in this case, we can formally show that standard intention-to-treat and instrumental variable estimands will allow us to evaluate only the so-called local average effects of Head Start access for the subgroup of compliers, i.e. those who comply with their assigned status and do not receive access to Head Start from outside the experiment.

Second, note that the experiment did not collect data on the set of preschools where any child received access. Instead, it only collected data on the treatment assignment status and the enrollment decision which partially reveal information on where access was received: namely, assignment to the treatment group reveals access to Head Start was received and enrollment in a given preschool reveals access to that preschool was received. As a result, to move beyond the local effects and evaluate the average effects of Head Start access for the whole population, we need to develop treatment effect tools that allow us to account for the fact that the treatment of interest corresponding to the receipt of Head Start access whose effect we want to evaluate is not directly observed, but only partially observed through the treatment assignment status and the enrollment decision for every child.

3 Identification Framework

In this section, I develop a framework to show how to evaluate the average effects of Head Start access using only data on the treatment assignment status, enrollment decision and the child's test score. The framework is based on the insight that receiving access to a preschool can be framed as receiving that preschool in the choice set of care settings from which enrollment decisions are made. As we will observe, this insight will allow us to use a selection model to relate where access was received to the data, and then use the structure of this model to define various parameters evaluating the average effects of Head Start access and exploit the partial information provided by the data to learn about them.

3.1 Model with Latent Choice Sets

I begin by proposing the selection model which is the building the block for the developed identification analysis. In this model, I assume that the set of feasible care settings that the parents can potentially enroll their child in is given by the following finite set

$$\mathcal{D} = \{n, a, h\} \ ,$$

where h denotes a Head Start preschool, a denotes an alternative non-Head Start preschool, and n denotes a no preschool or equivalently a home care setting. For a clearer exposition of the analysis that follows, it is useful to present the model in terms of the various stages of the experiment described in Section 2.

In Stage 1, parents receive access to Head Start and alternative preschools. Receiving access can equivalently be viewed as obtaining that preschool in their choice set of care settings from which enrollment decisions are made. Let C denote the value of this choice set. Based on the preschools to where access was received, the choice set can take values in

$$\mathcal{C} = \{\{n\}, \{n, a\}, \{n, h\}, \{n, a, h\}\} \ ,$$

i.e. the set of all possible choice sets that contain home care. The experiment exogenously altered the choice set by assigning Head Start access to those in the treatment group. Let Z denote the observed indicator for whether the child was assigned to the treatment group or not, and let $C(1)$ and $C(0)$ respectively denote the potential choice sets had the child been assigned to the treatment and control group. The obtained and potential choice sets are related by

$$C = \begin{cases} C(1) & \text{if} \quad Z = 1 \ , \\ C(0) & \text{if} \quad Z = 0 \ . \end{cases}$$

In Stage 2, parents decide where to enroll their child given their obtained choice set from Stage 1. Let D denote the observed enrollment decision. Further, let $U(d)$ denote the parents' indirect utility had their child been enrolled in care setting $d \in \mathcal{D}$. The observed enrollment decision is assumed to be given by the following utility maximization relationship

$$D = \arg\max_{d \in C} U(d) \ . \tag{1}$$

Note that since the utility under each care setting does not possess any cardinal value, different monotonic transformations of these utilities will generate observationally equivalent choices. For the purposes of the analysis, it is therefore more useful to directly refer to the underlying preference type that these utilities represent. To this end, assuming that the parents have a strict preference relation over the set of care settings, let U denote the parents' preference type corresponding to a strict preference relationship over the set of care settings $\mathcal{D}$. As there are three possible care settings, there are six possible preference types denoted by the following set

$$\mathcal{U} = \{(h \succ a \succ n),\ (h \succ n \succ a),\ (a \succ h \succ n),\ (a \succ n \succ h),\ (n \succ h \succ a),\ (n \succ a \succ h)\}\ .$$

Moreover, let $d(u, c)$ denote the known choice function that corresponds to what preference type $u \in \mathcal{U}$ would choose under a non-empty subset $c \subseteq \mathcal{D}$, i.e. under a given feasible choice set. For example, note that

$$d((h \succ a \succ n), \{n, a\}) = a$$

as preference type $(h \succ a \succ n)$ prefers a to n and hence would choose a when faced with the choice set containing n and a. Using the above notation, the utility maximization relationship in (1) can be alternatively re-written in terms of the preference type and the obtained choice set through the following relationship

$$D = \sum_{u \in \mathcal{U}, c \in \mathcal{C}} d(u, c) I\{U = u, C = c\} \equiv d(U, C)\ . \tag{2}$$

In Stage 3, the child's test score is observed under the enrolled care setting from Stage 2. Let Y denote the observed test score and let $Y(d)$ denote the potential test score had the child been enrolled in care setting $d \in \mathcal{D}$. The observed test score is related to the potential test scores and the enrolled care setting through the following relationship

$$Y = \sum_{d \in \mathcal{D}} Y(d) I\{D = d\} \equiv Y(D)\ . \tag{3}$$

In the model described above, note that the only variables that were presumed to be observed for a given child were the following

$$(Y, D, Z)\ .$$

Importantly, as the experiment did not collect data on where the parents received access, the value of the obtained choice set C is considered to be latent or unobserved in Stage 1. However, the structure of the model indirectly provides some partial information on the possible values that the choice set can take. In particular, the structure of the choice equation in (2) reveals that the observed choice must be in the choice set, i.e.

$$D = d \implies C \in \{c \in \mathcal{C} : d \in c\}\ .$$

Similarly, the experiment also provides some information on the values that the choice set can take as it ensured that the treatment group was provided Head Start access. This information can be exploited by imposing restrictions on the model, which I formally state in terms of the following assumption:

Assumption HSIS.

(i) $(Y(n), Y(a), Y(h), U, C(0), C(1)) \perp Z$.

(ii) $h \in C(1)$.

Assumption HSIS(i) states that children were randomly assigned to either the treatment or control group. Assumptions HSIS(ii) states that being assigned to the treatment group guaranteed Head Start access and captures the information that the experiment provides on the values that the choice set can take. In particular, it reveals that Head Start must be in the choice set when assigned to the treatment group, i.e.

$$Z = 1 \implies C \in \{c \in \mathcal{C} : h \in c\} \equiv \{\{n, h\}, \{n, a, h\}\} \ .$$

In the following sections, I illustrate how to exploit this partial information on where access was received and learn about different parameters evaluating the average effects of Head Start access in the context of the model.

Before proceeding, I make several general comments on the above described model. First, note that the model as stated is entirely nonparametric and places no restrictions on the dependence between the variables across the three stages, i.e. the potential test scores, the preference type and the choice sets are allowed to be correlated in an unrestricted manner. As a result, the model allows for settings where parents make enrollment decisions in Stage 2 based on unobservables correlated to potential test scores in Stage 3, and where parents receive preschool access in Stage 1 based on unobservables correlated to preferences in Stage 2 and potential test scores in Stage 3. In Section 3.4, I illustrate examples of several nonparametric assumptions that could additionally be imposed on this baseline model on how these variables may be related to each other.

Second, note that the model is presented in the context of the HSIS and, as a result, is specialized to capture features specific to this context. However, certain elements of the described model are not conceptually essential for the analysis developed in the following sections. In particular, the set of choice alternatives does not require to take only three elements, but can generally consist of any finite number of alternatives. Moreover, the information that an experiment provides on where access is received does not have to satisfy the form of Assumption HSIS(ii), but can potentially take alternative forms which may depend on the specific context of the experiment. In Section S.5 of the Supplement Appendix, I present the generalized version of the model which formalizes the type of extensions that the developed analysis generally allows for. I also present examples of two alternative experiments and illustrate how these extensions to the model allow the model to accommodate certain details specific to the setting of these experiments.

Finally, note that the model draws on certain elements from the literature on discrete choice analysis. In this literature, previous nonparametric analysis have employed models that specify the presence of only a choice set and preference type for each agent and place no parametric assumptions on the distribution of choice sets and preference types. In particular, Marschak (1960) first proposed using a model with latent preference types to analyze whether the distribution of observed choices can be rationlized by a model of utility maximizing agents with strict preferences. Manski (2007) then showed how this analysis can be extended more generally to also predict what agents would chose under counterfactual choice sets by introducing the general structure for a choice model with preference types and choice sets. Indeed, the above described model exploited the structure of this choice model when specifying the enrollment decision in Stage 2.

However, the model from Manski (2007), along with that from Marschak (1960), assumed that the choice sets were observed and statistically independent of preference types which was the only variable taken to be latent. In the setting of the HSIS, these assumptions are not valid as where parents' receive access is not observed and may be correlated to their preferences. As a result, to accommodate for such differences, the above described model made several important modifications to the model from Manski (2007). First, instead of taking the choice sets to be observed and statistically independent of the latent variables, the model allowed them to be latent and correlated in an unrestricted manner with the other latent variables. Second, it assumed, in contrast, that an alternative variable, i.e the treatment assignment status, was observed and statistically independent of the latent variables and which provided partial information on the choice set. Finally, apart from these differences in the structure of the choice model to accommodate the HSIS setting, the model also additionally introduced the presence of a Stage 3 that associated outcomes with each decision in Stage 2. As illustrated in the following section, this additional component to the model allows defining parameters that also study outcomes under different choice sets, and not solely choices which is usually the focus in the literature on discrete choice analysis.

3.2 Defining the Average Effects of Access

In the context of the proposed model, Head Start access corresponds to whether parents have Head Start in their choice set in Stage 1. As a result, the effect of Head Start access can be evaluated by comparing the parents' enrollment decisions in Stage 2 and their child's subsequent test scores in Stage 3 under choice sets with and without Head Start in Stage 1. Below, I define several parameters that capture these comparisons under choice sets of

$$\{n, h\} \quad \text{versus} \quad \{n\}$$

to measure the effect of Head Start access when access to an alternative preschool is absent, and

$$\{n, a, h\} \quad \text{versus} \quad \{n, a\}$$

to measure the effect of Head Start access when access to an alternative preschool is present.

In Stage 2, the provision of Head Start access may affect the parent's enrollment decision as it may induce a subgroup of parents to enroll their child into Head Start. In particular, when access to an alternative preschool is absent, this subgroup of parents correspond to those whose preference types belong to the following set

$$\mathcal{U}_{nh|n} = \{u \in \mathcal{U} : d(u, \{n, h\}) = h\} \ ,$$

i.e. the preference types who prefer Head Start to no preschool. Analogously, when access to an alternative preschool is present, this subgroup of parents corresponds to

$$\mathcal{U}_{nah|na} = \{u \in \mathcal{U} : d(u, \{n, a, h\}) = h\} \ ,$$

i.e. the preference types who prefer Head Start to an alternative preschool and no preschool. In turn, the average effect of Head Start access on inducing enrollment in Head Start corresponds to the proportion of parents who are induced to participate or enroll into Head Start, which is measured by the parameter

$$\text{PP}_{nh|n} \equiv \text{Prob}[U \in \mathcal{U}_{nh|n}] \tag{4}$$

when access to an alternative preschool is absent, and by the parameter

$$\text{PP}_{nah|na} \equiv \text{Prob}[U \in \mathcal{U}_{nah|na}] \tag{5}$$

when access to an alternative preschool is present.

Since the provision of Head Start access may affect the parents' decisions on where to enroll their child in Stage 2, it may subsequently affect their child's test scores in Stage 3. To formally describe the average effect of Head Start access on test scores, I introduce some additional notation. Had the parents' choice set been pre-specified to a non-empty subset $c \subseteq \mathcal{D}$ in Stage 1, let

$$D_c = \sum_{u \in \mathcal{U}} d(u, c) I\{U = u\}$$

denote the care setting in which the parent would then enroll their child in Stage 2 and let

$$Y_c = \sum_{d \in \mathcal{D}} Y(d) I\{D_c = d\}$$

denote the subsequent test score that their child would then acquire in Stage 3. Using this notation, the average (treatment) effect of Head Start access on test scores is measured by the parameter

$$\text{ATE}_{nh|n} \equiv E\left[Y_{\{n,h\}} - Y_{\{n\}}\right] \tag{6}$$

when access to an alternative preschool is absent, and by the parameter

$$\text{ATE}_{nah|na} \equiv E\left[Y_{\{n,a,h\}} - Y_{\{n,a\}}\right] \tag{7}$$

when access to an alternative preschool is present. These two parameters measure the average difference in a child's potential test score under two different choice sets with and without Head Start. However, if the provision of Head Start access did not affect the parents' enrollment decision in Stage 2 then the child's test score would remain unaffected in Stage 3, i.e.

$$\begin{aligned} U \notin \mathcal{U}_{nh|n} &\implies Y_{\{n,h\}} = Y_{\{n\}} \ , \\ U \notin \mathcal{U}_{nah|na} &\implies Y_{\{n,a,h\}} = Y_{\{n,a\}} \ . \end{aligned}$$

As a result, the average effect of Head Start access on test scores can be rewritten as

$$\begin{aligned} \text{ATE}_{nh|n} &= \text{PP}_{nh|n} \cdot E\left[Y_{\{n,h\}} - Y_{\{n\}} | U \in \mathcal{U}_{nh|n}\right] \ , \\ \text{ATE}_{nah|na} &= \text{PP}_{nah|na} \cdot E\left[Y_{\{n,a,h\}} - Y_{\{n,a\}} | U \in \mathcal{U}_{nah|na}\right] \ . \end{aligned}$$

To this end, by taking the ratio of the average effect of Head Start access on test score to that on enrollment, we can measure the average effect of Head Start access on test scores for the subgroup of parents participating or enrolling their child into Head Start by the parameter

$$\text{ATOP}_{nh|n} \equiv E\left[Y_{\{n,h\}} - Y_{\{n\}} | U \in \mathcal{U}_{nh|n}\right] \tag{8}$$

when access to an alternative preschool is absent, and by the parameter

$$\text{ATOP}_{nah|na} \equiv E\left[Y_{\{n,a,h\}} - Y_{\{n,a\}} | U \in \mathcal{U}_{nah|na}\right] \tag{9}$$

when access to an alternative preschool is present.

3.3 Nonparametric Characterization of the Identified Set

Given a parameter that evaluates the average effect of Head Start access across a specific dimension, I now analyze what we can learn about it given the distribution of the observed data and the assumptions imposed on the model.

For the purposes of the analysis, I assume that the chosen test score of interest take values in a known discrete set

$$\mathcal{Y} = \{y_1, \ldots, y_M\} \ . \tag{10}$$

As observed below, the use of a discretized test score outcome allows tractable characterizing what we can learn using a finite dimensional computational program. Given this discrete test score outcome, let

$$W = (Y(n), Y(a), Y(h), U, C(0), C(1)) \tag{11}$$

denote the random variable that summarizes the underlying latent variables for each child, which takes values on a discrete sample space $\mathcal{W} = \mathcal{Y}^3 \times \mathcal{U} \times \mathcal{C}^2$. Due to the discreteness, note that the distribution of this random variable can be characterized by a probability mass function Q with support contained in $\mathcal{W}$, i.e. $Q : \mathcal{W} \to [0, 1]$ such that

$$\sum_{w \in \mathcal{W}} Q(w) = 1 \ .$$

Similarly, let Q_z denote the probability mass function of the summary random variable conditional on the treatment group assignment indicator Z equal to $z \in \mathcal{Z} \equiv \{0, 1\}$. Let w denote

$$(y(n), y(a), y(h), u, c(0), c(1)) \in \mathcal{W} \ ,$$

i.e. a generic value in the sample space of the summary random variable.

Each of the parameters that we want to learn about can formally be rewritten as functions of Q, where this function shares a common underlying structure. In particular, denoting a given parameter by $\theta(Q)$, the parameter can be re-written in the following form

$$\theta(Q) = \frac{\sum_{w \in \mathcal{W}} a_{\text{num}}(w) \cdot Q(w)}{\sum_{w \in \mathcal{W}_{\text{den}}} Q(w)} \ , \tag{12}$$

where $a_{\text{num}} : \mathcal{W} \to \mathbf{R}$ is known function and $\mathcal{W}_{\text{den}}$ is a known subset of $\mathcal{W}$, i.e. each parameter is a fraction of linear functions of Q. A special case of this class of linear-fractional parameters are linear parameters, where by construction the denominator takes a value of one. For example, the proportion of parents who are induced to enroll their child into Head Start when access to an alternative preschool is absent as given in (4) can be re-written as a linear parameter

$$\text{PP}_{nh|n}(Q) = \sum_{w \in \mathcal{W}_{nh|n}} Q(w) \ ,$$

where $\mathcal{W}_{nh|n} = \{w \in \mathcal{W} : u \in \mathcal{U}_{nh|n}\}$, whereas the average effect of Head Start access in the absence of access to an alternative preschool in (6) can also be re-written as a linear parameter

$$\text{ATE}_{nh|n}(Q) = \sum_{w \in \mathcal{W}} \left[y\left(d(u, \{n, h\}) \right) - y\left(d(u, \{n\}) \right) \right] \cdot Q(w) \ .$$

In contrast, the average effect of Head Start access conditional on the parents enrolling their child into Head Start when access to an alternative preschool is absent as defined in (8) can be re-written as a linear-fractional parameter

$$\text{ATOP}_{nh|n}(Q) = \frac{\sum_{w \in \mathcal{W}} \left[y\left(d(u, \{n, h\}) \right) - y\left(d(u, \{n\}) \right) \right] \cdot Q(w)}{\sum_{w \in \mathcal{W}_{nh|n}} Q(w)} \ .$$

In a similar manner, the various parameters that evaluate the average effects of Head Start access when access to an alternative preschool is present can also be re-written as either linear or linear-fractional functions of Q.

What we can learn about $\theta(Q)$ is determined by the possible values that Q can take. These values are restricted by the distribution of the observed data and the imposed assumptions. The distribution of the data imposes restrictions on the conditional probability mass functions Q_z for $z \in \mathcal{Z}$ through the structure of the model. These restrictions can formally be stated as

$$\sum_{w \in \mathcal{W}_x} Q_z(w) = \text{Prob}\{Y = y, D = d | Z = z\} \tag{13}$$

for all $x = (y, d, z) \in \mathcal{Y} \times \mathcal{D} \times \mathcal{Z} \equiv \mathcal{X}$, where $\mathcal{W}_x$ is the set of all w in $\mathcal{W}$ such that $c = c(1)$ if $z = 1$ and $c = c(0)$ if $z = 0$, $d(u, c) = d$ and $y = y(d)$. More specifically, $\mathcal{W}_x$ is the set of all underlying values in $\mathcal{W}$ that could have possibly generated the observed value $x \in \mathcal{X}$ by the outcome equation in (3) and by the selection equation in (2). Intuitively, these restrictions capture what the parents' observed enrollment decisions and the child's test scores reveal through the structure of the model about the parents' underlying choice set and preferences, and their child's potential test scores.

Assumption HSIS(i) imposes restrictions on how the quantities Q_z for $z \in \mathcal{Z}$ are related to the quantity Q. Since Assumption HSIS(i) states that the treatment group assignment indicator Z is statistically independent of the underlying latent random variables summarized by W, the assumption can be equivalently restated as

$$Q_z(w) = Q(w) \tag{14}$$

for all $w \in \mathcal{W}$ and $z \in \mathcal{Z}$. The restrictions in (13) and (14) can then together be used to state the following restrictions imposed on Q by the observed data

$$\sum_{w \in \mathcal{W}_x} Q(w) = \text{Prob}\{Y = y, D = d | Z = z\} \tag{15}$$

for all $x = (y, d, z) \in \mathcal{X}$, where, as before, $\mathcal{W}_x$ is the set of all w in $\mathcal{W}$ such that $c = c(1)$ if $z = 1$ and $c = c(0)$ if $z = 0$, $d(u, c) = d$ and $y = y(d)$.

Assumption HSIS(ii) directly imposes restrictions on the possible values that Q can take. To illustrate these restrictions, note first that Assumption HSIS(ii) can be equivalently stated as

$$\text{Prob}\{h \notin C(1)\} = 0 \ ,$$

i.e. it imposes zero probability on the occurrence of events where Head Start is not in the potential choice set had the child been assigned to the treatment group. It is then straightforward to see that this statement can then be re-written in terms of a linear restriction on Q as

$$\sum_{w \in \mathcal{W}_{\text{HSIS}}} Q(w) = 0 \ , \tag{16}$$

where $\mathcal{W}_{\text{HSIS}} = \{w \in \mathcal{W} : h \notin c(1)\}$ denotes the set of all underlying values such that potential choice set under the treatment group does not contain Head Start. In Section 3.4, I illustrate additional nonparametric assumptions that can further restrict the baseline model, where these restrictions share a common underlying structure to those imposed by Assumption HSIS(ii). To this end, more generally, let $\mathcal{S}$ denote a finite set of such restrictions imposed on Q by these various assumptions such that each restriction $s \in \mathcal{S}$ satisfies

$$\sum_{w \in \mathcal{W}_s} Q(w) = 0 \ , \tag{17}$$

where $\mathcal{W}_s$ is a known subset of $\mathcal{W}$, i.e. there are only a finite number of restrictions imposed and that each restriction imposes zero probability on the occurrence of certain events.

Given the above restrictions imposed by the distribution of the observed data and by the assumptions on the unknown Q, what we can learn about $\theta(Q)$ is then formally defined by the identified set, i.e. the set of feasible parameter values such that Q satisfies the various imposed restrictions. In order to formally state the identified set, denote first by $\mathbf{Q}_{\mathcal{W}}$ the set of all probability mass functions on the sample space $\mathcal{W}$. The identified set can then be stated as follows

$$\Theta = \{\theta_0 \in \mathbf{R} : \theta(Q) = \theta_0 \text{ for some } Q \in \mathbf{Q}\} \ , \tag{18}$$

where

$$\mathbf{Q} = \{Q \in \mathbf{Q}_{\mathcal{W}} : Q \text{ satisfies (15), and (17) for each } s \in \mathcal{S} \ \} \tag{19}$$

is the set of all Q that satisfy the various restrictions imposed by the data and the assumptions.

In general, analytically characterizing the identified set for $\theta(Q)$ can be difficult due to the large number of restrictions that the data and the assumptions impose on Q and due to the structure of the function $\theta(Q)$. Nonetheless, in the following proposition, I show that this identified set can be tractably characterized as solutions to two linear programming problems. In this proposition below, I introduce the following additional quantity

$$\widetilde{\theta}(Q) = \sum_{w \in \mathcal{W}} a_{\text{num}}(w) \cdot Q(w) \ .$$

Proposition 3.1. Suppose that $\mathbf{Q}$ in (19) is non-empty and the parameter characterized by (12) is such that

$$\sum_{w \in \mathcal{W}_{\text{den}}} Q(w) > 0 \tag{20}$$

holds for every $Q \in \mathbf{Q}$. Then the identified set in (18) can be written as

$$\Theta = [\theta_l, \theta_u] \ , \tag{21}$$

where the (sharp) lower and upper bounds of this interval are solutions to the following two linear programming problems

$$\theta_l = \min_{\gamma, \{Q(w)\}_{w \in \mathcal{W}}} \widetilde{\theta}(Q) \quad \text{and} \quad \theta_u = \max_{\gamma, \{Q(w)\}_{w \in \mathcal{W}}} \widetilde{\theta}(Q) \ , \tag{22}$$

subject to the following constraints:

(i) $\gamma \geq 0$.

(ii) $0 \leq Q(w) \leq \gamma$ for every $w \in \mathcal{W}$.

(iii) $\sum_{w \in \mathcal{W}} Q(w) = \gamma$.

(iv) $\sum_{w \in \mathcal{W}_x} Q(w) = \gamma \cdot \text{Prob}\{Y = y, D = d | Z = z\}$ for every $x = (y, d, z) \in \mathcal{X}$.

(v) $\sum_{w \in \mathcal{W}_s} Q(w) = 0$ for every $s \in \mathcal{S}$.

(vi) $\sum_{w \in \mathcal{W}_{\text{den}}} Q(w) = 1$.

Given the structure of the linear-fractional parameters and the linear restrictions imposed on the probability mass function Q, Proposition 3.1 illustrates that the identified set for each parameter is an interval and a linear programming procedure can be used to compute it. Computing the identified set using Proposition 3.1 requires two conditions. First, it requires $\mathbf{Q}$ to be non-empty, i.e. the model is correctly specified, to ensure that the identified set is an interval. If this is not the case, the linear program automatically terminates. Second, it requires the denominator of the parameter of interest to be positive for every $Q \in \mathbf{Q}$ to ensure that the parameter is well-defined for all feasible model distributions. This condition can easily be verified in practice. To see how, note that the denominator is also a linear-fractional parameter and more specifically a linear parameter. The above proposition can then be employed to compute the lower bound for this auxiliary parameter to check if it is strictly positive.

Linear programming procedures have also been previously proposed by numerous authors in related treatment effect and discrete choice models as a convenient approach for obtaining the identified set for various parameters of interest—see, for example, Balke and Pearl (1997), Demuynck (2015), Freyberger and Horowitz (2015), Kline and Tartari (2016), Lafférs (2013), Manski (2007, 2014), Mogstad et al. (2018) and Torgovitsky (2016, 2018). However, note that the underlying arguments that justify these procedures in these previous papers do not immediately apply to all the parameters defined in the previous section. In particular, these arguments apply to parameters that can only be written as linear functions of Q, but not to those that can be written as linear-fractional functions of Q. In the latter case, as further illustrated in the proof of Proposition

3.1, one needs to first make the observation that the identified set can generally be instead written as solutions to two so-called linear-fractional programs. This observation then allows invoking results from Charnes and Cooper (1962) which imply that the following transformation

$$\widetilde{Q}(w) = \gamma \cdot Q(w) \quad \text{where} \quad \gamma = \frac{1}{\sum\limits_{w \in \mathcal{W}_{\text{den}}} Q(w)} ,$$

can be used to transform the linear-fractional programs to equivalent linear programs.

3.4 Additional Assumptions on the Baseline Model

Recall that the baseline model proposed in Section 3.1 imposed only Assumption HSIS and left the dependence between the variables across the three stages completely unrestricted. However, as we will observe when presenting the empirical results in Section 4, the bounds for some of the parameters under the baseline model can be wide in terms of reaching conclusions regarding the benefits of Head Start access. In such cases, we may therefore be interested in imposing assumptions that restrict the dependence between the various variables. To this end, an attractive feature of Proposition 3.1 is the flexibility with which we can impose such assumptions as long as the restrictions that they impose on Q satisfy (17).

Below, I discuss examples of three such assumptions that restrict the dependence between the variables across the three stages and that can have considerable identifying power as observed when presenting the empirical results. In Section S.6.3 of the Supplement Appendix, I show how, similar to Assumption HSIS(ii), each of these assumptions can be re-written as restrictions on Q in the form of (17).

Assumption UA. (Unaltered Alternative) $a \in C(0) \iff a \in C(1)$.

Assumption MTR. (Monotone Treatment Response) $Y(h) \geq Y(n)$ and $Y(a) \geq Y(n)$.

Assumption Roy. For each $d, d' \in \mathcal{D}$, if $Y(d') > Y(d)$ then $d(U, \{d, d'\}) = d'$.

Assumption UA states that assignment to either the treatment or control groups does not affect the parents' receipt of access to an alternative preschool. It restricts how the choice sets in Stage 1 may be related across the two experimental groups and, in turn, how parents receive preschool access in Stage 1. For example, it rules out cases where being assigned to the control group induces parents to receive access to an alternative preschool, i.e. a case where there is an alternative preschool present in $C(0)$ but not in $C(1)$. Assumption MTR states that the test scores under either Head Start or an alternative preschool cannot be worse off than that under home care. It is motivated by the notion that since preschools attempt to improve cognitive ability, it may be reasonable to believe that both Head Start and alternative preschools are at least as good as home care with respect to a child's test score. Note that this assumption is a version of the monotone treatment response assumption proposed by Manski (1997). Finally, Assumption Roy states that parents strictly prefer to enroll their child in the care setting with the higher test score. It restricts how preferences in Stage 2 may be related to test scores subsequently earned in Stage 3. In particular, it takes preferences to be based solely on maximizing outcomes and, in turn, obtains a version of the Roy selection model (Heckman and Honore, 1990; Mourifie et al., 2015).

4 Empirical Results

The analysis in the preceding section developed a framework to show how the data provided by the HSIS can be exploited to learn about the average effects of Head Start access. In this section, I apply this framework to the data and present the empirical results. Puma et al. (2010) provides a detailed description of the HSIS data. Section S.2.1 and Section S.2.5 of the Supplement Appendix present various summary statistics of interest and details on how the sample used in the empirical analysis presented below was constructed, respectively.

Following the analysis of Puma et al. (2010), I apply the analysis separately to the samples of the age three and four cohorts. The HSIS measured care setting enrollment by a administratively coded focal care setting that evaluated care setting attendance for the entire year. Following the analysis of Feller et al. (2016) and Kline and Walters (2016), I take the observed enrollment decision to be a categorized version of this administratively coded variable. I take the test score outcome to be a discretized version of that used in Kline and Walters (2016). In particular, they use a summary index of cognitive test scores as measured by the average of the Woodcock Johnson III (WJIII) and the Peabody Picture and Vocabulary Test (PPVT) test scores, where each score is standardized to have mean zero and variance one in the control group. Given this summary index, I discretize it using the quantiles of its empirical distribution to take ten support points.[2]

Table 1 reports estimated identified sets for the various parameters evaluating the average effects of Head Start access for the age three and four cohorts. Each row corresponds to a parameter described in Section 3.2, whereas each column corresponds to a specification of the model determined by which of the additional assumptions described in Section 3.4 are imposed on the baseline model. The estimated identified sets are obtained by applying the linear programming procedure in Proposition 3.1 to the empirical distribution of the data. I summarize below the findings by the different parameters for the age three cohort as those for the age four cohort are similar.

PP: The estimated identified sets for $\mathrm{PP}_{nh|n}$ imply that the provision of Head Start access induces between 85.5% and 91.1% of parents to enroll their child into Head Start when access

[2]In Section S.2.4 of the Supplement Appendix, I provide further details on how I construct the discretized test score. I also present results under alternative discretizations and find that the results are numerically similar.

Table 1: Estimated identified sets

	Assumption	(1)	(2)	(3)	(4)	(5)	(6)
	UA		✓			✓	✓
	MTR			✓		✓	
	Roy				✓		✓
	Parameter						
Age 3		Head Start access in absence of alternative preschool access					
	$\text{PP}_{nh\|n}$	0.855	0.855	0.855	0.855	0.855	0.855
		0.911	0.911	0.911	0.911	0.911	0.911
	$\text{ATE}_{nh\|n}$	-0.734	-0.734	0.062	0.062	0.171	0.171
		0.981	0.981	0.981	0.887	0.981	0.887
	$\text{ATOP}_{nh\|n}$	-0.859	-0.806	0.068	0.068	0.187	0.187
		1.135	1.076	1.088	1.037	1.076	0.981
		Head Start access in presence of alternative preschool access					
	$\text{PP}_{nah\|na}$	0.000	0.205	0.000	0.000	0.205	0.205
		0.855	0.855	0.855	0.855	0.855	0.855
	$\text{ATE}_{nah\|na}$	-1.533	-1.152	-1.479	0.000	-1.152	0.000
		1.511	1.207	0.707	0.707	0.426	0.426
	$\text{ATOP}_{nah\|na}$	-	-1.622	-	-	-1.622	0.000
			1.709			1.107	1.107
Age 4		Head Start access in absence of alternative preschool access					
	$\text{PP}_{nh\|n}$	0.788	0.788	0.788	0.788	0.788	0.788
		0.902	0.902	0.902	0.902	0.902	0.902
	$\text{ATE}_{nh\|n}$	-0.933	-0.933	0.007	0.007	0.140	0.140
		1.269	1.269	1.269	1.102	1.269	1.102
	$\text{ATOP}_{nh\|n}$	-1.125	-1.035	0.007	0.007	0.155	0.155
		1.479	1.408	1.454	1.398	1.408	1.241
		Head Start access in presence of alternative preschool access					
	$\text{PP}_{nah\|na}$	0.000	0.274	0.000	0.000	0.274	0.274
		0.788	0.788	0.788	0.788	0.788	0.788
	$\text{ATE}_{nah\|na}$	-1.244	-0.840	-1.225	0.000	-0.840	0.000
		1.412	0.978	0.771	0.769	0.348	0.348
	$\text{ATOP}_{nah\|na}$	-	-1.274	-	-	-1.274	0.000
			1.412			0.870	0.870

Notes: For each estimated identified set, the upper and lower panels correspond to the lower and upper bounds, respectively.

to an alternative preschool is absent. In contrast, those for $\text{PP}_{nah|na}$ imply that the proportion of parents induced to enroll their child into Head Start is always smaller when access to an alternative preschool is present than when absent. Under the baseline model, it is also possible that in this case no parent is induced to enroll their child into Head Start as zero is contained in the identified set. However, when Assumption UA is imposed, the lower bound increases and implies that at least 20.5% of parents are induced to enroll their child into Head Start.

ATE: Under the baseline model, the estimated identified sets for $\text{ATE}_{nh|n}$ and $\text{ATE}_{nah|na}$ imply that we cannot conclude whether there is a positive or negative average effect of Head Start access on test score as zero is contained in them. Intuitively, this uninformativeness arises as the baseline model imposes no assumptions on the potential test scores. To this end, Assumption MTR and Assumption Roy place restrictions on how the potential test scores may be related to each other or the other variables in the model, which then result in informative conclusions. When Assumption MTR is imposed, the estimated identified set for $\text{ATE}_{nh|n}$ implies that providing Head Start access improves test scores on average between 0.062 and 0.981 standard deviations when access to an alternative preschool is absent. When Assumption Roy is imposed, the lower bound is the same as that under Assumption MTR but the upper bound further shrinks to 0.887 standard deviations. Moreover, when Assumption UA is additionally imposed along with either of these assumptions, the lower bound further tightens to imply that there is an improvement of at least 0.171 standard deviations in average test scores when access to an alternative preschool is absent.

In contrast, the results for $\text{ATE}_{nah|na}$ imply that the provision of Head Start access can possibly have no improvements on average test scores when access to an alternative preschool is present. In particular, under the most informative specification given by imposing Assumption UA and Assumption Roy, the estimated identified sets imply that the possible improvement in test scores can possibly be between 0 and 0.426 standard deviations.

ATOP: Unlike $\text{ATE}_{nh|n}$ and $\text{ATE}_{nah|na}$, $\text{ATOP}_{nh|n}$ and $\text{ATOP}_{nah|na}$ condition on the subgroup of parents induced to enroll their child into Head Start and in turn affected by the provision of access when evaluating the average effect of Head Start access on test scores. Note, as explained in Section 3.3, $\text{ATOP}_{nh|n}$ and $\text{ATOP}_{nah|na}$ are linear-fractional parameters and their identified sets can be computed only when their corresponding denominators are strictly positive. Their denominators are respectively given by $\text{PP}_{nh|n}$ and $\text{PP}_{nah|na}$. As a result, since $\text{PP}_{nah|na}$ is not strictly positive unless Assumption UA is imposed, I only report estimated identified sets for $\text{ATOP}_{nah|na}$ when Assumption UA is imposed. The various conclusions based on these parameters are similar to those from $\text{ATE}_{nh|n}$ and $\text{ATE}_{nah|na}$, except that the magnitudes of the bounds are multiplied by some positive factor.

Overall, the conclusions from the estimated identified sets can be qualitatively summarized as follows: (i) the provision of Head Start access induces a positive proportion of parents to enroll their child into Head Start, where this proportion is larger when access to an alternative preschool is absent; (ii) the provision of Head Start access has a positive impact on average test scores of children whose parents are induced to enroll them into Head Start when access to an alternative preschool is absent; and (iii) the provision of Head Start access can possibly have no impact on average test scores of children when access to an alternative preschool is present.

Section S.2.3 and Section S.2.2 of the Supplement Appendix present additional empirical results. In Section S.2.2, I present confidence intervals for several of the parameters of interest; whereas, in Section S.2.3, I analyze the sensitivity of the above conclusions to when Assumption MTR and Assumption Roy are weakened such that they hold only for a given proportion of the population. In general, I find that the overall conclusions continue to hold after accounting for the sample uncertainty in the estimates and under mild departures of the imposed assumptions.

5 Conclusion

Many experiments are conducted by randomly assigning individuals to either a treatment group with program access or a control group without access. This paper developed a framework to evaluate the average effects of program access in such experiments when individuals do not comply with their assigned treatment status and when data on where individuals receive access is not collected. The main idea behind the framework is to use a selection model to relate where access was received to the observed data, and to then exploit the structure of the model to define and learn about various parameters evaluating the average effects of program access. I illustrated the framework by analyzing the average effects of providing Head Start access using the Head Start Impact Study (HSIS).

I conclude by highlighting the flexibility of the developed framework. Specifically, it can be broadly applied to different experiments where it can accommodate several institutional details unique to that experiment. In the paper, I demonstrated this feature by developing the framework in the context of the HSIS where I showed how the framework accommodated several details such as the presence of alternative preschools and the specific information that the experiment provided on where access was received. I demonstrate this flexibility further in Section S.5 of the Supplement Appendix, where I present the generalized version of the framework and show how it can be applied to accommodate alternative details in two other experiments.

References

Abaluck, J. and Adams, A. (2018). What do consumers consider before they choose? identification from asymmetric demand responses. Tech. rep.

Angelucci, M., Karlan, D. and Zinman, J. (2015). Microcredit impacts: Evidence from a randomized microcredit program placement experiment by compartamos banco. *American Economic Journal: Applied Economics*, **7** 151–82.

Balke, A. and Pearl, J. (1997). Bounds on treatment effects from studies with imperfect compliance. *Journal of the American Statistical Association*, **92** 1171–1176.

Barseghyan, L., Coughlin, M., Molinari, F. and Teitelbaum, J. (2018). Heterogeneous consideration sets and preferences. Tech. rep., Cornell University.

Ben-Akiva, M. and Boccara, B. (1995). Discrete choice models with latent choice sets. *International journal of Research in Marketing*, **12** 9–24.

Cattaneo, M. D., Ma, X., Masatlioglu, Y. and Suleymanov, E. (2018). A random attention model. *arXiv preprint arXiv:1712.03448*.

Charnes, A. and Cooper, W. W. (1962). Programming with linear fractional functionals. *Naval Research Logistics (NRL)*, **9** 181–186.

Crawford, G., Griffith, R. and Iaria, A. (2019). Preference estimation with unobserved choice set heterogeneity using sufficient sets.

Demuynck, T. (2015). Bounding average treatment effects: A linear programming approach. *Economics Letters*, **137** 75–77.

Feller, A., Grindal, T., Miratrix, L. and Page, L. C. (2016). Compared to what? variation in the impacts of early childhood education by alternative care type. *The Annals of Applied Statistics*, **10** 1245–1285.

Freyberger, J. and Horowitz, J. L. (2015). Identification and shape restrictions in nonparametric instrumental variables estimation. *Journal of Econometrics*, **189** 41–53.

Gaynor, M., Propper, C. and Seiler, S. (2016). Free to choose? reform, choice, and consideration sets in the english national health service. *American Economic Review*, **106** 3521–57.

Goeree, M. S. (2008). Limited information and advertising in the us personal computer industry. *Econometrica*, **76** 1017–1074.

Heckman, J. J. and Honore, B. E. (1990). The empirical content of the roy model. *Econometrica: Journal of the Econometric Society* 1121–1149.

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

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

Heckman, J. J., Urzua, S. and Vytlacil, E. (2008). Instrumental variables in models with multiple outcomes: The general unordered case. *Annales d'Economie et de Statistique* 151–174.

Ho, K. and Rosen, A. M. (2017). *Partial Identification in Applied Research: Benefits and Challenges*, vol. 2 of *Econometric Society Monographs*. Cambridge University Press, 307–359.

Horowitz, J. L. and Manski, C. F. (1998). Censoring of outcomes and regressors due to survey nonresponse: Identification and estimation using weights and imputations. *Journal of Econometrics*, **84** 37–58.

Horowitz, J. L. and Manski, C. F. (2000). Nonparametric analysis of randomized experiments with missing covariate and outcome data. *Journal of the American statistical Association*, **95** 77–84.

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

Kline, P. and Tartari, M. (2016). Bounding the labor supply responses to a randomized welfare experiment: A revealed preference approach. *The American Economic Review*, **106** 971–1013.

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

Lafférs, L. (2013). A note on bounding average treatment effects. *Economics Letters*, **120** 424–428.

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

Manski, C. F. (1997). Monotone treatment response. *Econometrica: Journal of the Econometric Society* 1311–1334.

Manski, C. F. (2007). Partial identification of counterfactual choice probabilities. *International Economic Review*, **48** 1393–1410.

Manski, C. F. (2014). Identification of income–leisure preferences and evaluation of income tax policy. *Quantitative Economics*, **5** 145–174.

Marschak, J. (1960). Binary choice constraints on random utility indicators. In *Stanford Symposium on Mathematical Methods in the Social Sciences*. Stanford University Press.

Mogstad, M., Santos, A. and Torgovitsky, A. (2018). Using instrumental variables for inference about policy relevant treatment parameters. *Econometrica*, **86** 1589–1619.

Mogstad, M. and Torgovitsky, A. (2018). Identification and extrapolation of causal effects with instrumental variables. *Annual Review of Economics*.

Molinari, F. (2010). Missing treatments. *Journal of Business & Economic Statistics*, **28** 82–95.

Mourifie, I., Henry, M. and Meango, R. (2015). Sharp bounds for the roy model.

Puma, M., Bell, S., Cook, R., Heid, C., Shapiro, G., Broene, P., Jenkins, F., Fletcher, P., Quinn, L., Friedman, J. et al. (2010). Head start impact study. final report. *Administration for Children & Families*.

Torgovitsky, A. (2016). Nonparametric inference on state dependence with applications to employment dynamics.

Torgovitsky, A. (2018). Partial identification by extending subdistributions.