Supplementary Appendix to "Estimating Welfare Effects in a Nonparametric Choice Model: The Case of School Vouchers"

Vishal Kamat
School of Economics and Finance
Queen Mary University of London
v.kamat@qmul.ac.uk

Samuel Norris
Department of Economics
University of British Columbia
sam.norris@ubc.ca

May 7, 2025

Abstract

This document presents proofs and additional details pertinent to the main analysis for the authors' paper titled "Estimating Welfare Effects in a Nonparametric Choice Model: The Case of School Vouchers." Section S.1 describes how to construct the partition in our baseline analysis. Section S.2 provides details on the various extensions to the baseline analysis. Section S.3 describes the procedure used to perform statistical inference in the empirical analysis. Section S.4 derives the functional forms on demand implied by the Logit specifications considered in the comparison. Section S.5 presents additional details relevant to the empirical analysis. Section S.6 presents proofs of all results. Section S.7 presents additional tables and figures.

S.1 Constructing $\mathcal{V}$

In this section, we show how to obtain a collection of sets $\mathcal{V}$ that partitions $\mathcal{P}^* = \bigcup_{l=1}^{L} \mathcal{P}_l^*$ satisfying Definition V. To this end, observe first that the union of the sets in $\mathcal{V}^{a,b}$ and $\{\{p\} : p \in \mathcal{P}_{\text{obs}}\}$ equals $\mathcal{P}^*$, where $\mathcal{V}^{a,b} = \left\{\mathcal{P}_l^{a,b}, \ldots, \mathcal{P}_{|\mathcal{J}|-1}^{a,b}, \{p^a\}, \{p^b\}\right\}$, and further that $\mathcal{V}^{a,b}$ and $\{\{p\} : p \in \mathcal{P}_{\text{obs}}\}$ each correspond to a partition of the union of the sets in it that satisfies Definition V(i)-(iii) with respect to its union. However, for a given $v \in \mathcal{V}^{a,b}$ and $v' \in \{\{p\} : p \in \mathcal{P}_{\text{obs}}\}$, it may be that Definition V(iii) is not satisfied. In particular, it may be the case that there exists $p \in \mathcal{P}_{\text{obs}}$ that intersects some $v \in \mathcal{V}^{a,b}$ in the jth dimension, i.e. $\{p_j\} \subset v_{[j]}$ for some $j \in \mathcal{J}$, and hence we have $\{p_j\} \neq v_{[j]}$ and $\{p_j\} \cap v_{[j]} \neq \emptyset$.

To obtain our partition, we therefore further partition the elements in $\mathcal{V}^{a,b}$ to obtain $\tilde{\mathcal{V}}^{a,b}$ such that Definition V(iii) is satisfied between the elements of $\tilde{\mathcal{V}}^{a,b}$ and $\{\{p\} : p \in \mathcal{P}_{\text{obs}}\}$ by accounting for whether $p \in \mathcal{P}_{\text{obs}}$ and $v \in \mathcal{V}^{a,b}$ intersect in each of the $j \in \mathcal{J}$ dimensions. In particular, let $\mathcal{T} = \{\Delta_l^{a,b} : 1 \leq l \leq |\mathcal{J}| - 1\} \cup \{\max\{\min\{p_j, p_j^a\}, p_j^b + \Delta_1^{a,b}\} - p_j^b : j \in \mathcal{J},\ p \in \mathcal{P}_{\text{obs}}\}\}$ denote the set of points such that $\{\min\{p_j^a, p_j^b + t\} : t \in \mathcal{T}\}$ correspond to the end-points of the sets in $\mathcal{P}_l^{a,b}$, $1 \leq l \leq |\mathcal{J}| - 1$, as well as where $p \in \mathcal{P}_{\text{obs}}$ may intersect the sets $\mathcal{V}^{a,b}$ in each of the $j \in \mathcal{J}$ dimensions—note here that the minimums and maximums ensure that we only include the intersection points if they are in $[p_j^b + \Delta_1^{a,b}, p_j^a]$. Denoting by $t_1 < \ldots < t_{|\mathcal{T}|}$ the ordered values of $\mathcal{T}$, then take $\tilde{\mathcal{V}}^{a,b} = \left\{\tilde{\mathcal{P}}_l^{a,b}, \ldots, \tilde{\mathcal{P}}_{|\mathcal{T}|-1}^{a,b}, \{p^a\}, \{p^b\}\right\}$, where $\tilde{\mathcal{P}}_l^{a,b} = \left\{p \in \mathcal{P} : p_j = \min\{p_j^a, p_j^b + t\},\ t \in (t_l, t_{l+1}) \text{ for } j \in \mathcal{J}\right\}$ for $1 \leq l \leq |\mathcal{T}| - 1$. As before, observe that $\tilde{\mathcal{V}}^{a,b}$ and $\{\{p\} : p \in \mathcal{P}_{\text{obs}}\}$ each correspond to a partition of the union of the sets in it that satisfy Definition V(i)-(iii) with respect to its union. However, now by construction, we also have that Definition V(iii) is satisfied for each $v \in \tilde{\mathcal{V}}^{a,b}$ and $v' \in \mathcal{P}^{\text{obs}}$. To this end, we take

$$\mathcal{V} = \tilde{\mathcal{V}}^{a,b} \bigcup \{\{p\} : p \in \mathcal{P}_{\text{obs}}\}\ ,$$

which corresponds to a partition of $\mathcal{P}^*$ satisfying Definition V.

Note that the above constructed $\mathcal{V}$ corresponds to the coarsest partition of $\mathcal{P}^*$ satisfying Definition V in the sense for any other partition $\mathcal{V}'$ satisfying Definition V we have that for all $v \in \mathcal{V}$ there exists $\mathcal{V}'' \subseteq \mathcal{V}'$ such that $\cup\{v' : v' \in \mathcal{V}''\} = v$, i.e. any other partition is finer and can be used to construct our choice of $\mathcal{V}$. To see this, suppose that this wasn't the case. This means that for some partition $\mathcal{V}'$ satisfying Definition V there exists $v \in \mathcal{V}$ and $v' \in \mathcal{V}'$ such that $v' \not\subseteq v$ and $v \cap v' \neq \emptyset$, i.e. v' is not contained in v and that they overlap. In this case, it must be that $v = \left\{p \in \mathcal{P} : p_j = \min\{p_j^a, p_j^b + t\},\ t \in (t_l, t_{l+1}) \text{ for } j \in \mathcal{J}\right\}$ for some $1 \leq l \leq |\mathcal{T}| - 1$ and $v' = \left\{p \in \mathcal{P} : p_j = \min\{p_j^a, p_j^b + t\},\ t \in (t', t'') \text{ for } j \in \mathcal{J}\right\}$ for some values of t', t'', where $t_{l+1} > t'$ or $t'' > t_l$. However, as the endpoints of v equal those of sets in $\mathcal{P}^*$ in some dimensions, we have that v' is not contained in and overlaps with these sets in $\mathcal{P}^*$ in these dimensions. In turn, $\mathcal{V}'$ cannot satisfy Definition V(i), which is a contradiction.

S.2 Additional Details on Extensions in Section 3

S.2.1 Example of a Set $\mathbf{B}^{\mathrm{r}}$

In this section, we describe a set $\mathbf{B}^{\mathrm{r}}$ such that $\phi(\mathbf{B}) \subseteq \mathbf{B}^{\mathrm{r}}$ that we use in our empirical analysis when implementing the linear programs in (28). To this end, it is first useful to consider an equivalent representation of the restrictions in (S.46) written in terms of pairs $w, w' \in \mathcal{W}$.

Proposition S.1. β satisfies (S.46) if and only if

$$\sum_{j \in \mathcal{J}^{>}_{w',w} \cup \mathcal{J}^{\dagger}} \beta_j(w) \geq \sum_{j \in \mathcal{J}^{>}_{w',w} \cup \mathcal{J}^{\dagger}} \beta_j(w') \tag{S.1}$$

for each $\mathcal{J}^{\dagger} \subseteq \mathcal{J}^{=}_{w,w'}$ and $w, w' \in \mathcal{W}$, where $\mathcal{J}^{>}_{w,w'} = \left\{ j \in \mathcal{J} : t > t' \text{ for all } t \in w_{[j]}\ ,\ t' \in w'_{[j]} \right\}$ and $\mathcal{J}^{=}_{w,w'} = \mathcal{J} \setminus \left(\mathcal{J}^{>}_{w,w'} \cup \mathcal{J}^{>}_{w',w} \right)$.

Given the equivalence between the restrictions in (S.46) and (S.1), we can alternatively write $\mathbf{B}$ in (S.48) as $\mathbf{B} = \left\{ \beta \in \mathbf{R}^{d_\beta} : \beta \text{ satisfies (S.44)} - \text{(S.45)},\ \text{(S.1)},\ \text{and (S.47)} \right\}$. In this set, observe that each restriction on β is for a given $w \in \mathcal{W}$ or for a pair of $w, w' \in \mathcal{W}$. Our choice of $\mathbf{B}^{\mathrm{r}}$ corresponds to the subset of these restrictions on β for $w \in \mathcal{W}^{\mathrm{r}}$ or for pairs of $w, w' \in \mathcal{W}^{\mathrm{r}}$, i.e. the subset of restrictions that directly correspond those that are in terms of β^{r}. More specifically, these restrictions correspond to the following

$$\beta^{\mathrm{r}}_j(w) \geq 0 \quad \text{for each } j \in \mathcal{J} \text{ and } w \in \mathcal{W}^{r}\ , \tag{S.2}$$

$$\sum_{j \in \mathcal{J}} \beta^{\mathrm{r}}_j(w) = 1 \quad \text{for each } w \in \mathcal{W}^{r}\ , \tag{S.3}$$

$$\sum_{j \in \mathcal{J}^{>}_{w',w} \cup \mathcal{J}^{\dagger}} \beta^{\mathrm{r}}_j(w) \geq \sum_{j \in \mathcal{J}^{>}_{w',w} \cup \mathcal{J}^{\dagger}} \beta^{\mathrm{r}}_j(w') \quad \text{for each} \quad \mathcal{J}^{\dagger} \subseteq \mathcal{J}^{=}_{w,w'} \quad \text{and} \quad w, w' \in \mathcal{W}^{\mathrm{r}}\ , \tag{S.4}$$

$$\beta^{\mathrm{r}}_j(\{p\}) = \mathrm{Prob}[D_i = j | P_i = p] \quad \text{for each } j \in \mathcal{J},\ p \in \mathcal{P}_{\mathrm{obs}}\ . \tag{S.5}$$

Then, denoting by $d_{\beta^{\mathrm{r}}}$ the dimension of β^{r}, the set we consider is given by $\mathbf{B}^{\mathrm{r}} = \{\beta^{\mathrm{r}} \in \mathbf{R}^{d_{\beta^{\mathrm{r}}}} : \beta^{\mathrm{r}} \text{ satisfies (S.2)} - \text{(S.5)}\}$.

S.2.2 Extension to Separability Assumptions

In this section, we show how to extend the arguments in Section 2.5 to compute the identified set under a general class of separability assumptions. For each $p \in \mathcal{P}$ and $\mathcal{J}' \subseteq \mathcal{J}$, let

$$p_{[\mathcal{J}']} = (p_j : j \in \mathcal{J}') \in \prod_{j \in \mathcal{J}'} [\underline{p}_j, \bar{p}_j]$$

denote the sub-vector of p with indices in $\mathcal{J}'$. Using this notation, the general separability assumption we consider can be stated as follows:

Assumption GS. (General Separability) For each $j \in \mathcal{J}$,

$$q_j(p) = \sum_{l \in \mathcal{L}_j} h_{jl}\left(p_{[\mathcal{J}_{jl}]}\right) , \tag{S.6}$$

where $\mathcal{L}_j$ and $\mathcal{J}_{jl}$ are pre-specified subsets of $\mathcal{J}$, and h_{jl} are some unknown functions.

Assumption GS imposes demand to be a sum of functions that are of lower dimension than demand. Observe that Assumption S is a special case of this assumption, where $\mathcal{L}_j = \mathcal{J}$ and $\mathcal{J}_{jl} = \{l\}$. We highlight that it also allows for the case where no dimension reduction is imposed and demand is as that in baseline non-separable case, by taking $\mathcal{L}_j$ to be a singleton set such as $\{j\}$ and taking $\mathcal{J}_{jl} = \mathcal{J}$—see proof of Proposition 2 for more details. Under this assumption, our admissible set of demand functions is given by

$$\mathbf{Q}_S = \{q \in \mathbf{F} : q \text{ satisfies } (11) - (12),\ (13) \text{ and } (S.6)\} , \tag{S.7}$$

i.e. $\mathbf{Q}_B$ in (14), but with the additional separability restriction in (S.6).

In order to describe how to compute the identified set $\theta(\mathbf{Q}_S)$, we show as in Proposition 2 that there is no loss of information in first replacing $\mathbf{Q}_S$ with a certain finite-dimensional space, given which the identified set can then be computed using linear programs. Let $\mathcal{W}$ denote the constructed partition of $\mathcal{P}$ from (18), and, for each $w \in \mathcal{W}$ and $\mathcal{J}' \subseteq \mathcal{J}$, let

$$w_{[\mathcal{J}']} = \left\{p_{[\mathcal{J}']} : p \in w\right\} \in \prod_{j \in \mathcal{J}'} \mathcal{V}_j \equiv \mathcal{W}_{[\mathcal{J}']}$$

denote the set that includes the sub-vector of prices in w with indices in $\mathcal{J}'$—note that this is with some abuse in notation as for a single index $j \in \mathcal{J}$, we simply use $w_{[j]}$ as previously introduced rather than $w_{[\{j\}]}$. The finite dimensional space we consider is then given by

$$\mathbf{Q}_S^{\text{fd}} = \left\{ q \in \mathbf{Q}_B : q_j(p) = \sum_{w \in \mathcal{W}} 1_w(p) \sum_{l \in \mathcal{L}_j} \psi_{jl}(w_{[\mathcal{J}_{jl}]}) \text{ for some } \{\psi_{jl}(w_{[\mathcal{J}_{jl}]})\}_{\substack{w_{[\mathcal{J}_{jl}]} \in \mathcal{W}_{[\mathcal{J}_{jl}]} \\ l \in \mathcal{L}_j}},\ j \in \mathcal{J} \right\} , \tag{S.8}$$

i.e. the same as (15) but with the additional restriction that the constant valued functions satisfy (S.6). Let ψ denote $\{\psi_{jl}(w_{[\mathcal{J}_{jl}]}) : w_{[\mathcal{J}_{jl}]} \in \mathcal{W}_{[\mathcal{J}_{jl}]},\ l \in \mathcal{L}_j,\ j \in \mathcal{J}\}$ in vector form. Observe that its dimension is given by $\sum_{j \in \mathcal{J}} \sum_{l \in \mathcal{L}_j} \prod_{m \in \mathcal{J}_{[jl]}} |\mathcal{V}_m|$, while that of β in (15) is $|\mathcal{J}| \prod_{m \in \mathcal{J}} |\mathcal{V}_m|$.

As in Section 2.5.2, to show that $\theta(\mathbf{Q}_S) = \theta(\mathbf{Q}_S^{\text{fd}})$ and how to compute $\theta(\mathbf{Q}_S^{\text{fd}})$, it is next useful to rewrite $\theta(\mathbf{Q}_S^{\text{fd}})$ in term of ψ. Let θ_S denote the continuous function that rewrites the parameter of interest in terms of ψ in the sense that $\theta(q) = \theta_S(\psi)$. Similarly, let

$$\mathbf{S} = \left\{ \psi \in \mathbf{R}^{d_\psi} : \left(\sum_{w \in \mathcal{W}} 1_w \sum_{l \in \mathcal{L}_j} \psi_{jl}(w_{[\mathcal{J}_{jl}]}) : j \in \mathcal{J} \right) \in \mathbf{Q}_B \right\} , \tag{S.9}$$

where d_ψ denotes the dimension of ψ, capture the restrictions on demand in terms of ψ. We can then write $\theta(\mathbf{Q}_S^{\mathrm{fd}})$ in terms of ψ by

$$\Theta_S \equiv \{\theta_0 \in \mathbf{R} : \theta_S(\psi) = \theta_0 \text{ for some } \psi \in \mathbf{S}\} \ . \tag{S.10}$$

As in Proposition 2, we show in the following proposition that $\theta(\mathbf{Q}_S)$ is equal to Θ_S, and that it can be compute by solving two optimization problem. We again highlight in its proof that θ_S is linear and that $\mathbf{S}$ is linear, which imply that the optimization problems are linear programs.

Proposition S.2. Suppose that $\mathbf{Q} = \mathbf{Q}_S$. Then, the identified set in (10) is equal to that in (S.10), i.e. $\Theta = \Theta_S$. In addition, if $\mathbf{S}$ is empty then by definition Θ_S is empty; whereas, if $\mathbf{S}$ is non-empty then the closure of Θ_S is given by $[\underline{\theta}_S, \bar{\theta}_S]$, where

$$\underline{\theta}_S = \inf_{\psi \in \mathbf{S}} \ \theta_S(\psi) \ \text{ and } \ \bar{\theta}_S = \sup_{\psi \in \mathbf{S}} \ \theta_S(\psi) \ . \tag{S.11}$$

S.2.3 Extension to Auxiliary Parametric Assumptions

In this section, we show to compute the identified set when Assumption P from Section 3.2 is additionally imposed. The admissible set of functions in this case is given by

$$\mathbf{Q}_A = \left\{q \in \bar{\mathbf{F}} : q \text{ satisfies } (8), (11) - (13) \text{ and } (29)\right\} \ . \tag{S.12}$$

i.e. $\mathbf{Q}_B$ along with the restriction in (29), and the identified set by $\theta(\mathbf{Q}_A)$. Unlike the nonparametric analysis in Section 2.5, note that computing this identified set is a finite-dimensional problem as $\mathbf{Q}_A$ is a finite-dimensional parameterized space given that the demand functions are parameterized by the variable α. In turn, $\theta(\mathbf{Q}_A)$ can be directly characterized by searching over q in $\mathbf{Q}_A$ and then taking their image under the function θ.

To show how to do this, it is useful, as in Section 2.5.2, to rewrite $\theta(\mathbf{Q}_A)$ in terms of α, the variable parameterizing $q \in \mathbf{Q}_A$. As θ is continuous in q and q is continuous in α, let θ_A be the continuous function of α such that $\theta(q) = \theta_A(\alpha)$. Moreover, $\mathbf{Q}_A$ can be written in terms of α by

$$\mathbf{A} = \left\{\alpha \in \mathbf{R}^{d_\alpha} : \left(\sum_{k=0}^{K_j} \alpha_{jk} \cdot b_{jk} : j \in \mathcal{J}\right) \in \mathbf{Q}_B\right\} \ . \tag{S.13}$$

where d_α denotes the dimension of α, i.e. the set of values of α that ensure that the corresponding q is in $\mathbf{Q}_B$. We can write $\theta(\mathbf{Q}_A)$ in terms of α by

$$\theta_A(\mathbf{A}) = \{\theta_0 \in \mathbf{R} : \theta_A(\alpha) = \theta_0 \text{ for some } \alpha \in \mathbf{A}\} \equiv \Theta_A \ . \tag{S.14}$$

In the following proposition, we show that when $\mathbf{A}$ is connected and non-empty, the closure of Θ_A is equal to an interval, where the endpoints can be characterized as solutions to two finite-dimensional optimization problems.

Proposition S.3. If $\mathbf{A}$ is empty then by definition Θ_A is empty; whereas, if $\mathbf{A}$ is connected and non-empty, then the closure of Θ_A is given by $[\underline{\theta}_A, \bar{\theta}_A]$, where

$$\underline{\theta}_A = \inf_{\alpha \in \mathbf{A}} \theta_A(\alpha) \quad \text{and} \quad \bar{\theta}_A = \sup_{\alpha \in \mathbf{A}} \theta_A(\alpha) \ . \tag{S.15}$$

Proposition S.3 shows how to characterize the identified set under a general class of parametric restrictions. While the problems in (S.15) are finite-dimensional, their tractability depends on the structure of the objective θ_A and the constraint set $\mathbf{A}$. Given the linear structure of θ and $\mathbf{Q}_B$ along with (29), we have that θ_A is linear and $\mathbf{A}$ is determined by linear restrictions. For example, under the parameterization we consider in (30), observe that (7) can be written in terms of α as

$$\begin{aligned}\theta_A(\alpha) = g^{a,b}\Delta_1^{a,b} + g^{a,b} \sum_{l=1}^{|\mathcal{J}|-1} \sum_{j \in \mathcal{J}_{l+1}^{a,b}} \Bigg(&\sum_{m \in \mathcal{J}_{l+1}^{a,b}} \sum_{k=0}^{K} \alpha_{jmk} \cdot (p_m^a)^k \cdot (\Delta_{l+1}^{a,b} - \Delta_l^{a,b}) + \\ &\sum_{m \in \mathcal{J} \setminus \mathcal{J}_{l+1}^{a,b}} \sum_{k=0}^{K} \alpha_{jmk} \cdot \left(\frac{(p_m^b + t)^{k+1}}{k+1} \Bigg|_{\Delta_l^{a,b}}^{\Delta_{l+1}^{a,b}} \right) \Bigg) \\ + \sum_{j \in \mathcal{J}} \sum_{m \in \mathcal{J}} \sum_{k=0}^{K} \alpha_{jmk} \cdot \Big(g_j^a \cdot (p_m^a)^k + g_j^b \cdot (p_m^b)^k \Big) \end{aligned} \tag{S.16}$$

which follows from substituting the relation between q and α from (30) in (7), and then evaluating the integrals in the expression. Similarly, to see the linear restrictions that determine $\mathbf{A}$ here, observe that by substituting the relation between q and α in (11)-(12), (13) and (8) we obtain

$$\sum_{m \in \mathcal{J}} \sum_{k=0}^{K} \alpha_{jmk} \cdot p_m^k \geq 0 \quad \text{for each } j \in \mathcal{J} \ , \tag{S.17}$$

$$\sum_{j \in \mathcal{J}} \sum_{m \in \mathcal{J}} \sum_{k=0}^{K} \alpha_{jmk} \cdot p_m^k = 1 \tag{S.18}$$

for all $p \in \mathcal{P}$,

$$\sum_{m \in \mathcal{J}} \sum_{k=0}^{K} \alpha_{jmk} \cdot \Big((p_m)^k - (p'_m)^k \Big) \geq 0 \tag{S.19}$$

for each $j \in \mathcal{J} \setminus \mathcal{J}'$ and $p, p' \in \mathcal{P}$ such that $p_j > p'_j$ for $j \in \mathcal{J}' \subseteq \mathcal{J}$ and $p_j = p'_j$ for $j \in \mathcal{J} \setminus \mathcal{J}'$, and

$$\sum_{m=1}^{J} \sum_{k=0}^{K} \alpha_{jmk} \cdot (p_m)^k = \text{Prob}[D = j | P_i = p] \ , \tag{S.20}$$

for each $j \in \mathcal{J}$ and $p \in \mathcal{P}_{\text{obs}}$, given which we have that $\mathbf{A} = \{\alpha \in \mathbf{R}^{d_\alpha} : \alpha \text{ satisfies } (\text{S.17}) - (\text{S.20})\}$.

However, note that the restrictions in (11)-(13) are evaluated at every price in the continuous space $\mathcal{P}$, which implies that the resulting optimization problems can be generally difficult to compute. In practice, we therefore propose to consider the following alternative optimization problems

$$\underline{\theta}_A^{\text{r}} = \min_{\alpha \in \mathbf{A}^{\text{r}}} \theta_A(\alpha) \quad \text{and} \quad \bar{\theta}_A^{\text{r}} = \max_{\alpha \in \mathbf{A}^{\text{r}}} \theta_A(\alpha) \ , \tag{S.21}$$

where $\mathbf{A}^{\mathrm{r}}$ corresponds to a subset of $\mathbf{A}$ with the restrictions in (11)-(13) evaluated on a discrete set of prices $\mathcal{P}^{\mathrm{r}} = \{p_0, \ldots, p_L\}$. For example, under the parameterization in (30), we have by simplifying and removing some jointly redundant restrictions using some algebra that the set of restrictions in (S.17)-(S.20) when evaluated on $\mathcal{P}^{\mathrm{r}}$ can be explicitly given by

$$\sum_{k=0}^{K} \alpha_{jjk} \cdot p_{l,j}^{k} + \sum_{m \in \mathcal{J} \setminus \{j\}} \sum_{k=0}^{K} \alpha_{jmk} \cdot p_{0,m}^{k} \geq 0 \quad \text{for each } j \in \mathcal{J},\ 0 \leq l \leq L\ , \tag{S.22}$$

$$\sum_{j \in \mathcal{J}} \sum_{m \in \mathcal{J}} \sum_{k=0}^{K} \alpha_{jmk} \cdot p_{0,m}^{k} = 1\ , \tag{S.23}$$

$$\sum_{j \in \mathcal{J}} \sum_{k=0}^{K} \alpha_{jmk} \cdot \Delta p_{l+1,m}^{k} = 0 \quad \text{for } m \in \mathcal{J},\ 0 \leq l \leq L-1\ , \tag{S.24}$$

$$\sum_{k=0}^{K} \alpha_{jmk} \cdot \Delta p_{l+1,m}^{k} \geq 0 \quad \text{for each } j \in \mathcal{J},\ m \neq j \in \mathcal{J},\ 0 \leq l \leq L-1\ , \tag{S.25}$$

and (S.20), given which $\mathbf{A}^{\mathrm{r}} = \{\alpha \in \mathbf{R}^{d_\alpha} : \alpha \text{ satisfies (S.22)}-\text{(S.25) and (S.20)}\}$, where $\{p_{0,j}, \ldots, p_{L,j}\}$ denotes the set of ordered values of $\mathcal{P}_j^{\mathrm{r}}$ for each $j \in \mathcal{J}$ and $\Delta p_{l,j}^{k} = p_{l,j}^{k} - p_{l-1,j}^{k}$ denotes the difference in two consecutive prices in this set with each of these two prices raised to the power of k. As the objective and the finite number of restrictions determining the constraint sets of these problems are linear in α, the problems in (S.21) are linear programs and hence generally computationally tractable. But, since $\mathbf{A} \subseteq \mathbf{A}^{\mathrm{r}}$, note that these problems only provide an outer set for Θ_A, i.e. $\Theta_A \subseteq [\underline{\theta}_A^{\mathrm{r}}, \bar{\theta}_A^{\mathrm{r}}]$, similar in spirit to those in (28) with respect to Θ_B.

To implement the above, we need to choose the discrete set of prices $\mathcal{P}^{\mathrm{r}}$. In our application in Section 4, we take $\mathcal{P}^{\mathrm{r}} = \prod_{j=1}^{J} \mathcal{P}_j^{\mathrm{r}}$, where $\mathcal{P}_j^{\mathrm{r}} = \{\underline{p}_j + \frac{l}{\bar{L}}(\bar{p}_j - \underline{p}_j) : 0 \leq l \leq \bar{L}\} \cup \{p_j : p \in \{p^a, p^b\} \cup \mathcal{P}_{\mathrm{obs}}\}$ for some pre-specified value of $\bar{L}$ and with $\underline{p}_j$ and $\bar{p}_j$ equal to the minimum and maximum values of jth dimension of the prices in $\{p^a, p^b\} \cup \mathcal{P}_{\mathrm{obs}}$, i.e. a set of $(\bar{L}+1)$ equidistant points between the minimum and maximum support of prices used in the definition of the parameter and observed in the data. The empirical results take $\bar{L} = 6$. In unreported results, we find that increasing the value of $\bar{L}$ generally tightens the bounds in a gradual manner, but at the cost of increased computational time, specifically for the confidence intervals.

S.2.4 Extension to Liquidity Constraints

In this section, we illustrate how to compute the identified set in the presence of liquidity constraints as modeled in Section 3.3. To this end, it is useful to explicitly state the analogs of various restrictions in (8) and (11)-(13), and the additional ones from Assumption LC in terms of the richer primitives $\tilde{q}$ and the resulting identified set. Exploiting the relation between $\tilde{q}$ and q given by $q_j(p) = \sum_{e \in \mathcal{E}} \tilde{q}_j(\tilde{p}(p,e), e)$, the data restrictions in (8) correspond to

$$\sum_{e \in \mathcal{E}} \tilde{q}_j(\tilde{p}(p,e), e) = P(D = j | P = p) \ , \tag{S.26}$$

for $j \in \mathcal{J},\ p \in \mathcal{P}_{\text{obs}}$; the logical restrictions that the richer version of demand must be positive and integrate out to one as in (11)-(12) correspond to

$$\tilde{q}(p,e) \geq 0 \tag{S.27}$$

for $j \in \mathcal{J}$, $e \in \mathcal{E}$, $p \in \mathcal{P}$, and

$$\sum_{e \in \mathcal{E}} \sum_{j \in \mathcal{J}} \tilde{q}_j(p_e, e) = 1 \tag{S.28}$$

where $p_e \in \mathcal{P}$ for each $e \in \mathcal{E}$, respectively; the shape restrictions capturing $\tilde{U}_{ij}$ being increasing as in (13) correspond to

$$\tilde{q}_j(p,e) \geq \tilde{q}_j(p', e) \tag{S.29}$$

for $e \in \mathcal{E}$, $p, p' \in \mathcal{P}$ and $j \in \mathcal{J} \setminus \mathcal{J}'$ such that $p_j \geq p'_j$ for $j \in \mathcal{J}'$ and $p_j = p'_j$ for $j \in \mathcal{J} \setminus \mathcal{J}'$; and, finally, the additional restrictions introduced when invoking Assumption LC that takes demand for certain alternatives to be zero when their price is greater than r, which we take to be given by

$$\tilde{q}_j(p,e) = 0 \tag{S.30}$$

for $p \in \mathcal{P}$ such that $p_j \geq r$, and $j \in \mathcal{J}_1$ and $e \in \mathcal{E}$ such that $p'_j > e$ for some $p' \in \mathcal{P}_{\text{obs}} \cup \{p^a, p^b\}$, i.e. when $j \notin C_i(p')$ for some $p' \in \mathcal{P}_{\text{obs}} \cup \{p^a, p^b\}$, which observe is when Assumption LC is invoked in our analysis to replace unaffordability in terms of prices by taking it to be equal to r and obtain the expression in (35) and the data restrictions in (S.26).[1] The admissible space of demand functions in this case then corresponds to

$$\tilde{\mathbf{Q}}_B = \left\{ \tilde{q} \in \tilde{\tilde{\mathbf{F}}} : \tilde{q} \text{ satisfies (S.26)} - \text{(S.30)} \right\} \tag{S.31}$$

where $\tilde{\tilde{\mathbf{F}}}$ denotes the set of all functions from $\mathcal{P} \times \mathcal{E}$ to $\mathbf{R}^{|\mathcal{J}|}$, and in turn, as in (10), the identified set is given by $\tilde{\theta}(\tilde{\mathbf{Q}}_B)$.

To compute the identified set, the main idea remains the same as that in Section 2.5, but performs the analysis conditional on each $e \in \mathcal{E}$ to capture the fact that we have a richer primitive $\tilde{q}$ that accounts for the individual's available income e. Specifically, for each $e \in \mathcal{E}$, we can construct a partition $\tilde{\mathcal{W}}(e)$ as in (18) using the set of prices in (16) given $e \in \mathcal{E}$, i.e.

$$\left\{ \tilde{\mathcal{P}}_l^{a,b}(e) : 1 \leq l \leq |\mathcal{J}| - 1 \right\} \cup \left\{ \{\tilde{p}(p,e)\} : p \in \left\{ p^a, p^b \right\} \cup \mathcal{P}_{\text{obs}} \right\} \tag{S.32}$$

[1] Note that we do not take (S.30) to hold for all $j \in \mathcal{J}_1$, but only if the alternative is unaffordable at certain relevant prices to ensure that no additional restrictions are imposed if it is affordable and so that the analysis in this case remains equivalent to our analysis without liquidity constraints.

with $\tilde{\mathcal{P}}_l^{a,b}(e) = \{p \in \mathcal{P} : p_j = \min\{\tilde{p}_j(p^a, e), \tilde{p}_j(p^b, e) + t\} \text{ for } t \in (\tilde{\Delta}_l^{a,b}(e), \tilde{\Delta}_{l+1}^{a,b}(e)),\ j \in \mathcal{J}\}$, and then consider the finite-dimensional space of functions given by

$$\tilde{\mathbf{Q}}_B^{\mathrm{fd}} = \left\{ q \in \tilde{\mathbf{Q}}_B : \tilde{q}_j(p, e) = \sum_{w \in \tilde{\mathcal{W}}(e)} 1_w(p) \cdot \tilde{\beta}_j(w, e) \text{ for some } \{\tilde{\beta}_j(w, e)\}_{w \in \tilde{\mathcal{W}}} \text{ for each } j \in \mathcal{J},\ e \in \mathcal{E} \right\} ,$$

where $\{\tilde{\beta}_j(w, e) : w \in \tilde{\mathcal{W}}(e),\ e \in \mathcal{E},\ j \in \mathcal{J}\}$ are unknown parameters that parameterizes the space here. As in Section 2.5.2, it can be similarly shown that this space leads to no loss of information, i.e. $\tilde{\theta}(\tilde{\mathbf{Q}}_B) = \tilde{\theta}(\tilde{\mathbf{Q}}_B^{\mathrm{fd}})$, and that the identified set can be computed by solving two analogous versions of the finite-dimensional linear programs in (26), where the objective and restrictions correspond to (35) and (S.26)-(S.30) written in terms of $\tilde{\beta}$, respectively.

Moreover, as in Section 3.1, we can also consider sub-programs based on sub-vectors of $\tilde{\beta}$ that are defined over subsets of $\tilde{\mathcal{W}}(e)$ used to define the parameter for each $e \in \mathcal{E}$, i.e.

$$\tilde{\mathcal{W}}^{\mathrm{r}}(e) = \left\{ w \in \tilde{\mathcal{W}}(e) : w = \prod_{j \in \mathcal{J}} v_{[j]} \text{ for some } v \in \tilde{\mathcal{V}}(e) \right\} ,$$

where $\tilde{\mathcal{V}}(e)$ denotes the partition of the union of the sets in (S.32) satisfying Definition V, and separability assumptions such as

$$\tilde{q}_j(p, e) = \sum_{m \in \mathcal{J}} \tilde{h}_j(p_m, e) \tag{S.33}$$

for each $j \in \mathcal{J}$ and $e \in \mathcal{E}$ for some unknown functions $\{\tilde{h}_{jm} : m \in \mathcal{J}\}$, i.e. requiring demand to be additively separable in prices of all alternatives. As in Section 3.2, we can also impose parametric restrictions such as

$$\tilde{q}_j(p, e) = \sum_{k=0}^{K_j} \tilde{\alpha}_{jk}(e) \cdot b_{jk}(p) \tag{S.34}$$

for each $j \in \mathcal{J}$ and $e \in \mathcal{E}$ for some unknown parameters $\{\tilde{\alpha}_{jk}(e) : 0 \le k \le K_j\}$ and known functions $\{b_{jk} : 0 \le k \le K_j\}$, i.e. require demand to be a linear function of some basis of prices. In particular, observe that this corresponds to the analyses in Sections S.2.2 and S.2.3 but for each value of $e \in \mathcal{E}$. In turn, we can analogously apply the arguments from these sections for each value of $e \in \mathcal{E}$.

To conclude this section, note that the above analysis assumed that the support of available income, $\mathcal{E}$, was discrete. Our analysis, however, can be extended to also allow for it to be continuous. In particular, we can consider a finite partition $\mathcal{F}$ of $\mathcal{E}$ such that the prices $\tilde{p}(e, p)$ for $p \in \{p^a, p^b\} \cup \mathcal{P}_{\mathrm{obs}}$ stay constant for values of $e \in f$ for each $f \in \mathcal{F}$. For example, denoting by $e_1^{a,b} \le \ldots \le e_{\tilde{M}}^{a,b}$ the order values of $\{p_j : p \in \{p^a, p^b\} \cup \mathcal{P}_{\mathrm{obs}},\ j \in \mathcal{J}\} \cup \{-\infty, \infty\}$, we can consider

$$\mathcal{F} = \{(e_1^{a,b}, e_2^{a,b}), [e_2^{a,b}, e_3^{a,b}), \ldots, [e_{\tilde{M}-1}^{a,b}, e_{\tilde{M}}^{a,b})\} .$$

As the value of income in elements of this partition face the same price, it is rich enough to equivalently define the data restrictions as well as the parameter of interest and in turn, similar to the arguments in Section 2.5, capture the same information as that of the underlying continuous income variable. We can then proceed as above by performing the analysis for each value of $f \in \mathcal{F}$. While this allows for a continuous support, note that in practice the cardinality of $\mathcal{F}$ can be quite large and hence make the resulting optimization problem very high dimensional and difficult to compute—this arises in our application as $|\mathcal{J}|$ is large. This is why in our empirical analysis we instead focus on a discretized version of $\mathcal{E}$, where we can choose it to have a low cardinality so that the optimization problems are tractable.

S.3 Statistical Inference

In this section, we describe how we construct confidence intervals for our parameters in our empirical analysis in Section 4 using the bootstrap procedure from Bugni et al. (2017). To this end, let $\{(D_i, P_i) : 1 \leq i \leq N\}$ denote our sample of N observations, assumed to be independently and identically distributed, on which our statistical tests are based.

Recall that our parameters of interest are generally bounded across our various specifications, where the lower and upper bounds are given by minimization and maximization problems, respectively. We construct confidence intervals such that each point in these bounds lies in the interval with probability at least $(1 - \alpha)$ for some pre-specified value of $\alpha \in (0, 1)$. In order to describe the common procedure that we use across all the parameters and specifications, it is useful to first define the common structure present in all these cases. To this end, note that each point in the bounds across these cases can be written as $c'\pi$ for some vector $\pi \in \mathbf{R}^{d_\pi}$ of dimension d_π that satisfies

$$A_1\pi = b_1 \ , \tag{S.35}$$

$$A_2\pi \leq b_2 \ , \tag{S.36}$$

where π corresponds to the optimizing variable in the minimization and maximization problems, c corresponds to the vector defining the objective in these problems that depends on the choice of parameter, (S.35) capture the restrictions imposed on the optimizing variable by the observed shares through (8), and (S.36) capture the restrictions imposed by the shape restrictions in (11)-(13). Note that across these cases the values c, A_1, A_2 and b_2 are known and deterministic, and only b_1 needs to be estimated as it corresponds to the observed enrollment shares.

We construct confidence intervals for the various parameters across the various specifications by test inversion. In particular, we test the null hypothesis at level α that there exists a π satisfying the restrictions in (S.35) and (S.36) such that $c'\pi = \theta_0$ for some given value of $\theta_0 \in \mathbf{R}$. Confidence intervals are then constructed by collecting the values of θ_0 that are not rejected.

We test this null hypothesis using a bootstrap procedure from Bugni et al. (2017), who show it can have several desirable theoretical properties that account for the fact that the parameter of interest is generally partially identified. Our above setup can be mapped into their general framework, given which the test procedure can be described in the following steps. First, compute the test statistic

$$TS_N(\theta_0) = N \cdot \min_{\pi}(A_1\pi - \hat{b}_1)'\hat{\Sigma}^{-1}(A_1\pi - \hat{b}_1)$$

subject to π satisfying $c'\pi = \theta_0$ and the restrictions in (S.36), where $\hat{b}_1$ corresponds to the empirical counterpart of b_1 using the data and $\hat{\Sigma}$ corresponds to a diagonal matrix consisting of the sample variances of the entries of $\hat{b}_1$. Second, for $l = 1, \ldots, B$, compute the so-called minimum resampling bootstrap test statistics of the form

$$TS_{l,N}^{MR}(\theta_0) = \min\left\{TS_{l,N}^{DR}(\theta_0), TS_{l,N}^{PR}(\theta_0)\right\}$$

where $TS_{l,N}^{DR}(\theta_0) = N \cdot (\hat{b}_1 - \hat{b}_{l,1})'\hat{\Sigma}^{-1}(\hat{b}_1 - \hat{b}_{l,1})$ and

$$TS_{l,N}^{PR}(\theta_0) = N \cdot \min_{\pi}\left(\hat{b}_1 - \hat{b}_{l,1} + \frac{1}{\kappa_N}(A_1\pi - \hat{b}_1)\right)'\hat{\Sigma}^{-1}\left(\hat{b}_1 - \hat{b}_{l,1} + \frac{1}{\kappa_N}(A_1\pi - \hat{b}_1)\right)$$

subject to π satisfying $c'\pi = \theta_0$ and the restrictions in (S.36), and $\hat{b}_{l,1}$ corresponds to the analog of $\hat{b}_1$ using the lth boostrap sample drawn with replacement from the data. Here note that κ_N is a tuning parameter that satisfies $\kappa_N \to \infty$ and $\kappa_N/\sqrt{N} \to 0$ as $N \to \infty$. For our empirical results, following Bugni et al. (2017), we take $\kappa_N = \sqrt{\ln N}$. Finally, compute the critical value of the test $\hat{c}(1-\alpha, \theta_0)$ by taking the $(1-\alpha)$-quantile of the distribution of computed bootstrap test statistics. The test procedure is given by $\phi_N(\theta_0) = 1\{TS_N(\theta_0) > \hat{c}(1-\alpha, \theta_0)\}$, i.e we reject if the test statistic is greater than the critical value. Given the above test procedure for a given value of θ_0, we can then use it construct confidence intervals by collecting the set all points we don't reject, namely $\{\theta_0 \in \mathbf{R} : \phi_N(\theta_0) = 0\}$.

S.4 Demand Expressions under Logit Specifications

In this section, we derive the functional forms on demand implied by the Logit specifications considered in Section 5.1. For the specifications without liquidity constraints, it follows from standard arguments that the expression under the Mixed Logit (ML) specification (e.g., Train, 2009, Section 6.1.) is given by

$$q_j(p) = \sum_{x \in \mathcal{X}} \left(\int \frac{e^{\xi_j - (\bar{\gamma}_0 + \bar{\gamma}_1' x + \nu) p_j}}{\sum\limits_{l \in \mathcal{J}} e^{\xi_l - (\bar{\gamma}_0 + \bar{\gamma}_1' x + \nu) p_l}} \phi\left(\frac{\nu}{\sigma}\right) d\nu \right) \text{Prob}[X_i = x] \tag{S.37}$$

for $j \in \mathcal{J}$, and under the Nested Logit (NL) specification (e.g., Train, 2009, Section 4.2.2.) by

$$q_j(p) = \sum_{x \in \mathcal{X}} \frac{e^{\left(\xi_j - (\bar{\gamma}_0 + \bar{\gamma}_1' x) p_j\right)/\lambda_{k(j)}} \left(\sum_{l \in \mathcal{N}_{k(j)}} e^{\left(\xi_l - (\bar{\gamma}_0 + \bar{\gamma}_l' x) p_l\right)/\lambda_{k(j)}} \right)^{\lambda_{k(j)}-1}}{\sum_{k \in \{1,2\}} \left(\sum_{l \in \mathcal{N}_k} e^{\left(\xi_l - (\bar{\gamma}_0 + \bar{\gamma}_l' x) p_l\right)/\lambda_k} \right)^{\lambda_k}} \mathrm{Prob}[X_i = x] \ , \tag{S.38}$$

for $j \in \mathcal{J}$, where $\mathcal{X}$ denotes the support of X_i, ϕ denotes the density of the standard normal distribution, and, for NL, $k(j) = 1$ if $j \in \mathcal{N}_1$ and $k(j) = 2$ otherwise.

For the specifications with liquidity constraints, we can similarly derive expressions for the richer definition of demand $\tilde{q}$ that also accounts for the maximum affordable price. As E_i is continuously distributed for these specifications, the primitive can be more appropriately defined here by

$$\tilde{q}_j(p, e) = \tilde{q}_j(p|e) f_E(e) \ , \tag{S.39}$$

where $\tilde{q}_j(p|e)$ denotes demand conditional on $E_i = e$ and f_E denotes the density of E_i. For both the ML and NL specifications, it follows that given (41), the density equals

$$f_E(e) = \sum_{x \in \mathcal{X}} \mathrm{Prob}[X_i = x] \cdot \phi\left(\frac{e - \tilde{\gamma}_0 - \tilde{\gamma}_1' X_i}{\tilde{\sigma}} \right) \ , \tag{S.40}$$

and by exploiting the fact that E_i is statistically independent of the utilities as $\tilde{\nu}_i$ is independent of ϵ_i and from similar arguments to (S.37) and (S.38), it follows that the conditional demand functions under ML and NL are given by

$$\tilde{q}_j(p|e) = \int \frac{e^{\xi_j - (\bar{\gamma}_0 + \nu)\tilde{p}_j(p,e)}}{\sum_{l \in \mathcal{J}} e^{\xi_l - (\bar{\gamma}_0 + \nu)\tilde{p}_l(p,e)}} \phi\left(\frac{\nu}{\sigma}\right) d\nu \ , \tag{S.41}$$

$$\tilde{q}_j(p|e) = \frac{e^{(\xi_j - \bar{\gamma}_0 \tilde{p}_j(p,e))/\lambda_{k(j)}} \left(\sum_{l \in \mathcal{N}_{k(j)}} e^{(\xi_l - \bar{\gamma}_0 \tilde{p}_l(p,e))/\lambda_{k(j)}} \right)^{\lambda_{k(j)}-1}}{\sum_{k \in \{1,2\}} \left(\sum_{l \in \mathcal{N}_k} e^{(\xi_l - \bar{\gamma}_0 \tilde{p}_l(p,e))/\lambda_k} \right)^{\lambda_k}} \ , \tag{S.42}$$

respectively, where $\tilde{p}$ is defined as in Section 3.3 with $r \to \infty$ as Assumption LC is satisfied here with $r \to \infty$, and which accounts for the fact that alternatives with prices above e will not be affordable and hence cannot be chosen.[2] The parameter of interest in (35) in this case that accounts for E_i being continuous is given by

$$\tilde{\theta}(\tilde{q}) = g^{a,b} \int_{-\infty}^{\infty} \left(\tilde{\Delta}_1^{a,b}(e) f_E(e) + \sum_{l=1}^{|\mathcal{J}|-1} \sum_{j \in \tilde{\mathcal{J}}_{l+1}^{a,b}(e)} \int_{\tilde{\Delta}_l^{a,b}(e)}^{\tilde{\Delta}_{l+1}^{a,b}(e)} \tilde{q}_j\left(\min\{\tilde{p}(p^a, e), \tilde{p}(p^b, e) + t\}, e \right) dt \right) de$$

[2]Formally, for specification ML, Assumption LC is satisfied with $r \to \infty$ only when $\bar{\gamma}_0 + \nu$ is always positive, which is not the case under the normality assumption on ν. In unreported results, we also consider a log normal distribution on γ, which ensures it is always positive, and find similar results to ML. But we focus on the ML specification in the main text as it is the more common specification considered in the literature. Similarly, for specification NL, this is true provided $\bar{\gamma}_0 > 0$, which is the case in our estimated model as revealed by the estimates in Table S.5.

$$+\sum_{j\in J} g_j^a \int_{-\infty}^{\infty} \tilde{q}_j(\tilde{p}(p^a, e), e)de + g_j^b \int_{-\infty}^{\infty} \tilde{q}_j(\tilde{p}(p^b, e), e)de \ . \tag{S.43}$$

For purposes of computation, note that the above expression can be simplified by accounting for the fact that the price function $\tilde{p}$ does not change for certain values of e. Specifically, by taking $e_1^{a,b} \leq \ldots \leq e_{\tilde{M}}^{a,b}$ to denote the ordered values of $\{p_j : j \in \mathcal{J},\ p \in \{p^a, p^b\}\} \cup \{-\infty, \infty\}$, we can rewrite the above expression as

$$\tilde{\theta}(\tilde{q}) = \sum_{m=1}^{\tilde{M}-1} \text{Prob}[E_i \in (e_m^{a,b}, e_{m+1}^{a,b})] \cdot$$
$$\left(g^{a,b} \left(\tilde{\Delta}_1^{a,b}(e_m^{a,b}) + \sum_{l=1}^{|\mathcal{J}|-1} \sum_{j\in \tilde{\mathcal{J}}_{l+1}^{a,b}(e_m^{a,b})} \int_{\tilde{\Delta}_l^{a,b}(e_m^{a,b})}^{\tilde{\Delta}_{l+1}^{a,b}(e_m^{a,b})} \tilde{q}_j \left(\min\{\tilde{p}(p^a, e_m^{a,b}), \tilde{p}(p^b, e_m^{a,b}) + t\} | e_m^{a,b} \right) dt \right) + \right.$$
$$\left. \sum_{j\in J} g_j^a \tilde{q}_j(\tilde{p}(p^a, e_m^{a,b}) | e_m^{a,b}) + g_j^b \tilde{q}_j(\tilde{p}(p^b, e_m^{a,b}) | e_m^{a,b}) \right) ,$$

where

$$\text{Prob}[E_i \in (e_m^{a,b}, e_{m+1}^{a,b})] = \sum_{x\in\mathcal{X}} \text{Prob}[X_i = x] \left(\Phi \left(\frac{e_{m+1}^{a,b} - \tilde{\gamma}_0 - \tilde{\gamma}_1' x}{\tilde{\sigma}} \right) - \Phi \left(\frac{e_m^{a,b} - \tilde{\gamma}_0 - \tilde{\gamma}_1' x}{\tilde{\sigma}} \right) \right)$$

For the purposes of our analysis in Section 5, observe that, for the specification without liquidity constraints, as the underlying parameters are point identified, we can point identify the demand functions q in (S.37) or (S.38), and then the parameter of interest in (7). Similarly, for the specifications with liquidity constraints, as the underlying parameters are point identified, we can alternatively point identify the richer definition of demand in (S.39), and then the parameter of interest in (S.43).

S.5 Additional Details on Empirical Analysis

S.5.1 Data Construction

In this section, we describe how we construct the data used in our empirical analysis in Section 4. The original data sample comes from the replication files for the evaluation of OSP, which are available from the US Department of Education (Wolf et al., 2010). Recall that our analysis focuses on the initial school choice for students who entered the experiment in 2005. Beginning with this subsample, we make the following data-cleaning choices to reach our final analysis data.

Our analysis requires only the prices (as measured by the tuition) of the participating private schools and the school choices of the students (to compute their enrollment shares). For all participating private schools, we observe tuition in either the first or second year of the study, but

Table S.1: Student and family characteristics by voucher receipt

	With voucher	Without voucher	Difference
Mother married (=1)	0.52	0.55	-0.034
Mother years education	12.20	12.25	-0.057
Mother works full time (=1)	0.35	0.38	-0.028
Mother works part time (=1)	0.11	0.11	0.007
Family income ($)	16,725	17,372	-647
HH receives govt transfers (=1)	0.03	0.01	0.016
Household size	4.11	4.14	-0.030
Black (=1)	1.00	1.00	-0.001
Male (=1)	0.50	0.47	0.030
Grade $\leq$ 5 (=1)	0.65	0.65	0.000
Grade 6-8 (=1)	0.21	0.21	-0.000
Grade $\geq$ 9 (=1)	0.14	0.14	0.000
Child learning disabilities (=1)	0.09	0.09	-0.004
Observations	1,090	730	

Table shows mean student and family characteristics by treatment group, weighted using the baseline weights. Observations rounded to the nearest 10. SOURCE: *Evaluation of the DC Opportunity Scholarship Program: Final Report (NCEE 2010-4018)*, U.S. Department of Education, National Center for Education Statistics previously unpublished tabulations.

not necessarily both. If we observe tuition only in the first year, we assume that it was unchanged between the first and second year (recall we use only the second year of data). For school choices, we have that they are missing for around 36% of the students in our data. By a fortunate quirk of the research design, however, participating private schools reported all voucher students to the researchers. Unobserved school choices must therefore be either in non-participating private schools, or government-funded schools. For these students, we assume they enroll in these two groups at the same relative rate that students with observed choices enroll in non-participating private and government-funded schools. Once we obtain these school choices, we weight these observed choices using the baseline weights of the original evaluation—see Wolf et al. (2010, Appendix A.7) for details on how these weights were constructed.

S.5.2 Summary Statistics on School Setting

In this section, we describe additional information on the students and schools in our analysis data. Table S.1 reports mean characteristics of students and their families. Only families making less than 185% of the federal poverty line were eligible for the program, and so unsurprisingly the students are relatively disadvantaged. Approximately 50% of the students' mothers were married, and fewer than 50% were employed at baseline. Family income was slightly less than $17,000. Baseline achievement reflects both positive and negative selection: families chose to participate in the experiment, but

Table S.2: Characteristics of sample schools

	Private	Government-funded	Difference
Share minority	0.73	0.96	-0.227
School size	222.97	325.90	-102.924
Student/teacher ratio	8.92	12.82	-3.897
Catholic (=1)	0.35	0.00	0.354
Other religious (=1)	0.20	0.00	0.200
Secular (=1)	0.45	1.00	-0.554
Gifted program (=1)	0.35	0.39	-0.040
Learning difficulties program (=1)	0.48	0.93	-0.447
Individual tutors available (=1)	0.64	0.69	-0.052
Students tracked by ability (=1)	0.79	0.60	0.192
Remedial classes available (=1)	0.61	0.68	-0.070
Observations	60	160	

Displays school characteristics for private and government-run schools. We do not break out the private schools by participation status because the non-participating schools almost never responded. Observations rounded to the nearest 10. SOURCE: *Evaluation of the DC Opportunity Scholarship Program: Final Report (NCEE 2010-4018)*, U.S. Department of Education, National Center for Education Statistics previously unpublished tabulations.

they also had to be relatively poor to qualify. The table also reveals that voucher recipients and non-recipients are balanced in terms of the various predetermined characteristics. This suggests that the receipt of the voucher was random, in line with Assumption E.

Table S.2 reports characteristics of the private and government-funded schools in the sample. It reveals that the private schools are substantially whiter, have smaller student/teacher ratios, and are more likely to track students by ability. Most strikingly, many of the private schools are religious—35% of them are Catholic, and an additional 20% another religion. In addition, private schools tend to have lower minority shares, lower student/teacher ratios, smaller school sizes, and have more students tracked by ability.

S.6 Proofs

S.6.1 Proof of Proposition 1

The proof of this proposition follows from Bhattacharya (2018, Proposition 1) and Bhattacharya (2018, Theorem 1). We reproduce these proofs here in the context of our setup and notation for completeness. For convenience, we drop the i sub-index here.

To see why the variable $B^{a,b}$ exists and is unique, note first that the right hand side of (4) is continuous in $B^{a,b}$ as U_j is a continuous function for each $j \in \mathcal{J}$. In addition, since $Y - p_j^a >$

$Y-p_j^b-B^{a,b}$ and $Y-p_j^a<Y-p_j^b-B^{a,b}$ for $j\in\mathcal{J}$ when $B^{a,b}>\Delta_{|\mathcal{J}|}^{a,b}$ and $B^{a,b}<\Delta_1^{a,b}$, respectively, note that it follows from the fact that U_j is strictly increasing for each $j\in\mathcal{J}$ that if $B^{a,b}>\Delta_{|\mathcal{J}|}^{a,b}$ then the right hand side of (4) is strictly smaller than its left hand side, whereas if $B^{a,b}<\Delta_1^{a,b}$ then the right hand side will be strictly greater than the left hand side. From these two points, it then follows by the intermediate value theorem that there exists a $B^{a,b}\in[\Delta_1^{a,b},\Delta_{|\mathcal{J}|}^{a,b}]$ such that the right hand side equals the left hand side, i.e. a solution to (4) exists. Furthermore, given that U_j is strictly increasing for each $j\in\mathcal{J}$, it follows that the solution must be unique.

To see why the average value of $B^{a,b}$ is given by (6), note from above that $B^{a,b}\in[\Delta_1^{a,b},\Delta_{|\mathcal{J}|}^{a,b}]$ and hence that $\text{Prob}[B^{a,b}\le t]=0$ for $t<\Delta_1^{a,b}$ and $\text{Prob}[B^{a,b}\le t]=1$ for $t\ge\Delta_{|\mathcal{J}|}^{a,b}$. To calculate this probability for $t\in[\Delta_1^{a,b},\Delta_{|\mathcal{J}|}^{a,b})$, note that since U_j is strictly increasing for each $j\in\mathcal{J}$, we have that $B^{a,b}\le t$ is equivalent to $\max_{j\in\mathcal{J}}U_j\left(Y-p_j^a\right)\ge\max_{j\in\mathcal{J}}U_j\left(Y-p_j^b-t\right)$. It follows that for $t\in[\Delta_l^{a,b},\Delta_{l+1}^{a,b})$ for $l=1,\ldots,|\mathcal{J}|-1$, we have

$$\begin{aligned}
\text{Prob}[B^{a,b}\le t]&=\sum_{j\in\mathcal{J}}\text{Prob}\left[U_j(Y-p_j^a)\ge\max\left\{\max_{m\in\mathcal{J}\setminus\{j\}}U_m(Y-p_m^a),\ \max_{m\in\mathcal{J}}U_m(Y-p_m^b-t)\right\}\right]\\
&=\sum_{j\in\mathcal{J}\setminus\mathcal{J}_{l+1}^{a,b}}\text{Prob}\left[U_j(Y-p_j^a)\ge\max\left\{\begin{array}{c}\max\limits_{m\in\mathcal{J}\setminus\mathcal{J}_{l+1}^{a,b}\cup\{j\}}U_m(Y-p_m^a),\\ \max\limits_{m\in\mathcal{J}_{l+1}^{a,b}}U_m(Y-p_m^b-t)\end{array}\right\}\right]\\
&=\sum_{j\in\mathcal{J}\setminus\mathcal{J}_{l+1}^{a,b}}q_j\left(\min\{p^a,p^b+t\}\right)\ ,
\end{aligned}$$

where the second equality follows from the fact that U_j is strictly increasing for each $j\in\mathcal{J}$ along with $Y-p_j^b-t\ge Y-p_j^a$ for $j\in\mathcal{J}_{l+1}^{a,b}$ and $Y-p_j^b-t\le Y-p_j^a$ for $j\in\mathcal{J}\setminus\mathcal{J}_{l+1}^{a,b}$, and the final equality follows from the definition of the average demand functions along with noting that $\min\{p_j^a,p_j^b+t\}=p_j^b+t$ for $j\in\mathcal{J}_{l+1}^{a,b}$ and $\min\{p_j^a,p_j^b+t\}=p_j^a$ for $j\in\mathcal{J}\setminus\mathcal{J}_l^{a,b}$. Finally, given that $B^{a,b}$ is a positive random variable, we have that its expectation is given by

$$E[B^{a,b}]=\int_0^\infty\left[1-\text{Prob}[B^{a,b}\le t]\right]dt\ ,$$

it follows from the above characterization of $\text{Prob}[B^{a,b}\le t]$ along with noting that

$$\sum_{j\in\mathcal{J}_{l+1}^{a,b}}q_j\left(\min\{p^a,p^b+t\}\right)=1-\sum_{j\in\mathcal{J}\setminus\mathcal{J}_{l+1}^{a,b}}q_j\left(\min\{p^a,p^b+t\}\right)\ ,$$

that the average value of $B^{a,b}$ is given by (6). This concludes the proof.

S.6.2 Proof of Proposition 2

In order to prove the proposition, we need to show that $\Theta\subseteq\Theta_B$ and $\Theta_B\subseteq\Theta$, and that $\Theta_B=[\underline{\theta}_B,\bar{\theta}_B]$ if $\mathbf{B}$ is non-empty. The proof of the first two parts, i.e. $\Theta\subseteq\Theta_B$ and $\Theta_B\subseteq\Theta$, respectively follow directly from the first two parts of Proposition S.2 because $\mathbf{Q}_B$ corresponds to a special case of $\mathbf{Q}_S$ in (S.7) when taking L_j to be a singleton set such as $\{j\}$ and $\mathcal{J}_{jl} = \mathcal{J}$ in (S.6) for each $j \in \mathcal{J}$. It remains to prove the third part, i.e. $\Theta_B = [\underline{\theta}_B, \bar{\theta}_B]$ if $\mathbf{B}$ is non-empty, as the third part of Proposition S.2 only shows that the closure of the set is a closed interval. Here we additionally show that $\mathbf{B}$ is not only convex but also compact, which implies that image of it through θ_B is a closed interval with end points given by (26).

In order to show $\mathbf{B}$ is convex and compact, it is useful to rewrite (S.54)-(S.57) in terms of β when taking L_j to be a singleton set such as $\{j\}$ and $\mathcal{J}_{jl} = \mathcal{J}$. In this case, we have that (S.54)-(S.55) corresponds to

$$\beta_j(w) \geq 0 \quad \text{for each } j \in \mathcal{J} \ , \tag{S.44}$$

$$\sum_{j \in \mathcal{J}} \beta_j(w) = 1 \tag{S.45}$$

for each $w \in \mathcal{W}$; (S.56) corresponds to

$$\beta_j(w) \geq \beta_j(w') \tag{S.46}$$

for each $w, w' \in \mathcal{W}$ and $j \in \mathcal{J} \setminus \mathcal{J}'$ such that $t > t'$ for all $t \in w_{[m]}$, $t' \in w'_{[m]}$ for $m \in \mathcal{J}' \subseteq \mathcal{J}$ and $w_{[m]} = w'_{[m]}$ for $m \in \mathcal{J} \setminus \mathcal{J}'$; and (S.57) corresponds to

$$\beta_j(\{p\}) = \text{Prob}[D_i = j | P_i = p] \ , \tag{S.47}$$

for each $j \in \mathcal{J}$ and $p \in \mathcal{P}_{\text{obs}}$. Hence, it follows that

$$\mathbf{B} = \left\{ \beta \in \mathbf{R}^{d_\beta} : \beta \text{ satisfies (S.44)} - \text{(S.47)} \right\} \ . \tag{S.48}$$

Given that $\mathbf{B}$ is determined by linear inequality restrictions and further that each variable is bounded by the restrictions in (S.44)-(S.45), it follows that it is bounded and compact, which concludes the proof.

S.6.3 Proof of Proposition 3

Given Assumption LC, it follows from (33) that the willingness to pay in (32) can be defined as

$$\max_{j \in \mathcal{J}} \tilde{U}_{ij}(Y_i - \tilde{p}_j(p^a, E_i)) = \max_{j \in \mathcal{J}} \tilde{U}_{ij}(Y_i - \tilde{p}_j(p^b, E_i) - B_i^{a,b}) \ . \tag{S.49}$$

Conditional on $E_i = e \in \mathcal{E}$, it then follows from Proposition 1 as $\tilde{U}_{ij}$ is continuous and strictly increasing for each $j \in \mathcal{J}$ than $B_i^{a,b}$ exists and is unique.

We can similarly apply Proposition 1 to obtain the expression in (34). Specifically, for each $e \in \mathcal{E}$, let

$$\tilde{q}_j(p|e) = \tilde{q}_j(p, e) / \tilde{q}(e) \tag{S.50}$$

denote the demand for alternative $j \in \mathcal{J}$ under prices $p \in \mathcal{P}$ conditional on individuals having a maximum affordable income of e. For each $e \in \mathcal{E}$, we can then apply Proposition 1 to obtain the average willingness to pay conditional on $E_i = e$ given by

$$E[B_i^{a,b}|E_i = e] = \tilde{\Delta}_1^{a,b}(e) + \sum_{l=1}^{|\mathcal{J}|-1} \sum_{j \in \tilde{\mathcal{J}}_{l+1}^{a,b}(e)} \int_{\tilde{\Delta}_l^{a,b}(e)}^{\tilde{\Delta}_{l+1}^{a,b}(e)} \tilde{q}_j \left(\min\{\tilde{p}(p^a, e), \tilde{p}(p^b, e) + t\} | e \right) dt \; . \tag{S.51}$$

Noting that $E[B_i^{a,b}] = \sum_{e \in \mathcal{E}} \tilde{q}(e) \cdot E[B_i^{a,b}|E_i = e]$ and then plugging in the expression from (S.51) and using the relation in (S.50) gives the expression in (34), which completes the proof.

S.6.4 Proof of Proposition S.1

To see why the restrictions in (S.1) imply those in (S.46), consider $w, w' \in \mathcal{W}$ such that $t > t'$ for all $t \in w_{[j]}$, $t' \in w'_{[j]}$ for $j \in \mathcal{J}' \subseteq \mathcal{J}$ and $w_{[j]} = w'_{[j]}$ for $j \in \mathcal{J} \setminus \mathcal{J}'$. In this case, note that $\mathcal{J}^{>}_{w,w'} = \mathcal{J}'$, $\mathcal{J}^{>}_{w',w} = \emptyset$, and $\mathcal{J}^{=}_{w,w'} = \mathcal{J} \setminus \mathcal{J}'$. Then, taking $\mathcal{J}^{\dagger} = \{j\}$ for each $j \in \mathcal{J}^{=}_{w,w'}$ in (S.1) implies that (S.46) holds for each $j \in \mathcal{J}'$. To see why the restrictions in (S.46) imply those in (S.1), consider $w, w' \in \mathcal{W}$ as well as a $w'' \in \mathcal{W}$ such that we have $w''_{[j]} = w'_{[j]} = w_{[j]}$ for $j \in \mathcal{J} \setminus \left(\mathcal{J}^{>}_{w,w'} \cup \mathcal{J}^{>}_{w',w} \right)$, $w''_{[j]} = w_{[j]}$ for $j \in \mathcal{J}^{>}_{w,w'}$, and $w''_{[j]} = w'_{[j]}$ for $j \in \mathcal{J}^{>}_{w',w}$. Since this implies that $t'' > t$ for all $t \in w_{[j]}$, $t'' \in w''_{[j]}$ for $j \in \mathcal{J}^{>}_{w',w}$ and $w_{[j]} = w''_{[j]}$ for $j \in \mathcal{J} \setminus \mathcal{J}^{>}_{w',w}$, it follows from (S.46) that

$$\beta_j(w'') \geq \beta_j(w) \tag{S.52}$$

for each $j \in \mathcal{J}^{=}_{w,w'} \cup \mathcal{J}^{>}_{w,w'}$. Similarly, since it also implies that $t'' > t'$ for all $t \in w'_{[j]}$, $t'' \in w''_{[j]}$ for $j \in \mathcal{J}^{>}_{w,w'}$ and $w'_{[j]} = w''_{[j]}$ for $j \in \mathcal{J} \setminus \mathcal{J}^{>}_{w,w'}$, it also follows from (S.46) that

$$\beta_j(w'') \geq \beta_j(w') \tag{S.53}$$

for each $j \in \mathcal{J}^{=}_{w,w'} \cup \mathcal{J}^{>}_{w',w}$. Then, for each $\mathcal{J}^{\dagger} \subseteq \mathcal{J}^{=}_{w,w'}$, this implies that (S.1) holds as

$$\begin{aligned}
\sum_{j \in \mathcal{J}^{>}_{w',w} \cup \mathcal{J}^{\dagger}} \beta_j(w') \leq \sum_{j \in \mathcal{J}^{>}_{w',w} \cup \mathcal{J}^{\dagger}} \beta_j(w'') &= 1 - \sum_{j \in \mathcal{J} \setminus \left(\mathcal{J}^{>}_{w',w} \cup \mathcal{J}^{\dagger} \right)} \beta_j(w'') \\
&\leq 1 - \sum_{j \in \mathcal{J} \setminus \left(\mathcal{J}^{>}_{w',w} \cup \mathcal{J}^{\dagger} \right)} \beta_j(w) = \sum_{j \in \mathcal{J}^{>}_{w',w} \cup \mathcal{J}^{\dagger}} \beta_j(w)
\end{aligned}$$

where the first inequality follows from (S.53), the second equality follows from (S.45), the third inequality from (S.52), and the final equality from (S.45).

S.6.5 Proof of Proposition S.2

To prove the proposition, we need to show that $\Theta \subseteq \Theta_S$ and $\Theta_S \subseteq \Theta$, and that $\text{closure}(\Theta_S) = [\underline{\theta}_S, \bar{\theta}_S]$ if $\mathbf{S}$ is non-empty. Below, we divide the proof into three parts respectively showing each of these statements. First, we show $\Theta \subseteq \Theta_S$, i.e for every $\theta_0 \in \Theta$ there exists a $\psi \in \mathbf{S}$ such that $\theta_S(\psi) = \theta_0$. Second, we show $\Theta_S \subseteq \Theta$, i.e. for every $\theta_0 \in \Theta_S$ there exists a $q \in \mathbf{Q}_S$ such that $\theta(q) = \theta_0$. Third, we show that if $\mathbf{S}$ is non-empty then it its closure is given by $\Theta_S = [\underline{\theta}_S, \bar{\theta}_S]$.

Before proceeding, it is useful to first explicitly characterize $\mathbf{S}$ and θ_S in terms of ψ. To this end, note that $\mathbf{S}$ corresponds to all ψ such that the corresponding q determined by (S.8) satisfy (11)-(12), (13) and (8). Given that (S.8) requires q to be a constant valued function over $w \in \mathcal{W}$, it follows that (11)-(12), (13) and (8) need to be satisfied across only values in $\mathcal{W}$. In particular, observe that (11) and (12) equivalently correspond to

$$\sum_{l \in \mathcal{L}_j} \psi_{jl}(w_{[\mathcal{J}_{jl}]}) \geq 0 \;\; \text{for each } j \in \mathcal{J} \; , \tag{S.54}$$

$$\sum_{j \in \mathcal{J}} \sum_{l \in \mathcal{L}_j} \psi_{jl}(w_{[\mathcal{J}_{jl}]}) = 1 \tag{S.55}$$

for each $w \in \mathcal{W}$, and, given the way $\mathcal{W}$ was constructed, (13) corresponds to

$$\sum_{l \in \mathcal{L}_j} \psi_{jl}(w_{[\mathcal{J}_{jl}]}) \geq \sum_{l \in \mathcal{L}_j} \psi_{jl}(w'_{[\mathcal{J}_{jl}]}) \tag{S.56}$$

for each $w, w' \in \mathcal{W}$ and $j \in \mathcal{J} \setminus \mathcal{J}'$ such that $t > t'$ for all $t \in w_{[m]}$, $t' \in w'_{[m]}$ for $m \in \mathcal{J}' \subseteq \mathcal{J}$ and $w_{[m]} = w'_{[m]}$ for $m \in \mathcal{J} \setminus \mathcal{J}'$. Similarly, since $\{p\} \in \mathcal{W}$ for each $p \in \mathcal{P}_{\text{obs}}$ given the way $\mathcal{W}$ was constructed, observe that (8) corresponds to

$$\sum_{l \in \mathcal{L}_j} \psi_{jl}\left(\{p_{[\mathcal{J}_{jl}]}\}\right) = \text{Prob}\left[D_i = j | P_i = p\right] \; , \tag{S.57}$$

for each $j \in \mathcal{J}$ and $p \in \mathcal{P}_{\text{obs}}$. Then, it follows $\mathbf{S}$ can be written as

$$\mathbf{S} = \left\{\psi \in \mathbf{R}^{d_\psi} : \psi \text{ satisfies (S.54)} - \text{(S.57)}\right\} \; . \tag{S.58}$$

For the parameter of interest, observe that (6) can be written in terms of ψ as

$$E[B_i^{a,b}] = \Delta_1^{a,b} + \sum_{l=1}^{|\mathcal{J}|-1} \sum_{j \in \mathcal{J}_{l+1}^{a,b}} \int_{\Delta_l^{a,b}}^{\Delta_{l+1}^{a,b}} q_j\left(\min\{p^a, p^b + t\}\right) dt \; , \tag{S.59}$$

$$= \Delta_1^{a,b} + \sum_{l=1}^{|\mathcal{J}|-1} \sum_{j \in \mathcal{J}_{l+1}^{a,b}} \sum_{v \in \mathcal{V}_l^{a,b}} \int_{\underline{t}_v}^{\bar{t}_v} q_j\left(\min\{p^a, p^b + t\}\right) dt \; , \tag{S.60}$$

$$= \Delta_1^{a,b} + \sum_{l=1}^{|\mathcal{J}|-1} \sum_{j \in \mathcal{J}_{l+1}^{a,b}} \sum_{v \in \mathcal{V}_l^{a,b}} (\bar{t}_v - \underline{t}_v) \sum_{l \in \mathcal{L}_j} \psi_{jl}(w(v)_{[\mathcal{J}_{jl}]}) \tag{S.61}$$

where $\mathcal{V}_l^{a,b} \subseteq \mathcal{V}$ is such that $\bigcup_{v \in \mathcal{V}_l^{a,b}} v = \mathcal{P}_l^{a,b}$, which exists given Definition V(i), and $w(v) \equiv \prod_{j \in \mathcal{J}} v_{[j]} \in \mathcal{W}$ given $v \in \mathcal{V}$. In particular, the first line simply recalls (6), the second line follows from rewriting it in terms of the collection of sets $\mathcal{V}$ and exploiting the fact that given the form of $\mathcal{P}_l^{a,b}$, we have that, for $v \in \mathcal{V}_l^{a,b}$, $v = \{p \in \mathcal{P} : p = \min\{p^a, p^a + t\}$ for $t \in (\underline{t}_v, \bar{t}_v)$ or $t \in [\underline{t}_v, \bar{t}_v)$ or $t \in (\underline{t}_v, \bar{t}_v]\}$ for some values of $\underline{t}_v$ and $\bar{t}_v$, and the third line then directly follows from substituting in the equation from (S.8). In turn, along with the fact that $\{p^b\}, \{p^a\} \in \mathcal{W}$ given how $\mathcal{W}$ was constructed, observe that we have that (7) can be written in terms of ψ as follows

$$\theta_S(\psi) = g^{a,b}\Delta_1^{a,b} + g^{a,b} \sum_{l=1}^{|\mathcal{J}|-1} \sum_{j \in \mathcal{J}_{l+1}^{a,b}} \sum_{v \in \mathcal{V}_l^{a,b}} (\bar{t}_v - \underline{t}_v) \sum_{l \in \mathcal{L}_j} \psi_{jl}(w(v)_{[\mathcal{J}_{jl}]})$$
$$+ \sum_{j \in \mathcal{J}} \left(g_j^a \sum_{l \in \mathcal{L}_j} \psi_{jl}\left(\{p^a_{[\mathcal{J}_{jl}]}\}\right) + g_j^b \sum_{l \in \mathcal{L}_j} \psi_{jl}\left(\{p^b_{[\mathcal{J}_{jl}]}\}\right) \right) . \tag{S.62}$$

Given the explicit characterizations of $\mathbf{S}$ and θ_S, we now proceed to presenting the proofs of each of the three parts.

Part 1: Since $\theta_0 \in \Theta$, there exists by definition a $q \in \mathbf{Q}$ such that $\theta(q) = \theta_0$. Using this q, we construct a $\psi^\dagger$ such that $\psi^\dagger \in \mathbf{S}$ and $\theta_S(\psi^\dagger) = \theta_0$. In particular, given that q satisfies (S.6), we take $\psi^\dagger$ to be such that

$$\psi^\dagger_{jl}(w_{[\mathcal{J}_{jl}]}) = \int_0^1 h_{jl}\left(p_{[\mathcal{J}_{jl}]}(t, w)\right) dt \tag{S.63}$$

for each $w \in \mathcal{W}$, $j \in \mathcal{J}$ and $l \in \mathcal{L}_j$, where, for $t \in (0,1)$, $p(t,w) = (p_j(t,w) : j \in \mathcal{J})$ with $p_j(t,w) = \underline{w}_{[j]} + (\bar{w}_{[j]} - \underline{w}_{[j]})t$, and $\bar{w}_{[j]} = \sup\{t : t \in w_{[j]}\}$ and $\underline{w}_{[j]} = \inf\{t : t \in w_{[j]}\}$.

To show $\psi^\dagger \in \mathbf{S}$, we need to show it satisfies the restrictions in (S.54)-(S.57). The restriction in (S.54) is satisfied for each $w \in \mathcal{W}$ and $j \in \mathcal{J}$ as

$$\sum_{l \in \mathcal{L}_j} \psi^\dagger_{jl}(w_{[\mathcal{J}_{jl}]}) = \sum_{l \in \mathcal{L}_j} \int_0^1 h_{jl}\left(p_{[\mathcal{J}_{jl}]}(t,w)\right) dt = \int_0^1 q_j\left(p(t,w)\right) dt \geq 0 ,$$

where the equalities follow from (S.63) and (S.6), respectively, and the inequality from (11). Similarly, the restriction in (S.55) is satisfied for each $w \in \mathcal{W}$ and $j \in \mathcal{J}$ as

$$\sum_{j \in \mathcal{J}} \sum_{l \in \mathcal{L}_j} \psi^\dagger_{jl}(w_{[\mathcal{J}_{jl}]}) = \int_0^1 \sum_{l \in \mathcal{L}_j} h_{jl}\left(p_{[\mathcal{J}_{jl}]}(t,w)\right) dt = 1 ,$$

where the equalities follows from (S.63) and (12), respectively. To see why (S.56) is satisfied, take $w, w' \in \mathcal{W}$ such that $t > t'$ for all $t \in w_{[j]}$, $t' \in w'_{[j=]}$ for $j \in \mathcal{J}' \subseteq \mathcal{J}$ and $w_{[j]} = w'_{[j]}$ for $j \in \mathcal{J} \setminus \mathcal{J}'$. For $t \in (0,1)$, observe that $p_j(t,w) > p_j(t,w)$ for $j \in \mathcal{J}'$ and $p_j(t,w) = p_j(t,w')$ for $j \in \mathcal{J} \setminus \mathcal{J}'$. In turn, it follows from (13) that

$$q_j(t,w) \geq q_j(t,w')$$

for each $t \in (0,1)$ and $j \in \mathcal{J} \setminus \mathcal{J}'$. Taking the integral over $t \in (0,1)$ and then rewriting using (S.6) and (S.63), it directly follows that (S.56) is satisfied. Finally, for the restriction in (S.57), observe that it is satisfied as

$$\sum_{l \in \mathcal{L}_j} \psi^{\dagger}_{jl}\left(\{p_{[\mathcal{J}_{jl}]}\}\right) = \sum_{l \in \mathcal{L}_j} h_{jl}\left(p_{[\mathcal{J}_{jl}]}\right) = q_j(p) = P_{j|0} \ ,$$

for each $j \in \mathcal{J}$ and $p \in \mathcal{P}_{\text{obs}}$, where the equalities follow from (S.63), (S.6), and (8), respectively.

To show $\psi^{\dagger}$ satisfies $\theta_S(\psi^{\dagger}) = \theta_0$, note that it is sufficient to show $\theta_S(\psi^{\dagger}) = \theta(q)$ as $\theta(q) = \theta_0$. To this end, observe that the various components in (S.60) can be written in terms of $\psi^{\dagger}$ as

$$\begin{aligned}
\int_{\underline{t}_v}^{\bar{t}_v} q_j\left(\min\{p^a, p^b + t\}\right) dt &= (\bar{t}_v - \underline{t}_v) \int_0^1 q_j\left(p(t, w(v))\right) dt \ , \\
&= (\bar{t}_v - \underline{t}_v) \int_0^1 \sum_{l \in \mathcal{L}_j} h_{jl}\left(p_{[\mathcal{J}_{jl}]}(t, w(v))\right) dt \ , \\
&= (\bar{t}_v - \underline{t}_v) \sum_{l \in \mathcal{L}_j} \psi^{\dagger}_{jl}\left(w_{[\mathcal{J}_{jl}]}(v)\right) dt
\end{aligned}$$

for each $v \in \mathcal{V}^{a,b}_l$, $1 \leq l \leq |\mathcal{J}| - 1$, where the first equality follows from the change of variables $t = \underline{t}_v + (\bar{t}_v - \underline{t}_v) \cdot t'$ for $t' \in [0,1]$ along with the above definition of $p(t, w(v))$, and the second and third equalities follow from (S.63) and (S.6), respectively; and that

$$\begin{aligned}
\sum_{j \in J} g^a_j q_j(p^a) + g^b_j q_j(p^b) &= \sum_{j \in J} \left(g^a_j \sum_{l \in \mathcal{L}_j} h_{jl}\left(p^a_{[\mathcal{J}_{jl}]}\right) + g^b_j \sum_{l \in \mathcal{L}_j} h_{jl}\left(p^b_{[\mathcal{J}_{jl}]}\right) \right) \\
&= \sum_{j \in J} \left(g^a_j \sum_{l \in \mathcal{L}_j} \psi^{\dagger}_{jl}\left(\{p^a_{[\mathcal{J}_{jl}]}\}\right) + g^b_j \sum_{l \in \mathcal{L}_j} \psi^{\dagger}_{jl}\left(\{p^b_{[\mathcal{J}_{jl}]}\}\right) \right)
\end{aligned}$$

where the equalities follow from (S.6) and (S.63), respectively. Substituting these terms in (7), we obtain the expression in (S.62) from which it follows that $\theta(q) = \theta_S(\psi^{\dagger})$. This completes the first part of the proof.

Part 2: Since $\theta_0 \in \Theta_S$, there exist by definition a $\psi \in \mathbf{S}$ such that $\theta_S(\psi) = \theta_0$ and, in turn, by how Θ_S and θ_S were constructed, a $q^{\dagger} \in \mathbf{Q}^{\text{fd}}_S$ that is related to ψ by the equation in (S.8) such that $\theta(q^{\dagger}) = \theta_S(\psi) = \theta_0$. Since it holds that $\mathbf{Q}^{\text{fd}}_S \subseteq \mathbf{Q}_S$, it follows that $q \in \mathbf{Q}$. This completes the second part of the proof.

Part 3: Given the various linear restrictions that define $\mathbf{S}$ in (S.58), observe that $\mathbf{S}$ is a convex set. In addition, it is also a non-empty set by assumption. Then, since θ_S is a continuous real-valued scalar function, it follows that the image of this function over $\mathbf{S}$ given by Θ_S is a convex and non-empty set on the real line, i.e. an interval whose closure has endpoints given by (26). This completes the final part of the proof.

S.6.6 Proof of Proposition S.3

Note $\mathbf{A}$ is a connected and non-empty set. Then, since θ_A is a continuous real-valued scalar function, it follows that the image of this function over $\mathbf{A}$ given by Θ_A is a connected and non-empty set on the real line, i.e. an interval whose closure has endpoints given by (S.15).

S.7 Additional Table and Figures

Table S.3: Dimension of optimizing variable under different specifications for the programs estimating welfare effects under the status quo voucher amount

Without Liquidity Constraints					
Nonparametric			Parametric Separable, K		
True Baseline	Outer Baseline	Separable	1	2	3
1.01e+69	1,856	65,774	5,619	8,428	11,237
With Liquidity Constraints					
Nonparametric			Parametric Separable, K		
True Baseline	Outer Baseline	Separable	1	2	3
2.26e+71	6,417	200,609	22,476	33,712	44,948

True Baseline true denotes the programs under baseline specification in (26), and Outer Baseline outer denotes the programs under the baseline specification in (28). SOURCE: *Evaluation of the DC Opportunity Scholarship Program: Final Report (NCEE 2010-4018)*, U.S. Department of Education, National Center for Education Statistics previously unpublished tabulations.

Table S.4: Average school characteristics by tuition level for voucher private schools

	Tuition $\leq$ \$3,500	Tuition $>$ \$3,500	Difference
School size	196.103	281.843	-85.740
Student/teacher ratio	13.000	9.860	3.140
Catholic (=1)	0.829	0.159	0.671
Other religious (=1)	0.003	0.256	-0.253
Gifted program (=1)	0.243	0.397	-0.155
Learning difficulties program (=1)	0.446	0.547	-0.101
Individual tutors available (=1)	0.609	0.774	-0.165
Students tracked by ability (=1)	0.763	0.729	0.034
Remedial classes available (=1)	0.646	0.619	0.027
Number of schools	20	50	

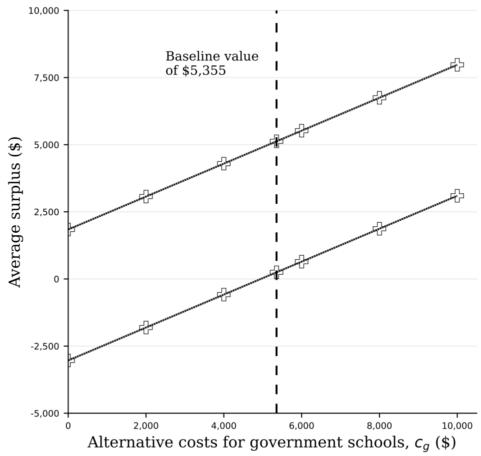
SOURCE: *Evaluation of the DC Opportunity Scholarship Program: Final Report (NCEE 2010-4018)*, U.S. Department of Education, National Center for Education Statistics previously unpublished tabulations. Observations rounded to the nearest 10.

Table S.5: Estimates of the underlying parameters for the various logit specifications

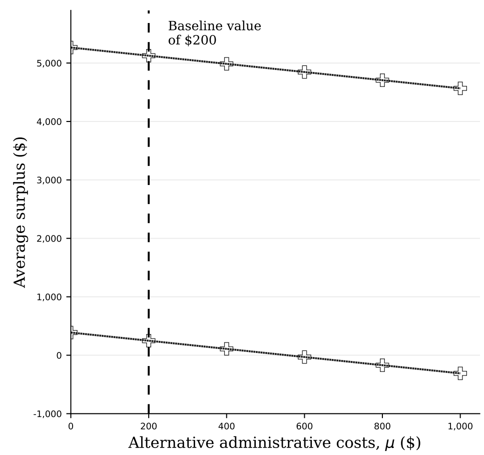
	Specification			
	With LC		Without LC	
	Mixed Logit	Nested Logit	Mixed Logit	Nested Logit
$\bar{\gamma}_0$	12.18	8.45	13.82	7.45
	[2.05]	[0.89]	[8.98]	[0.74]
$\bar{\gamma}_{11}$	-0.88	-0.24		
	[1.15]	[1.10]		
$\bar{\gamma}_{12}$	-0.88	0.01		
	[1.05]	[1.12]		
$\bar{\gamma}_{13}$	-3.09	-2.25		
	[1.23]	[0.96]		
$\tilde{\gamma}_0$			5,094.68	6,342.16
			[1,453.52]	[4,033.85]
$\tilde{\gamma}_1$			1,310.06	966.52
			[2,095.44]	[4,509.57]
$\tilde{\gamma}_2$			7,431.60	1,586.91
			[4,963.80]	[3,904.02]
$\tilde{\gamma}_3$			3,944.48	25,918.62
			[1,257.73]	[10,586.12]
σ	4.37		7.51	
	[1.31]		[7.41]	
λ_1		1.87		2.00
		[0.47]		[0.46]
$\tilde{\sigma}$			3,944.48	5,166.43
			[1,257.73]	[2,969.95]

$\bar{\gamma}_0$, $\bar{\gamma}_{11}$, $\bar{\gamma}_{12}$, $\bar{\gamma}_{13}$, and σ are multiplied by 10^4 so that they can be visually presented in the same scale as the remaining parameters. Standard errors reported in square brackets computed using bootstrap with 1,000 draws. $\bar{\gamma}_{1l}$ corresponds to the coefficient on the lth income bin, i.e. incomes between the l and $l+1$ quartiles of the empirical distribution. The coefficient on the fourth income bin is normalized to 0 due to avoid perfect multicollinearity with the respect to the constant term. λ_2 is normalized to 1 as there is a single alternative in $\mathcal{N}_2$. LC denotes that the model allows students to be liquidity constrained. SOURCE: *Evaluation of the DC Opportunity Scholarship Program: Final Report (NCEE 2010-4018)*, U.S. Department of Education, National Center for Education Statistics previously unpublished tabulations.

Figure S.1: Estimated bounds on average surplus under nonparametric baseline specification for alternative values of government school costs (c_g) and administrative costs (μ)



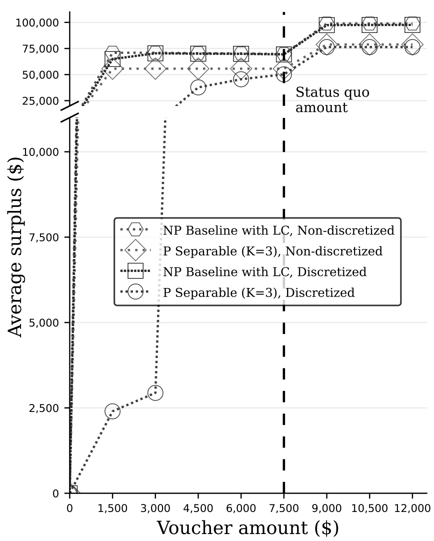
(a) Government school costs



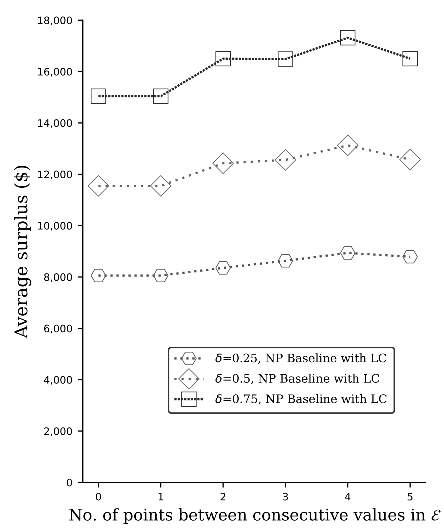
(b) Administrative costs

Observe that as long as we assume that c_g is at least around $5,000, and that γ is at most around $500, we continue to robustly find positive net average benefits. SOURCE: *Evaluation of the DC Opportunity Scholarship Program: Final Report (NCEE 2010-4018)*, U.S. Department of Education, National Center for Education Statistics previously unpublished tabulations.

Figure S.2: Estimated upper bounds on average surplus with and without discretizing E_i and when $R_i = E_i(1+\delta)$ and we increase support of $\mathcal{E}$



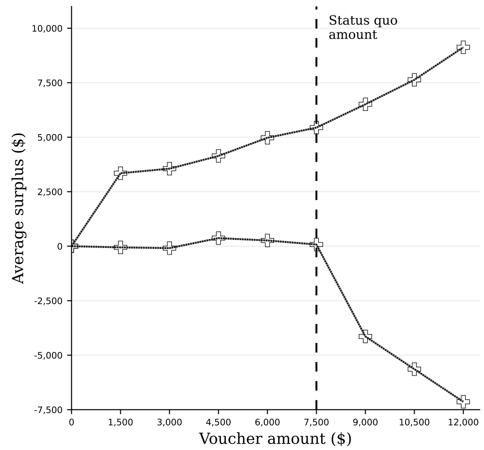
(a) With and without discretization



(b) $R_i = E_i(1+\delta)$ and increasing $|\mathcal{E}|$

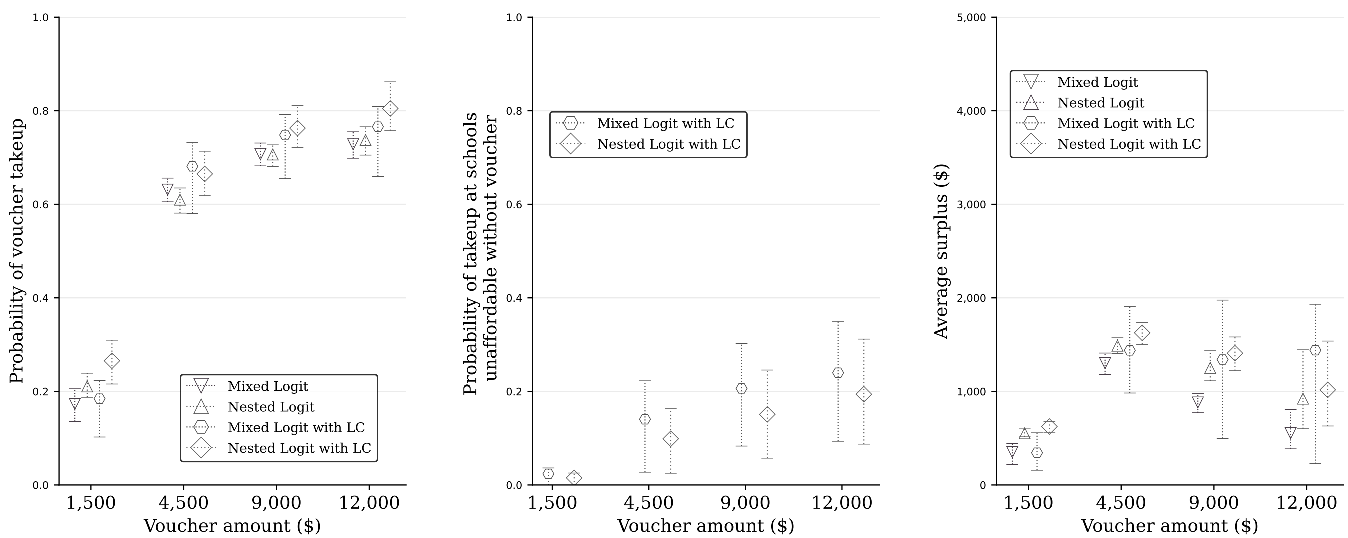
In Panel (b), we increase $|\mathcal{E}|$ by introducing equidistant support points between consecutive ones in $\mathcal{E}$ and between 30,000 and 100,000. Both panels plot only upper bounds as the lower ones with and without discretization remain the same. SOURCE: *Evaluation of the DC Opportunity Scholarship Program: Final Report (NCEE 2010-4018)*, U.S. Department of Education, National Center for Education Statistics previously unpublished tabulations.

Figure S.3: 90% confidence intervals under nonparametric baseline specification on average surplus for a range of voucher amounts



SOURCE: *Evaluation of the DC Opportunity Scholarship Program: Final Report (NCEE 2010-4018)*, U.S. Department of Education, National Center for Education Statistics previously unpublished tabulations.

Figure S.4: Estimates along with 90% confidence intervals for the logit specifications for various demand and welfare parameters at various counterfactual voucher amounts



(a) Probability of voucher takeup

(b) Probability of takeup at schools with tuition at least voucher amount

(c) Average surplus

For the various logit specifications, the markers denote the point estimates and the dashed intervals denote 90% confidence intervals computed using the percentile bootstrap with 1,000 draws. LC denotes that the model allows students to be liquidity constrained. SOURCE: *Evaluation of the DC Opportunity Scholarship Program: Final Report (NCEE 2010-4018)*, U.S. Department of Education, National Center for Education Statistics previously unpublished tabulations.

References

Bhattacharya, D. (2018). Empirical welfare analysis for discrete choice: Some general results. *Quantitative Economics*, **9** 571–615.

Bugni, F. A., Canay, I. A. and Shi, X. (2017). Inference for subvectors and other functions of partially identified parameters in moment inequality models. *Quantitative Economics*, **8** 1–38.

Train, K. E. (2009). *Discrete choice methods with simulation*. Cambridge university press.

Wolf, P., Gutmann, B., Puma, M., Kisida, B., Rizzo, L., Eissa, N., Carr, M. and Silverberg, M. (2010). Evaluation of the dc opportunity scholarship program: Final report. *National Center for Education Evaluation, Institute for Education Sciences Report*, **4018**.