Estimating the Welfare Effects of School Vouchers*

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

Samuel Norris
Harris School of Public Policy
University of Chicago
samnorris@uchicago.edu

November 11, 2020

Abstract

We analyze the welfare effects of voucher provision in the DC Opportunity Scholarship Program (OSP), a school voucher program in Washington, DC, that randomly allocated a voucher to students. To do so, we develop new discrete choice tools to learn about the average willingness to pay for a voucher of a given amount and its average costs in a nonparametric model of school choice. Our tools exploit the insight that these welfare parameters can be expressed as functions of the underlying demand for the different schools. However, while the random allocation of the voucher reveals the value of demand at two prices, the parameters generally depend on their values beyond these prices. Our tools show how to sharply characterize what we can learn when demand is allowed to remain entirely nonparametric or to be parameterized in a flexible manner, both of which imply that the parameters are not necessarily point identified. Applying our tools to the OSP, we find that provision of the status-quo as well as a wide range of counterfactual voucher amounts has a positive net average benefit. These positive results arise due to the presence of many low-tuition schools in the program; removing these schools from the program can result in a negative net average benefit.

KEYWORDS: School vouchers, welfare analysis, discrete choice analysis, demand analysis, Opportunity Scholarship Program.

JEL classification codes: C14, C25, D12, D61, I21.

*We thank Thierry Magnac, Adam Rosen, Alex Torgovitsky, and participants at several seminars and conferences for useful comments and discussions. Vishal Kamat gratefully acknowledges funding from ANR under grant ANR-17-EURE-0010 (Investissements d'Avenir program). First version: arXiv:2002.00103 dated January 31, 2020.

1 Introduction

School vouchers are a topic of active policy debate across several countries. In their basic form, they are government-funded certificates of a certain amount that students can use to offset tuition at any eligible private school of their choice. By reducing the price of private schools and making these schools more affordable, voucher advocates argue that they foster school choice and can make voucher recipients better off (Friedman, 1962).

To empirically investigate this claim, numerous studies have estimated the effects of vouchers on various outcomes using data from programs that randomly allocate vouchers (e.g., Angrist et al., 2002; Dynarski et al., 2018; Howell et al., 2000; Krueger and Zhu, 2004; Mayer et al., 2002; Mills and Wolf, 2017; Muralidharan and Sundararaman, 2015; Wolf et al., 2010). However, as surveyed in Epple et al. (2017), the evidence from these studies is mixed. Some find positive effects, while others find null or even negative effects. Nonetheless, despite this mixed evidence on the effects on outcomes, the data in each of these studies indicate that a non-trivial proportion of recipients choose to use the voucher. By revealed preference arguments, this suggests that recipients in general value vouchers and, in turn, that vouchers may generally be welfare-enhancing. Yet, little empirical work has attempted to quantify these welfare benefits and analyze whether they can justify the costs of providing vouchers.

In this paper, we analyze the potential welfare benefits of the status-quo voucher provided to recipients in the DC Opportunity Scholarship Program (OSP), as well as the benefits of counterfactual vouchers of varying amounts. The OSP provided school vouchers to disadvantaged students in Washington, DC. The vouchers could be used to offset tuition up to \$7,500 at eligible DC private schools and, due to oversubscription, were randomly allocated to students. The data from the program reveal that around 70% of the recipients choose to use the voucher, implying that a large proportion of recipients may in fact value the voucher.

Our analysis starts by showing how to generally use data with random allocation of vouchers to learn about the welfare benefits of providing a voucher of a given amount. To measure the welfare benefits of such a voucher, we use the average of individual willingness to pay for that voucher, i.e. the amount of money an individual is willing to pay to receive that voucher such that they are indifferent to having not received it. This measure provides a natural money metric for the welfare benefits and, in particular, relates to the average compensating variation of the decrease in school prices induced by the voucher. To benchmark these benefits and perform a cost-benefit analysis, we also measure the costs the government faces when providing the voucher.

To characterize what can learn about these quantities using the data, we develop new tools. In our model of school choice, which is entirely nonparametric, each cost and benefit parameter can be expressed as functions of the demand for each school. Given the random allocation of vouchers, the observed choices of recipients and non-recipients reveal the value of demand at two prices, namely the prices with and without the application of the status-quo voucher. However, our parameters of interest generally depend on demand values beyond these two prices. The tools we develop aim to show how to sharply characterize what we can learn about these parameters under a given specification of demand. We consider two demand specifications. In the first, the demand for each school is only nonparametrically restricted to be decreasing with its own price and increasing with the prices of other schools; whereas, in the second, the demand is additionally allowed to be generally parameterized through a flexible functional form restriction on how it varies with prices. For both specifications, we develop easy-to-implement computational procedures that characterize what we can learn about our parameters.

Importantly, our procedures account for the fact that under both specifications there may not exist a single point-identified demand but potentially multiple demand functions consistent with the data. Indeed, this is generally the case unless one focuses attention solely on arguably restrictive parametric specifications of demand. Our procedures generate the unique parameter value in these more restrictive cases, while continuing to generate the set of all parameter values consistent with the multiple admissible demand functions in the more general case. As we discuss below, this generality of our developed tools is a novel feature and of potential interest to discrete choice analysis beyond the voucher setup we consider in this paper.

Applying the developed tools to the OSP data, our estimates reveal that provision of the status-quo voucher amount can have a positive net average benefit. We find that this conclusion is robust to several choices of flexible parametric demand specifications and continues to hold even under the nonparametric specification. In particular, under our most flexible parametric specification, we find that the average benefit net of costs is bounded between $645 and $2,887, whereas, under the nonparametric specification, it is bounded between $213 and $5,088. For a wide range of counterfactual voucher amounts, we find that the voucher program would continue to have a positive net average benefit.

A closer inspection of the data reveals a high number of low-tuition schools in the program. These schools potentially induce a high welfare benefit for recipients relative to the net costs the government faces to fund a voucher when redeemed at them. Indeed, Friedman (1962) argued a key rationale for school vouchers is that they may subsidize private schools that provide services individuals value more efficiently than government-funded schools. Our analysis concludes by investigating the importance of these low-tuition schools in the OSP. We estimate how the welfare effects of the program would change if these schools were removed from the program. Our estimates reveal the presence of such schools plays an essential role in explaining our positive findings; absent schools with tuition of at most $3,500, the program can result in a negative net benefit.

We contribute to the literature on the evaluation of school voucher programs. As highlighted above, most papers in this area estimate the effects of vouchers on outcomes. However, these estimates do not address the welfare implications of these programs. We complement these papers by providing welfare estimates of a specific program, and developing general tools that can be used to analyze programs beyond the one we study. A smaller group of papers uses structural choice models to study various other voucher-related school choice questions of interest (e.g., Allende, 2019; Arcidiacono et al., 2016; Carneiro et al., 2019; Gazmuri, 2019; Neilson, 2013). These analyses require exogenous variation beyond the random allocation of vouchers, and a fully-parameterized model such that the underlying primitives are point-identified. We complement these papers by focusing solely on the welfare effects of the voucher and showing precisely what can be learned without these more demanding data and modeling requirements.

In developing our tools, we exploit recent advances from the literature on nonparametric welfare analysis in an important intermediate step. Specifically, Bhattacharya (2015, 2018) show that the average compensating variation of a price decrease can be nonparametrically expressed as a function of each good's demand. If these demand functions are point identified, then one can directly apply these results. We show how to exploit these results even when the demand functions are not point identified. Recently, Bhattacharya (2019) also derives analytic nonparametric bounds for welfare parameters in such cases with two goods where one of them is a numeraire good. These results however do not straightforwardly extend to the case with multiple goods and prices. We show how the geometry of the parameters in our context can be exploited to propose a computational procedure that similarly derives nonparametric bounds. In addition, we also show how flexible parametric specifications can be incorporated into the analysis. In this direction, we exploit ideas from Mogstad et al. (2018), who show how parametric restrictions can be incorporated in the alternative setting of a treatment effect model.

More broadly, we contribute to the literature on identification in nonparametric discrete choice models. Point identification arguments in such models often require large amounts of exogenous variation in the data (e.g., Berry and Haile, 2009, 2014; Briesch et al., 2010; Chiappori and Komunjer, 2009; Matzkin, 1993). In our setting, however, the random allocation of the voucher induces only binary exogenous variation in prices. We therefore develop tools that allow us to more generally partially identify our parameters. A growing body of papers has similarly developed tools to nonparametrically evaluate various discrete choice questions—such as estimating the effect of different prices and choice sets on demand, characterizing the underlying utility functions, and testing the premise of utility maximization—in setups that permit partial identification (e.g., Chesher et al., 2013; Kamat, 2019; Kitamura and Stoye, 2018; Manski, 2007; Tebaldi et al., 2019). Our analysis contributes to these papers by showing how to evaluate a question of interest not addressed in them, namely the average willingness to pay for a given decrease in prices.

We organize our analysis as follows. Section 2 describes our model of school choice and demand specifications. Section 3 defines our parameters measuring the welfare effects of voucher provision. Section 4 presents our identification analysis. Section 5 presents our empirical results. Section 6 concludes. Proofs and additional details are presented in the Supplementary Appendix.

2 Model and Demand Functions

2.1 Model of School Choice

Suppose the set of schools where individuals can enroll can be partitioned into government-funded schools, and private schools that do and do not participate in the voucher program. Let $\mathcal{J}_g$ denote the set of government-funded schools, $\mathcal{J}_n$ denote the set of private schools not participating in the voucher program, and $\mathcal{J}_v$ denote the set of private schools participating in the voucher program. The status-quo voucher program provides an amount of at most $\tau_{\text{sq}} \in \mathbf{R}_+$ to cover the price (the tuition) for any school in $\mathcal{J}_v$. For the jth school in $\mathcal{J}_v$, let $p_j^* \in \mathbf{R}_+$ denote its original price before applying the voucher and let $p_j(\tau)$ denote its price after applying a voucher of amount $\tau \in \mathbf{R}_+$, where these two prices are related by the relationship $p_j(\tau) = \max\{0, p_j^* - \tau\}$. Under this notation, the original price and that under the status-quo amount for the jth school in $\mathcal{J}_v$ are given by $p_j(0)$ and $p_j(\tau_{\text{sq}})$, respectively. For notational convenience, we use $\mathcal{J}_s = \mathcal{J}_g \cup \mathcal{J}_n \cup \mathcal{J}_v$ to denote the set of all schools. In addition, we take $\mathcal{J}_v = \{1, \ldots, J\}$, where the schools in this list are ordered in terms of their original prices, i.e. $p_1^* \leq \ldots \leq p_J^*$, and we take $p(\tau) = (p_1(\tau), \ldots, p_J(\tau))$ to denote the vector of prices for these schools under a voucher of amount τ.

Each individual i in the population is associated with observables Z_i and D_i, which respectively denote an indicator for whether the individual received a voucher and the school in $\mathcal{J}_s$ where the individual enrolled. We assume that the observed enrollment choice is the product of an underlying individual-level utility maximization decision. To this end, let Y_{ij} denote the individual's underlying disposable income under the jth school in $\mathcal{J}_g$ or $\mathcal{J}_n$, and let $U_{ij}(Y_{ij})$ denote the corresponding indirect utility under that school. For schools in $\mathcal{J}_v$, we can define similar quantities but we need to explicitly account for the role their prices play as they are altered by the receipt of the voucher. Specifically, let $Y_{ij} - p_j$ denote the individual's underlying disposable income under the jth school in $\mathcal{J}_v$ had the price of that school been set to p_j, and let $U_{ij}(Y_{ij} - p_j)$ denote the corresponding indirect utility under that school given that price. Using these quantities, we can define the individual's utility maximizing choice had the prices of the schools in $\mathcal{J}_v$ been set to the vector $p = (p_1, \ldots, p_J)$ by

$$D_i(p) = \begin{cases} \arg\max\limits_{j \in \mathcal{J}_g \cup \mathcal{J}_n} U_{ij}(Y_{ij}) & \text{if} \quad \max\limits_{j \in \mathcal{J}_g \cup \mathcal{J}_n} U_{ij}(Y_{ij}) > \max\limits_{j \in \mathcal{J}_v} U_{ij}(Y_{ij} - p_j) , \\ \arg\max\limits_{j \in \mathcal{J}_v} U_{ij}(Y_{ij} - p_j) & \text{if} \quad \max\limits_{j \in \mathcal{J}_g \cup \mathcal{J}_n} U_{ij}(Y_{ij}) \leq \max\limits_{j \in \mathcal{J}_v} U_{ij}(Y_{ij} - p_j) . \end{cases}$$

The observed enrollment choice is then assumed to be related to the underlying utility maximizing choices and voucher receipt by the relationship

$$D_i = D_i\left(p(\tau_{\text{sq}})\right) \cdot Z_i + D_i\left(p(0)\right) \cdot (1 - Z_i) \ . \tag{1}$$

In our above model for each individual, the most substantial assumption we make is that the voucher program affects utilities solely through a change in disposable income. Apart from this, we do not impose any additional assumptions such as additively separable utilities, how individual and school (observed and unobserved) covariates may affect utilities, or how utilities across schools may be related. In addition, we leave the model entirely nonparametric, and do not assume any functional form on the utilities or any parametric distribution for the unobservables across individuals and schools. It is useful to highlight that this is in contrast to standard differentiated-products demand models (e.g., Berry et al., 1995) that impose a combination of such assumptions, and are often the basis for the structural models of school choice used to study vouchers that we referenced in Section 1.

For the purposes of our analysis, we emphasize that we model solely the variation in school prices induced by the receipt of the voucher, as formally captured by (1). Indeed, as highlighted above, we do not model and impose additional restrictions on how individual and school characteristics affect the enrollment decision. Importantly, this means that our analysis exploits only the binary variation in prices that the voucher induces across individuals—namely, the original school prices and these prices under the status-quo voucher. Along with the nonparametric nature of our model, this discrete variation will generally imply that our various parameters of interest are only partially identified. Our identification analysis in Section 4 captures this point more precisely.

It is useful to highlight certain features of the school choice setting that our model abstracts away from, and their implications for our analysis. We do not explicitly model individual application and school admission decisions that determine where an individual can be admitted and hence enroll. Nonetheless, note that our model allows for the potential scenario that these decisions may imply, namely one where an individual may not be able to be admitted to some schools. Indeed, such a scenario can be relevant in many school choice settings (Agarwal and Somaini, 2020). The model does so implicitly by allowing $U_{ij}(\cdot) = -\infty$ for the schools in which the individual cannot be admitted and hence enroll. In this sense, our model interprets the utility under a given school to be net of whether the individual can be admitted there and, in turn, our resulting analysis should be more appropriately interpreted taking where an individual can be admitted as fixed.

In addition, we also do not model general equilibrium effects of the voucher program that may affect utilities under schools through changes beyond that in the disposable income. Examples of such effects include changes in the set of schools to where an individual can be admitted due to sufficiently large changes in individual application and school admission behaviour; and changes in the utilities under different schools due to changes in school incentives to invest in quality (Allende, 2019; Neilson, 2013) or those in peer composition (Allende, 2019; Gazmuri, 2019). In this sense, our model interprets the utilities as partial equilibrium quantities where these various features remain fixed. Our resulting analysis should therefore be interpreted taking these features as fixed. As we highlight in Section 6, extending our analysis to study these potential general equilibrium type effects is an interesting direction for future research.

2.2 Average Demand Functions

Our analysis is based on the demand functions for the different schools in the sense that we use them to state our assumptions and define our parameters of interest. These functions correspond to the distribution of enrollment choices across individuals for the different schools at each potential price vector. More formally, let $\mathcal{P} = \prod_{j=1}^{J} [0, p_j(0)] \subset \mathbf{R}_+^J$ denote the domain of price vectors for the schools in $\mathcal{J}_v$ over which we define these functions. Then, for a given $p \in \mathcal{P}$, let

$$
\begin{aligned}
q_j(p) &= \text{Prob}\{D_i(p) = j\} \ , \\
q_g(p) &= \text{Prob}\{D_i(p) \in \mathcal{J}_g\} \ , \\
q_n(p) &= \text{Prob}\{D_i(p) \in \mathcal{J}_n\}
\end{aligned}
$$

respectively define the demand for the jth school in $\mathcal{J}_v$, for any school in $\mathcal{J}_g$ and for any school in $\mathcal{J}_n$. Analogously, let $q_j(p|z)$ for $j \in \mathcal{J}_v$, $q_n(p|z)$ and $q_g(p|z)$ respectively define these demand functions conditional on the receipt of the voucher $Z = z \in \{0, 1\}$. Note we only define demand for any school in $\mathcal{J}_g$ and $\mathcal{J}_n$, and not for each specific school in these sets of schools. As we will observe, this is because defining demand over this more parsimonious grouping is sufficient for the definition of our welfare parameters. For notational convenience, let $\mathcal{J} = \{g, n\} \cup \mathcal{J}_v$ denote the set of indices over which the demand functions are defined.

In the following assumption, we state the restrictions we impose on the demand functions under our baseline specification. In particular, note that this specification is entirely nonparametric.

Assumption B. (Baseline)

(i) For each $j \in \mathcal{J}$, $q_j(p|z) = q_j(p)$ for all $p \in \mathcal{P}$ and $z \in \{0, 1\}$.

(ii) For each $j \in \mathcal{J}$, q_j is weakly increasing in p_m for each $m \in \mathcal{J}_v$ such that $m \neq j$.

Assumption B(i) states that the demand functions are invariant to the receipt of the voucher. It follows from this assumption that the underlying demand functions can be uniquely captured by the vector $q \equiv (q_g, q_n, q_1, \ldots, q_J)$ of unconditional demand functions. As a result, in the remainder of our analysis, we focus solely on the unconditional demand; whenever we refer to demand, it is understood we are referring to the unconditional demand. Assumption B(ii) imposes shape restrictions on how demand behaves with the prices of the private schools in the voucher program. In particular, it imposes that for each $p, p' \in \mathcal{P}$ such that $p_j > p'_j$ for $j \in \mathcal{J}' \subseteq \mathcal{J}_v$ and $p_j = p'_j$ for $j \in \mathcal{J}_v \setminus \mathcal{J}'$, we have that

$$
q_j(p) \geq q_j(p') \tag{2}
$$

for each $j \in \mathcal{J} \setminus \mathcal{J}'$. Since by definition we have that

$$
q_j(p) = 1 - \sum_{m \in \mathcal{J} \setminus \{j\}} q_m(p)
$$

for each $j \in \mathcal{J}_v$, note that it directly follows from Assumption B(ii) that q_j is weakly decreasing in p_j for $j \in \mathcal{J}_v$, i.e. the standard shape restriction from demand theory that states demand for each good is weakly decreasing with respect to its own price. Moreover, note that Assumption B(ii) corresponds to a version of the weak substitutes assumption from Berry et al. (2013). While Assumption B imposes restrictions directly on the demand functions, note that these restrictions follow from assumptions on the underlying variables of the model. For example, Assumption B(i) follows from assuming the voucher is randomly allocated, i.e. Z_i is statistically independent of the remaining underlying variables of the model. On the other hand, Assumption B(ii) follows from assuming U_{ij} is weakly increasing for each $j \in \mathcal{J}_s$.

A common approach in the literature on discrete choice analysis is to consider specifications that place parametric functional form restrictions on the demand functions—see, for example, Train (2009, Chapter 2) for a textbook introduction on such parameterizations. These specifications are often chosen to ensure that the demand functions are point identified. In our analysis, we also consider auxiliary specifications that impose such parametric restrictions in addition to those in Assumption B, but we do not restrict attention to only those that ensure point identification. In the following assumption, we state the general class of parametric specifications we consider.

Assumption A. (Auxiliary) For each $j \in \mathcal{J}$,

$$q_j(p) = \sum_{k=0}^{K_j} \alpha_{jk} \cdot b_{jk}(p) \tag{3}$$

for some $\{\alpha_{jk} : 0 \leq k \leq K_j\}$, where $\{b_{jk} : 0 \leq k \leq K_j\}$ denote some known functions.

Assumption A states that the demand functions are linear functions of some known functions of prices, where the variable $\alpha \equiv \left(\alpha'_g, \alpha'_n, \alpha'_1, \ldots, \alpha'_J\right)'$, with $\alpha_j = \left(\alpha_{j1}, \ldots, \alpha_{jK_j}\right)'$ for each $j \in \mathcal{J}$, parameterizes the demand functions. As we further discuss in Section 4.3, this assumption allows for several types of flexible parametric specifications. For example, it allows for those that result in point identification of the demand functions such as

$$q_j(p) - q_g(p) = \alpha_{j0} + \alpha_{j1} \cdot p_j \quad \text{for } j \in \mathcal{J}_v \ , \tag{4}$$

$$q_n(p) - q_g(p) = \alpha_n \ , \tag{5}$$

for some $\{\alpha_{jk} : j \in \mathcal{J}_v,\ 0 \leq k \leq 1\}$ and α_n, i.e. the difference in demand for a given school in $\mathcal{J}_v$ and any school in $\mathcal{J}_g$ is a linear functions of that schools own price and the difference in demand for any school in $\mathcal{J}_n$ and any school in $\mathcal{J}_g$ is constant—see Appendix S.2.1 for details on how this specification imposes restrictions similar to a logit specification and, like the logit, achieves point identification. However, Assumption A also allows for more flexible specifications where demand for each school is a polynomial function of own prices as well as prices of all schools, which do not imply point identification.

3 Welfare Effects of Voucher Provision

In the context of our model, the provision of a voucher can make individuals better off by increasing their disposable income when enrolled in schools in the voucher program. In this section, we define the main parameter of interest of our analysis that aims to quantify these potential welfare benefits. We define this parameter for a generic voucher amount $\tau \in \mathbf{R}_+$. As mentioned below, this generality, by choosing different values of τ, allows us to analyze the welfare effects of the status-quo voucher amount as well as alternative counterfactual voucher amounts.

To quantify the benefit for each individual i, we use a money metric for the welfare gains from the receipt of the voucher. Specifically, we use the amount of money that the individual would pay to receive the voucher or, equivalently, the negative of the compensating variation of the reduction in prices induced by the voucher. Formally, the individual's willingness to pay for a voucher of amount τ is defined by the variable $B_i(\tau)$ that solves

$$\begin{gathered}\max\left\{\max_{j\in\mathcal{J}_g\cup\mathcal{J}_n} U_{ij}(Y_{ij}),\ \max_{j\in\mathcal{J}_v} U_{ij}\left(Y_{ij}-p_j(0)\right)\right\}\\ = \\ \max\left\{\max_{j\in\mathcal{J}_g\cup\mathcal{J}_n} U_{ij}(Y_{ij}-B_i(\tau)),\ \max_{j\in\mathcal{J}_v} U_{ij}\left(Y_{ij}-p_j(\tau)-B_i(\tau)\right)\right\}\ ,\end{gathered} \tag{6}$$

i.e. the amount of money to be subtracted from the individual's income under the receipt of the voucher so that they obtain the same utility as in the absence of the voucher. We then quantify the average benefit of a voucher that provides an amount τ by

$$AB(\tau) = E[B_i(\tau)]\ , \tag{7}$$

i.e. the average willingness to pay across individuals to receive that voucher.

As mentioned, our analysis is based on the fact that our parameters of interest can be written as functions of the demand functions introduced in the previous section. In order to show this for the average benefit parameter defined above, we exploit results from Bhattacharya (2015, 2018), who showed in a more general setup that the average value of a variable such as that defined in (6) can be written as a closed form expression of the demand functions. In the following proposition, we formally state this result in terms of our setup and notation. In the statement of this proposition, we use $j(\tau)$ to denote the jth school in $\mathcal{J}_v$ such that $p_{j(\tau)}(0) < \tau$ and $p_{j(\tau)+1}(0) \geq \tau$, i.e. the last school in $\mathcal{J}_v$ for which the voucher amount τ is strictly greater than the tuition amount. In addition, we take $\{a_l(\tau) : 0 \leq l \leq J\}$ to be a set of values such that $a_0(\tau) = 0$, $a_l(\tau) = p_l(0)$ for $1 \leq l \leq j(\tau)$ and $a_l(\tau) = \tau$ for $l > j(\tau)$.

Proposition 3.1. For each individual i, suppose U_{ij} is continuous and strictly increasing for each

$j \in \mathcal{J}_s$. Then we have that $B_i(\tau)$ defined in (6) exists and is unique, and that

$$E[B_i(\tau)] = \sum_{l=0}^{j(\tau)} \int_{a_l(\tau)}^{a_{l+1}(\tau)} \left(\sum_{j=l+1}^{J} q_j \left(p_1(0), \ldots, p_l(0), p_{l+1}(\tau) + a, \ldots, p_J(\tau) + a\right) \right) da \ . \tag{8}$$

While voucher provision is beneficial for individuals, it can be costly to the government financing the voucher. To benchmark the benefits and perform a cost-benefit analysis, we also consider parameters that measure these potential costs. To this end, observe that the provision of a voucher introduces costs to the government when individuals enroll in a school in the program, but can also bring about savings depending on the costs the government faces at schools where individuals enroll in the absence of the voucher. To formally capture these net costs, let $c_j(\tau)$ denote the cost that the government associates with the jth demand function in $\mathcal{J}$ under a voucher of amount τ. For example, in our baseline empirical analysis, we take

$$c_j(\tau) = \begin{cases} c_g & \text{for } j = g \ , \\ 0 & \text{for } j = n \ , \\ \min\{p_j(0), \tau\} + \gamma \cdot 1\{\tau > 0\} & \text{for } j \in \mathcal{J}_v \ , \end{cases}$$

i.e. the cost associated with each government-funded school is some known value c_g, the cost associated with each private school not participating in the program is zero, and the cost associated with each private school participating in the program is the voucher amount spent to cover tuition plus some known administrative cost γ of operating the program (i.e. charged only when the voucher amount is positive). We then measure the average net costs from the provision of a voucher of amount τ by

$$AC(\tau) = \sum_{j \in \mathcal{J}} c_j(\tau) \cdot q_j(p(\tau)) - \sum_{j \in \mathcal{J}} c_j(0) \cdot q_j(p(0)) \ , \tag{9}$$

i.e. the average costs the government faces when individuals receive the voucher net of those it faces when individuals do not receive the voucher. Along with the average benefit parameter, we can then also define the average surplus parameter, which can be used to perform a cost-benefit analysis. Specifically, for a voucher of amount τ, let

$$AS(\tau) = AB(\tau) - AC(\tau) \tag{10}$$

denote the average surplus of the voucher, i.e. the average benefit across individuals of receiving the voucher net of the average cost for the government of providing that voucher. Note that the average cost parameter is a function of q and, since the average benefit parameter is a function of q, so is the average surplus parameter.

The benefit, cost and surplus parameters we described above were defined for a generic voucher of amount τ. By taking different values of τ, we can evaluate these parameters for both the status-quo voucher amount as well as alternative counterfactual amounts. More specifically, by taking

$\tau = \tau_{\text{sq}}$, we can evaluate these parameters for the status-quo voucher amount, whereas, by taking $\tau = \tau_{\text{c}} \neq \tau_{\text{sq}}$, we can evaluate these parameters for a counterfactual amount of τ_{c}. In our analysis, we also study the difference of the parameters under these amounts, i.e.

$$\Delta AB\left(\tau_{\text{c}}\right) = AB\left(\tau_{\text{c}}\right) - AB\left(\tau_{\text{sq}}\right) \ , \tag{11}$$

$$\Delta AC\left(\tau_{\text{c}}\right) = AC\left(\tau_{\text{c}}\right) - AC\left(\tau_{\text{sq}}\right) \ , \tag{12}$$

$$\Delta AS\left(\tau_{\text{c}}\right) = AS\left(\tau_{\text{c}}\right) - AS\left(\tau_{\text{sq}}\right) \ , \tag{13}$$

which allows us to directly compare the benefit, cost and surplus between the counterfactual and status-quo voucher amounts.

4 Identification Analysis

In the previous section, we described our parameters of interest and noted that each of them was a function of the demand functions. In this section, we study what we can learn about each of these parameters given what we know about the demand functions from the imposed assumptions and the distribution of the data, taken to be known for the purposes of this section.

4.1 General Setup

We begin by formally describing the general setup for the identification analysis we develop below. To this end, let $\theta(q)$ denote a pre-specified parameter of interest from Section 3 that we want to learn about.

Since θ is a known function, it follows what we can learn about our parameter depends on what we know about the function q. As q is defined to be a function whose image is a vector of probabilities, we know by construction that for each $p \in \mathcal{P}$ we have

$$0 \leq q_j(p) \leq 1 \ \text{ for each } j \in \mathcal{J} \ , \tag{14}$$

$$\sum_{j \in \mathcal{J}} q_j(p) = 1 \ , \tag{15}$$

i.e., for all prices, each demand function lies in the unit interval and their sum together equals one. Under our baseline specification, we know that q satisfies Assumption B(ii), i.e. it satisfies the nonparametric shape restrictions stated in (2). Under our auxiliary specifications, we additionally know that q satisfies Assumption A, i.e. it satisfies the parametric restrictions stated in (3). Finally, under both specifications, the distribution of the data across individuals also restricts the values that q can take. Specifically, it follows from (1) and Assumption B(i) that the distribution of the data reveals

$$q_j(p(0)) = \text{Prob}[D_i = j | Z_i = 0] \equiv P_{j|0} \ , \tag{16}$$

$$q_j\left(p\left(\tau_{\text{sq}}\right)\right) = \text{Prob}[D_i = j | Z_i = 1] \equiv P_{j|1} \tag{17}$$

for $j \in \mathcal{J}_v$, and

$$q_j(p(0)) = \text{Prob}[D_i \in \mathcal{J}_j | Z_i = 0] \equiv P_{j|0} \ , \tag{18}$$

$$q_j\left(p\left(\tau_{\text{sq}}\right)\right) = \text{Prob}[D_i \in \mathcal{J}_j | Z_i = 1] \equiv P_{j|1} \tag{19}$$

for $j \in \{g, n\}$, i.e. the enrollment shares across schools conditional on the receipt of voucher reveal the values the demand functions take at the vector of prices with and without the status-quo voucher amount. To summarize the above information on what we know about q, let $\mathbf{F}$ denote the set of all functions from $\mathcal{P}$ to $\mathbf{R}^{|\mathcal{J}|}$. Then, let

$$\mathbf{Q}_B = \{q \in \mathbf{F} : q \text{ satisfies } (14) - (15),\ (2) \text{ and } (16) - (19)\} \tag{20}$$

denote the admissible set of all demand functions that satisfy the various restrictions imposed by the assumptions and data under our baseline specification, and let

$$\mathbf{Q}_A = \{q \in \mathbf{F} : q \text{ satisfies } (14) - (15),\ (2),\ (3) \text{ and } (16) - (19)\} \tag{21}$$

denote the analogous set of such demand functions under our auxiliary specification.

Given what we know about q, our objective is to characterize what we can then learn about our parameter $\theta(q)$. In some cases, observe that there exists a single admissible value of q under the chosen specification. In such cases, it follows that we can exactly learn the value of $\theta(q)$. For example, as we noted before, this is the case under the specification described in (4)-(5). However, under more flexible parametric specifications as well as the baseline nonparametric specification, there generally exist multiple admissible values of q. In these more general cases, it follows that we can learn a set of values that $\theta(q)$ lies in.

Our analysis aims to show what we can learn across both these two cases. We generally do so by showing how to characterize the identified set. Formally, for a given admissible set of demand functions $\mathbf{Q}$, the identified set is defined by

$$\theta(\mathbf{Q}) = \{\theta_0 \in \mathbf{R} : \theta(q) = \theta_0 \text{ for some } q \in \mathbf{Q}\} \equiv \Theta \ , \tag{22}$$

i.e. the image of the set of admissible functions $\mathbf{Q}$ under the function θ. Intuitively, the identified set corresponds to the set of all parameter values that could have been generated by the admissible values of q. By construction, it sharply captures all that we can learn about the parameter given the data and the chosen specification. Indeed, if the parameter is point identified then the identified set corresponds to a single point. Alternatively, if the parameter is partially identified then the identified set corresponds to the sharpest set of all possible parameter values consistent with the data and specification.

In what follows, we develop procedures to compute the identified set under each of our specifications: first, in Section 4.2, under our baseline specification, i.e. Θ in (22) when $\mathbf{Q} = \mathbf{Q}_B$; and then, in Section 4.3, under our auxiliary specification, i.e. Θ in (22) when $\mathbf{Q} = \mathbf{Q}_A$.

4.2 Identified Set under Baseline Nonparametric Specification

In principle, observe that characterizing the identified set corresponds to searching over the various q in $\mathbf{Q}$ and taking their image under the function θ. Under the baseline specification, this problem can be challenging due to the fact that $\mathbf{Q}_B$ is an infinite-dimensional space. Below, we show how to feasibly proceed in this case. In particular, we exploit the idea that we can replace $\mathbf{Q}_B$ by a finite-dimensional space $\mathbf{Q}_B^{\text{fd}}$ without any loss of information with respect to what we can learn about the parameter in the sense that $\theta(\mathbf{Q}_B) = \theta\left(\mathbf{Q}_B^{\text{fd}}\right)$. This allows us to indirectly characterize the identified set by searching only through q in $\mathbf{Q}_B^{\text{fd}}$, which is a finite-dimensional problem and, hence, potentially feasible in practice.

We begin by defining the finite-dimensional $\mathbf{Q}_B^{\text{fd}}$ we consider. Our choice of $\mathbf{Q}_B^{\text{fd}}$ takes q to be constant on some finite partition of the space of prices. The partition is intuitively chosen such that the resulting q based on it is sufficiently rich to equivalently define the various parameters of interest and the data restrictions in (16)-(19) as well as preserve the information provided by the shape restrictions in (2).

In order to define the partition, we need to first define a collection of sets that play a role in the definition of the parameters and data restrictions, and allow the preservation of the information provided by the shape restrictions. To this end, observe that

$$\mathcal{P}_l(\tau) = \{p \in \mathcal{P} : p_j = \min\{p_j(0), p_j(\tau) + a\} \text{ for } a \in [a_l(\tau), a_{l+1}(\tau)] \text{ for each } j \in \mathcal{J}_v\} \tag{23}$$

for $0 \leq l \leq j(\tau)$ correspond to the various sets of prices that play a role in the definition of the parameter $AB(\tau)$, and

$$\{p(0), p\left(\tau_{\text{sq}}\right), p(\tau)\} \tag{24}$$

corresponds to the set of prices that play a role in the definition of the parameter $AC(\tau)$ as well as the data restrictions in (16)-(19). Note it then follows that

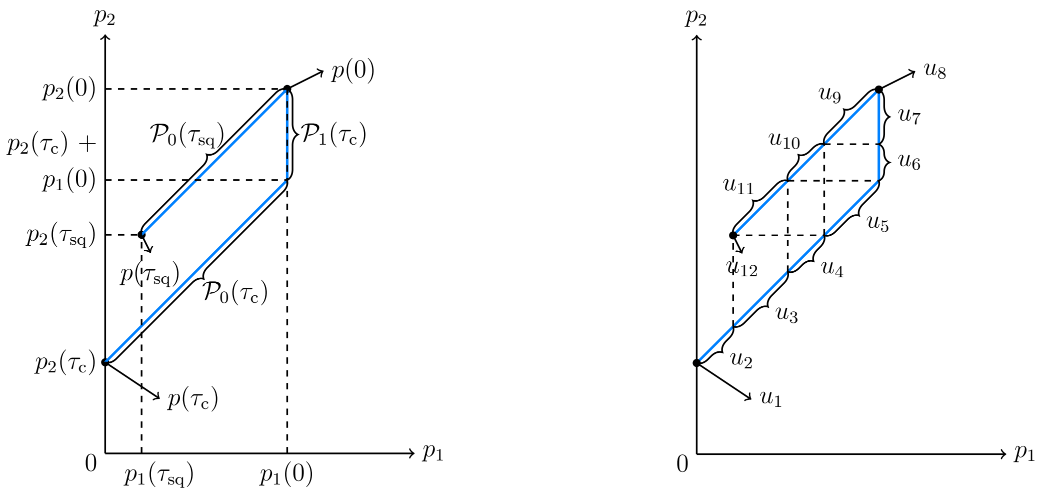
$$\mathcal{P}^* = \bigcup_{l=0}^{j\left(\tau_{\text{sq}}\right)} \mathcal{P}_l\left(\tau_{\text{sq}}\right) \bigcup_{l=0}^{j(\tau_c)} \mathcal{P}_l\left(\tau_{\text{c}}\right) \bigcup \{p(0), p\left(\tau_{\text{sq}}\right), p\left(\tau_{\text{c}}\right)\} \tag{25}$$

corresponds to the subset of $\mathcal{P}$ that plays a role in the definition of all parameters for the status-quo voucher amount and a counterfactual voucher amount of τ_{c} along with the restrictions imposed by the data. Given this set of prices, we define in the following definition the collection of sets that we later use in the definition of the partition.

Definition 4.1. Let $\mathcal{U}$ denote a finite partition of the set of prices $\mathcal{P}^*$ in (25) such that for all $u \in \mathcal{U}$ we have either

(i) $u = \{p \in \mathcal{P} : p_j = \min\{p_j(0), p_j(\tau) + a\} \text{ for } a \in (\underline{a}_u, \bar{a}_u] \text{ or } (\underline{a}_u, \bar{a}_u) \text{ for each } j \in \mathcal{J}_v\}$ where $\underline{a}_u$ and $\bar{a}_u$ are such that $u \subseteq \mathcal{P}_l(t)$ for some $0 \leq l \leq j(\tau)$ and $\tau \in \{\tau_{\text{sq}}, \tau_{\text{c}}\}$; or

Figure 1: Various sets of prices for an example with $J = 2$ and $\tau_{\text{sq}} < p_1(0) < \tau_{\text{c}} < p_2(0)$



(a) Sets that play a role in defining the parameters and data restrictions

(b) A partition $\mathcal{U} \equiv \{u_1, \ldots, u_{12}\}$ of the union of sets in (a) that satisfies Definition 4.1

(ii) $u = \{p(\tau)\}$ for some $\tau \in \{0, \tau_{\text{sq}}, \tau_{\text{c}}\}$,

and for all $u, u' \in \mathcal{U}$ we have either

$$u(j) = u'(j) \quad \text{or} \quad u(j) \cap u'(j) = \emptyset \tag{26}$$

for each $j \in \mathcal{J}_v$, where $u(j) = \{t \in \mathbf{R} : p_j = t \text{ for some } p \in u\}$ for each $u \in \mathcal{U}$ and $j \in \mathcal{J}_v$.

Definition 4.1 states that $\mathcal{U}$ corresponds to a finite partition of $\mathcal{P}^*$, where each element of the partition satisfies specific properties. In particular, Definition 4.1(i)-(ii) states that each element is a connected subset of that in (23) or (24). In addition, it states in (26) that any pair of sets in this partition are such that they either completely overlap or are disjoint in each price coordinate. Intuitively, note that the first property, as the sets in (23) and (24) are specifically based on the parameters and data restrictions, is what ensures that the finite-dimensional q will be sufficiently rich to define the parameters and data restrictions. On the other hand, note that the latter property, which implies that the sets can be ordered and pairwise compared across each price coordinate, is what ensures that the finite-dimensional q will be able to preserve the information provided by the shape restrictions in (2) that are indeed based on pairwise comparisons of prices.

To better understand these various set of prices, Figure 1(a) first graphically illustrates the sets of prices in (23) and (24) in the context of a simple example with two voucher schools and a specific combination of status-quo and counterfactual voucher amounts. Figure 1(b) then shows how the union of the sets in Figure 1(a) can be partitioned to obtain a collection of sets satisfying Definition 4.1. In particular, it sequentially divides any two sets in Figure 1(a) that partially overlap in a given coordinate until the condition in (26) is satisfied. In Appendix S.2.2, we describe a computational procedure that sequentially divides sets in such a manner to obtain a partition satisfying Definition 4.1 in the case of more than two goods.

Using the above defined collection of sets, we can now define the partition of the space of prices and our choice of $\mathbf{Q}_B^{\text{fd}}$. To define the partition, observe that for each $j \in \mathcal{J}_v$, the collection of sets determined by the prices in $u \in \mathcal{U}$ for the jth school, i.e. $\{u(j) : u \in \mathcal{U}\}$, generates a partition of $[p_j(\max\{\tau_{\text{sq}}, \tau_{\text{c}}\}), p_j(0)] \subseteq [0, p_j(0)]$. Given this implies that $\mathcal{U}_j = \{[0, p_j(\max\{\tau_{\text{sq}}, \tau_{\text{c}}\}))\} \bigcup \{u(j) : u \in \mathcal{U}\}$ corresponds to a partition of $[0, p_j(0)]$ for each $j \in \mathcal{J}_v$, observe that

$$\mathcal{W} = \prod_{j=1}^{J} \mathcal{U}_j \equiv \{w_1, \ldots, w_M\} \ ,$$

denotes a partition of the space of prices $\mathcal{P}$ over which q is defined, where, for each element of the partition, the prices for the jth school in $\mathcal{J}_v$ take values in a set that corresponds to an element of $\mathcal{U}_j$. Then, using this partition, we take

$$\mathbf{Q}_B^{\text{fd}} = \left\{ q \in \mathbf{Q}_B : q_j(p) = \sum_{w \in \mathcal{W}} 1_w(p) \cdot \beta_j(w) \text{ for some } \{\beta_j(w)\}_{w \in \mathcal{W}} \text{ for each } j \in \mathcal{J} \right\} \ , \tag{27}$$

where $1_w(p) \equiv 1\{p \in w\}$, i.e. the space we consider corresponds to a subset of $\mathbf{Q}_B$ such that each q is parameterized to be a constant function over the elements of the partition $\mathcal{W}$.

We next show that replacing $\mathbf{Q}_B$ with this choice of $\mathbf{Q}_B^{\text{fd}}$ leads to no loss of information with respect to what we can learn about the parameter of interest, i.e. $\theta(\mathbf{Q}_B) = \theta\left(\mathbf{Q}_B^{\text{fd}}\right)$. In addition, we also show that characterizing $\theta\left(\mathbf{Q}_B^{\text{fd}}\right)$, which is a finite-dimensional problem, can be solved using two finite-dimensional optimization problems. In order to state this result, it is useful to first restate $\theta\left(\mathbf{Q}_B^{\text{fd}}\right)$ in terms of the variable $\beta \equiv \left(\beta_g', \beta_n', \beta_1', \ldots, \beta_J'\right)'$, where $\beta_j = (\beta_j(w_1), \ldots, \beta_j(w_M))$ for each $j \in \mathcal{J}$, that parameterizes a given $q \in \mathbf{Q}_B^{\text{fd}}$. To this end, note that given each parameter θ is continuous in q and that q is continuous in β, it follows that θ can be written in terms of a continuous function of β in the sense that there exists a continuous function θ_B of β such that $\theta(q) = \theta_B(\beta)$. Similarly, note that $\mathbf{Q}_B$ can also be written in terms of β by

$$\mathbf{B} = \left\{ \beta \in \mathbf{R}^{d_\beta} : \left(\sum_{w \in \mathcal{W}} 1_w \cdot \beta_j(w) : j \in \mathcal{J} \right) \in \mathbf{Q}_B \right\} \ , \tag{28}$$

where d_β denotes the dimension of β, i.e. the set of values of β that ensure that the corresponding q is in $\mathbf{Q}_B$. Then, we can write $\theta\left(\mathbf{Q}_B^{\text{fd}}\right)$ in terms of β by

$$\{\theta_0 \in \mathbf{R} : \theta_B(\beta) = \theta_0 \text{ for some } \beta \in \mathbf{B}\} \equiv \Theta_B \ . \tag{29}$$

In the following proposition, we state the result that the identified set under the baseline specification, i.e. Θ in (22) when $\mathbf{Q} = \mathbf{Q}_B$, is equal to Θ_B. In addition, the proposition also shows that we can characterize Θ_B by solving two finite-dimensional optimization problems.

Proposition 4.1. Suppose that $\mathbf{Q} = \mathbf{Q}_B$. Then, the identified set in (22) is equal to that in (29), i.e. $\Theta = \Theta_B$. In addition, if $\mathbf{B}$ is empty then by definition Θ_B is empty; whereas, if $\mathbf{B}$ is non-empty then $\Theta_B = [\underline{\theta}_B, \bar{\theta}_B]$, where

$$\underline{\theta}_B = \min_{\beta \in \mathbf{B}} \ \theta_B(\beta) \ \text{ and } \ \bar{\theta}_B = \max_{\beta \in \mathbf{B}} \ \theta_B(\beta) \ . \tag{30}$$

Proposition 4.1 shows that the identified set under the baseline specification when not empty is given by a closed interval, where the endpoints can be obtained by solving the two optimization problems stated in (30). In the proof of the proposition, we explicitly derive $\mathbf{B}$, the constraint set of these optimization problems, and observe that it is determined by constraints that are all linear in β. In addition, we also explicitly derive θ_B, the objectives of these optimization problems, for each of our parameters of interest and observe that they all correspond to linear functions of β. These two observations then imply that these optimization problems are, in fact, linear programming problems, a useful observation in their practical implementation. Lastly, observe that to characterize the identified set using these linear programs, we specifically require that $\mathbf{B}$ is non-empty or, equivalently, that the model is not misspecified. However, when this is not the case, these linear programs automatically terminate, indicating that the model is misspecified.

While the optimization problems in (30) are linear programs, they can nonetheless be computationally expensive in cases where the dimension of the optimizing variable β is large. Such a case arises especially in settings when J is large as in our empirical analysis, where we have that J is equal to 68. To ensure tractability in such cases, it is useful to consider alternative lower-dimensional linear programs that are easier to compute and can continue to allow us to learn about our parameters. To this end, observe that, given how $\mathcal{U}$ captured all sets relevant in defining our parameters, only a restricted subset of $\mathcal{W}$ given by

$$\mathcal{W}^{\mathrm{r}} = \left\{ w \in \mathcal{W} : w = \prod_{j=1}^{J} u(j) \text{ for some } u \in \mathcal{U} \right\} \equiv \{w_1^{\mathrm{r}}, \ldots, w_{M^{\mathrm{r}}}^{\mathrm{r}}\} \ ,$$

corresponds to the sets of prices that play a role in the definition of our parameters. In turn, observe that only a subvector of β defined over these sets given by $\beta^{\mathrm{r}} = \left(\beta_g^{\mathrm{r}\prime}, \beta_n^{\mathrm{r}\prime}, \beta_1^{\mathrm{r}\prime}, \ldots, \beta_J^{\mathrm{r}\prime}\right)' \equiv \phi(\beta)$, where $\beta_j^{\mathrm{r}} = \left(\beta_j(w_1^{\mathrm{r}}), \ldots, \beta_j(w_{M^{\mathrm{r}}}^{\mathrm{r}})\right)$ for each $j \in \mathcal{J}$, plays a role in determining θ_B in the sense that there equivalently exists a linear function θ_B^{r} such that $\theta_B^{\mathrm{r}}(\beta^{\mathrm{r}}) = \theta_B(\beta)$. Then, the lower-dimensional linear programs we consider are those in terms of the subvector β^{r} given by

$$\underline{\theta}_B^{\mathrm{r}} = \min_{\beta^{\mathrm{r}} \in \mathbf{B}^{\mathrm{r}}} \ \theta_B^{\mathrm{r}}(\beta^{\mathrm{r}}) \ \text{ and } \ \bar{\theta}_B^{\mathrm{r}} = \max_{\beta^{\mathrm{r}} \in \mathbf{B}^{\mathrm{r}}} \ \theta_B^{\mathrm{r}}(\beta^{\mathrm{r}}) \ , \tag{31}$$

where $\mathbf{B}^{\mathrm{r}}$ denotes a set of β^{r} determined by linear constraints. By an appropriate choice of $\mathbf{B}^{\mathrm{r}}$, these alternative linear programs can continue to allow us to learn about our parameters. To see how, observe first that if $\mathbf{B}^{\mathrm{r}} = \phi(\mathbf{B})$, we have by construction that these programs are equivalent to those in (30). In turn, by taking $\mathbf{B}^{\mathrm{r}}$ to be such that $\phi(\mathbf{B}) \subseteq \mathbf{B}^{\mathrm{r}}$, it follows that we have $\underline{\theta}_B^{\mathrm{r}} \leq \underline{\theta}_B$

and $\bar{\theta}^{\mathrm{r}}_B \geq \bar{\theta}_B$, and can therefore continue to learn about our parameters by obtaining a set that contains the identified set, i.e. $\Theta_B \in \left[\underline{\theta}^{\mathrm{r}}_B, \bar{\theta}^{\mathrm{r}}_B\right]$. In Appendix S.2.3, we provide a natural choice of such a $\mathbf{B}^{\mathrm{r}}$ determined by restrictions on β^{r} implied by those in $\mathbf{B}$, which we find in our empirical analysis can be tractably implemented and also result in informative conclusions.

4.3 Identified Set under Auxiliary Parametric Specifications

We now proceed to show how to characterize the identified set under our auxiliary specification. Under this specification, in contrast to the baseline, note that the problem is finite-dimensional in nature due to the fact that $\mathbf{Q}_A$ is a finite-dimensional parameterized space. As a result, in this case, the identified set can be directly characterized by searching over q in $\mathbf{Q}_A$ and then taking their image under the function θ.

In order to state the result that shows how to do this, it useful to first restate the identified set in terms of the variable α that parameterizes a given $q \in \mathbf{Q}_A$ through (3). Given each parameter θ is continuous in q and that q is continuous in α, note that it follows that θ can be written in terms of a continuous function of α in the sense that there exists a continuous function θ_A of α such that $\theta(q) = \theta_A(\alpha)$. Similarly, note that $\mathbf{Q}_A$ can also be written in terms of α by

$$\mathbf{A} = \left\{ \alpha \in \mathbf{R}^{d_\alpha} : \left(\sum_{k=0}^{K_j} \alpha_{jk} \cdot b_{jk} : j \in \mathcal{J} \right) \in \mathbf{Q}_B \right\} . \tag{32}$$

where d_α denotes the dimension of α, i.e. the set of values of α that ensure that the corresponding q is in $\mathbf{Q}_B$. Then, the identified set under the auxiliary specification, i.e. Θ in (22) when $\mathbf{Q} = \mathbf{Q}_A$, can equivalently be given by

$$\theta_A(\mathbf{A}) = \{\theta_0 \in \mathbf{R} : \theta_A(\alpha) = \theta_0 \text{ for some } \alpha \in \mathbf{A}\} \equiv \Theta_A \ , \tag{33}$$

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

Proposition 4.2. 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

$$\underline{\theta}_A = \inf_{\alpha \in \mathbf{A}} \ \theta_A(\alpha) \ \text{ and } \ \bar{\theta}_A = \sup_{\alpha \in \mathbf{A}} \ \theta_A(\alpha) \ . \tag{34}$$

Proposition 4.2 shows how to characterize the identified set under a general class of parametric restrictions. As we mentioned before, this class allows various types of more flexible versions of the parametric specification in (4)-(5) that ensured point identification of the demand functions. We conclude this section by discussing three types of such specifications we later consider in our empirical analysis that can be implemented using Proposition 4.2.

Assumption O. (Own-price) For each $j \in \mathcal{J}_v$,

$$q_j(p) - q_g(p) = \sum_{k=0}^{K} \alpha_{jk} \cdot p_j^k$$

for some $\{\alpha_{jk} : 0 \leq k \leq K\}$, and $q_n(p) - q_g(p) = \alpha_n$ for some α_n.

Assumption AS. (Additively Separable) For each $j \in \mathcal{J}$,

$$q_j(p) = \sum_{m=1}^{J} \sum_{k=0}^{K} \alpha_{jmk} \cdot p_m^k$$

for some $\{\alpha_{jmk} : m \in \mathcal{J}_v,\ 0 \leq k \leq K\}$.

Assumption NS. (Nonseparable) For each $j \in \mathcal{J}$,

$$q_j(p) = \sum_{\substack{m=1 \\ m \neq j}}^{J} \sum_{k=0}^{K} \sum_{l=0}^{K} \alpha_{jmkl} \cdot p_j^k \cdot p_m^l$$

for some $\{\alpha_{jmkl} : m \in \mathcal{J}_v,\ 0 \leq k, l \leq K\}$, and for each $j \in \{g, n\}$,

$$q_j(p) = \sum_{m=1}^{J} \sum_{k=0}^{K} \alpha_{jmk} \cdot p_m^k$$

for some $\{\alpha_{jmk} : m \in \mathcal{J}_v,\ 0 \leq k \leq K\}$.

Assumption O states that the difference in demand for each $j \in \mathcal{J}_v$ and any $j \in \mathcal{J}_g$ is a function only of its own price, where this function is a polynomial of degree K, and the difference in demand for any $j \in \mathcal{J}_n$ and any $j \in \mathcal{J}_g$ is constant. When K equals one, this corresponds to the linear specification in (4)-(5). However, for larger values of K, it allows for more flexible patterns in prices. Nonetheless, while more flexible, it can still be viewed as restrictive as it assumes the demand for a given school with respect to government-funded schools to be invariant to prices of other schools. To this end, Assumption AS and Assumption NS consider more flexible parametric specifications that allow the demand for each school to depend on the prices of all voucher schools. Assumption AS takes the demand for each $j \in \mathcal{J}$ to be an additively separable function in the prices of each $j \in \mathcal{J}_v$, where these functions are polynomials of degree K. Assumption NS further parsimoniously relaxes the requirement of additive separability by allowing for nonseparability in its own price. In particular, it takes the demand for $j \in \mathcal{J}_v$ to be an additively separable function only in the prices of each $m \in \mathcal{J}_v \setminus \{j\}$, where these bivariate functions are bivariate polynomials of degree K.

While the optimization problems in (34) are finite-dimensional, their computational tractability depends on the structure of the objective θ_A and constraint set $\mathbf{A}$. In Appendix S.2.4, we illustrate that, under each of the different parametric specifications considered above, θ_A for each parameter is a linear function of α and that $\mathbf{A}$ is characterized by linear equality and inequality restrictions on α. However, we observe here that some of the linear restrictions in these cases are evaluated at every possible price vector in $\mathcal{P}$, which implies that the resulting optimization problems in (34) can be generally difficult to compute. To this end, similar in spirit to those in (31), we consider the following alternative optimization problems

$$\underline{\theta}^{\mathrm{r}}_A = \min_{\alpha \in \mathbf{A}^{\mathrm{r}}} \theta_A(\alpha) \quad \text{and} \quad \bar{\theta}^{\mathrm{r}}_A = \max_{\alpha \in \mathbf{A}^{\mathrm{r}}} \theta_A(\alpha) \tag{35}$$

in our empirical analysis, where $\mathbf{A}^{\mathrm{r}}$ corresponds to a subset of $\mathbf{A}$ that evaluates some of the restrictions on only a finite set of prices in $\mathcal{P}$—the exact form of $\mathbf{A}^{\mathrm{r}}$ for each specification is provided in Appendix S.2.4. Indeed, since the objective and the finite number of restrictions determining the constraint sets of these problems are linear in α, they are linear programming problems and hence generally computationally tractable. But, since $\mathbf{A} \subseteq \mathbf{A}^{\mathrm{r}}$, these problems only provide a set that contains the identified set, i.e. $\Theta_A \subseteq [\underline{\theta}^{\mathrm{r}}_A, \bar{\theta}^{\mathrm{r}}_A]$, similar to how those in (31) do so for the identified set under the baseline specification. In our empirical analysis, we nonetheless find that these sets result in informative conclusions.

5 Empirical Analysis

In this section, we now use the tools developed in the previous sections to estimate the welfare effects of the DC Opportunity Scholarship Program.

5.1 The DC Opportunity Scholarship Program

We begin by providing some brief background information on the program. The DC Opportunity Scholarship Program (OSP) was a federally-funded school voucher program established by Congress in January 2004, and which started accepting students for the 2004-2005 (henceforth, 2004) school year. The OSP was structured similarly to other voucher programs that existed at the time (Epple et al., 2017). It was open to students residing in Washington, DC, and whose family income was no higher than 185% of the federal poverty line ($18,850 for a family of four in 2004).[1] It could be used only for K-12 education, and at the time of initial receipt was renewable for up to five years. It provided students a voucher worth $7,500 that could be used to offset tuition, fees, and transportation at any private school of their choice participating in the program.

The law that established the program also mandated its evaluation, which culminated with a final report to Congress (Wolf et al., 2010). The report exploited the fact that the OSP randomly allocated vouchers to participating students. In particular, Congress expected the program to be

[1] All dollar amounts throughout have been deflated to 2004 dollars.

oversubscribed, i.e. the number of applicants would exceed the number of available slots in participating private schools. As a result, it required that vouchers be randomly allocated to applicants through a lottery whenever the program was oversubscribed—see Wolf et al. (2010) for details on the lottery. Wolf et al. (2010) exploited this random allocation by comparing various outcomes of voucher recipients to non-recipients to experimentally evaluate the effect of voucher receipt on these outcomes. The main findings from this report, as listed in its executive summary, can be broadly summarized as follows. First, they find no conclusive evidence that the receipt of the voucher had any significant effects on various outcomes corresponding to student achievement. Second, they find that the receipt of the voucher significantly improved students' chances of graduating from high school. Finally, they find that the receipt of the voucher raised parents' ratings of school safety and satisfaction.

In what follows, we use the tools developed in the previous sections to complement these findings by analyzing the welfare effects of providing the status-quo voucher amount as well as alternative counterfactual amounts. Our analysis is based on the premise that while the receipt of the voucher revealed mixed evidence on outcomes in the sense that there are zero as well as some positive effects, parents may nonetheless value the voucher, potentially across dimensions not easily captured by the outcomes. Indeed, as we highlight below, the data from the program reveals that a non-trivial proportion of voucher recipients used the voucher, which, by revealed preference arguments, implies that recipients may value receiving the voucher. Our analysis below estimates these potential welfare benefits using data collected by the OSP.

5.2 Data and Summary Statistics

The OSP collected detailed data for the first two years of the program, 2004 and 2005, and tracked students for at least four years. Across these years, the school settings were different—the composition of applicants and private schools participating in the program changed. Wolf et al. (2010) provide a detailed description on how the data was collected and various summary statistics for the various years. To keep prices and the set of available schools the same for all students, in our analysis we focus on the second year of the program, 2005, which contains around 80% of the entire sample. In addition, we focus on the initial year of the data for students entering the program this year. As we note in Section 6, this avoids complications that arise from the dynamics of the setup. In Appendices S.4.1-S.4.2, we provide details on how our analysis sample was constructed from the original evaluation data and some statistics on the school setting. Below, we present summary statistics for the main variables our analysis exploits, namely the enrollment shares and the prices as measured by the tuition of private schools participating in the program.

Table 1 presents the empirical enrollment shares across the three types of schools, i.e. government-funded schools (which includes charter schools) and private schools participating and not participat-

Table 1: Enrollment shares across school type by voucher receipt

	With voucher	Without voucher	Difference
Government-funded	0.288	0.901	-0.613
	[0.453]	[0.299]	(0.019)
Non-participating private	0.014	0.020	-0.006
	[0.117]	[0.140]	(0.007)
Participating private	0.698	0.079	0.619
	[0.459]	[0.270]	(0.018)
Observations	1,090	730	

Observations rounded to the nearest 10. Standard deviations in square brackets and robust standard errors in parentheses.

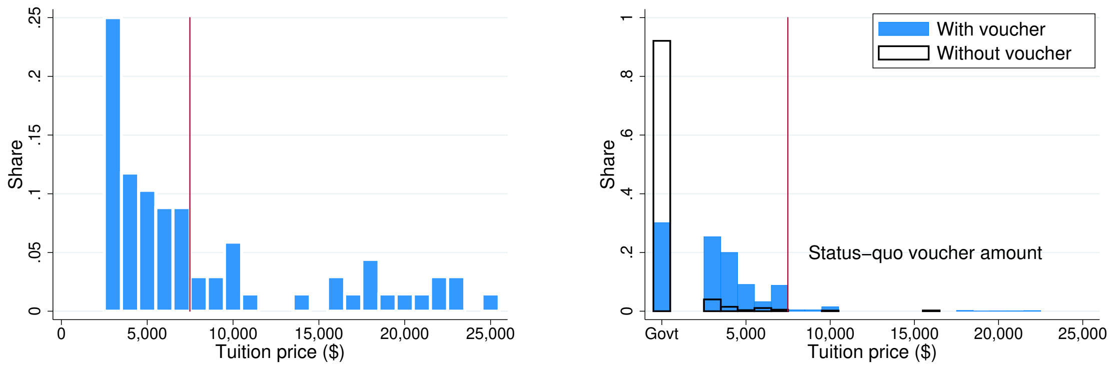
SOURCE: *Evaluation of the DC Opportunity Scholarship Program: Final Report (NCEE 2010-4018)*, U.S. Department of Education, National Center for Education Statistics previously unpublished tabulations.

ing in the program, by voucher receipt. The proportion of voucher recipients who use the voucher, corresponding to those enrolled in participating private schools, is relatively large (69.8%). By revealed preference, this implies that recipients value the voucher. In addition, the voucher increases the proportion enrolling in private schools by 61.9 percentage points, suggesting that prices play an important role in inducing private school enrollment. Finally, observe the nearly symmetric decline in the proportion enrolled in government-funded schools caused by voucher receipt (-61.3 percentage points), which reveals that nearly all students induced into participating private schools by the voucher would be in government-funded schools absent the voucher.

In 2005, there were 68 private schools participating in the program (out of a total of 109 in DC). Figure 2 presents histograms that summarize the variation in prices across these schools as well as the enrollment shares across these prices. Figure 2(a) reveals that a large number of participating private schools had low prices—around 80% had prices below the status-quo voucher amount. Figure 2(b) reveals that the voucher induced a significant proportion to enroll in these low-price schools—out of the 61.9 percentage point increase in the number of students attending a participating private school, a full 59 percentage points (95%) was into schools with prices less than the status-quo voucher amount. As we highlight below, these observations play an important role in better understanding the welfare effects of the voucher.

To also provide some evidence on why recipients may be choosing participating private schools and, in turn, value the voucher, Table S.2 in Appendix S.4.2 compares characteristics of these schools with those of with government-funded schools, where the majority of students enroll absent the voucher. Private and government-funded schools differ across several attributes. The private schools tend to be more religious and specifically Catholic, have smaller school sizes, are more likely to track students by ability, and are less likely to have learning difficulties programs. This suggests that recipients may value these attributes and, hence, the voucher that makes these schools more

Figure 2: Prices and enrollment shares by voucher receipt across participating private schools



(a) Prices of participating private schools

(b) Enrollment shares over prices

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.

affordable. However, our analysis directly estimates the welfare effects of the voucher and does not quantify the effect of these attributes on school valuations. The latter analysis usually requires more demanding variation in the data and modeling assumptions beyond what our analysis exploits (e.g., Carneiro et al., 2019).

Recall from Section 3 that our analysis also uses a value of c_g for the costs the government faces when a student enrolls in government-funded schools as well as a value of γ for administrative costs. In our main analysis, we take c_g = \$5,355, which corresponds to the educational expenditure reported by the US Census (2005). This is lower than total per-pupil expenditure (\$12,979, which includes some fixed costs), or educational expenditure as measured in other sources (\$8,105, Sable and Hill (2006)). However, given that our surplus parameters are increasing in c_g, we choose the smaller, more conservative value. On the other hand, we take γ = \$200, which corresponds to cost of administration, adjudication and providing information to families for an alternative school voucher program reported in Levin and Driver (1997).[2] For our baseline government-funded schools cost of \$5,355, Figure 2(b) reveals that a large proportion of recipients (81%) redeem the voucher at schools with prices below this value. Given that Table 1 revealed that the majority of these recipients would have enrolled in government-funded schools absent the voucher, this suggests that the government may face only small net costs or even savings from the provision of a voucher, even accounting for the administrative costs. Our estimates below make this point more precisely.

[2] As a sensitivity analysis, we also present results for a range of other values of c_g and γ in Appendix S.4.3.

Table 2: Estimated welfare effects of the status-quo voucher

	Nonparametric	Own-price K			Additively separable K			Nonseparable K		
		1	2	3	1	2	3	1	2	3
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	239				1,669	1,000	720	1,669	1,000	720
AB(τ_{sq})	364	∅	∅	∅	1,769	1,100	795	1,769	1,100	795
	5,239				1,873	2,518	2,912	1,920	2,608	3,038
	5,414				1,973	2,618	3,037	2,020	2,733	3,163
	-20				0	0	0	0	0	0
AC(τ_{sq})	150	∅	∅	∅	150	150	150	150	150	150
	300				290	290	290	290	290	290
	113				1,519	825	545	1,519	825	545
AS(τ_{sq})	213	∅	∅	∅	1,619	950	645	1,619	950	645
	5,088				1,723	2,368	2,762	1,770	2,458	2,887
	5,313				1,848	2,493	2,912	1,870	2,583	3,012
Spec. p-value	-	0.00	0.00	0.00	-	-	-	-	-	-

For each parameter, each panel reports the lower endpoint of the 95% CI, the estimated lower bound, the estimated upper bound, and the upper endpoint of the 95% CI, respectively. Upper and lower bound not repeated if they coincide. The ∅ denotes the empty set and indicates that the specification was rejected by the data. For rejected specifications, we provide the specification test p-value.

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.

5.3 Welfare Estimates for the Status-quo Voucher Amount

Table 2 presents the estimates of the welfare effects for the status-quo voucher amount. Each row of the table corresponds to a parameter from (7), (9) or (10), taking $\tau = \tau_{\text{sq}} \equiv \$7,500$. Each column corresponds to a specification of demand, which is either the baseline nonparametric specification defined by Assumption B or an auxiliary parametric specification that additionally imposes either Assumption O, Assumption AS or Assumption NS for some value of K. We consider $K = 1, 2, 3$. The estimates under the nonparametric specification are computed using the optimization problems in (31) with the choice of $\mathbf{B}^{\text{r}}$ described in Appendix S.2.3 and those under the parametric specifications are computed using the optimization problems in (35) with the choices of $\mathbf{A}^{\text{r}}$ described in Appendix S.2.4, where in both cases the enrollment shares in the restrictions in (16)-(19) are replaced by their empirical counterparts. We also report 95% confidence intervals, and specification test p-values in the case of misspecification. These are both constructed using a subsampling procedure from Kalouptsidi et al. (2020), which we describe in Appendix S.3.

The empty sets reveal that some of the specifications may be misspecified. Specifically, the specification in (4)-(5) in Column (2) that implies point identification of the demand function as well as more flexible versions in the form of Assumption O in Columns (3) and (4) may be misspecified. To see why this arises, observe that Assumption O requires the difference between $q_g(p)$ and $q_n(p)$ to be constant for all values of p, which then implies that the difference in enrollment shares for any government-funded and non-participating private school with and without the voucher be equal; however, their empirical counterparts in Table 1 reveal these values are in fact different. The p-values in the final row of Table 2 reveal that this difference is also statistically significant. In contrast, the data do not reject the nonparametric specification in Column (1) or the more flexible parametric specifications in Columns (5)-(10). In these cases, as highlighted in Section 4, there exist multiple demand functions consistent with data and, as a result, we can generally only obtain bounds for the parameters. Nonetheless, as we discuss below, these bounds are quite tight and allow us to reach informative conclusions.

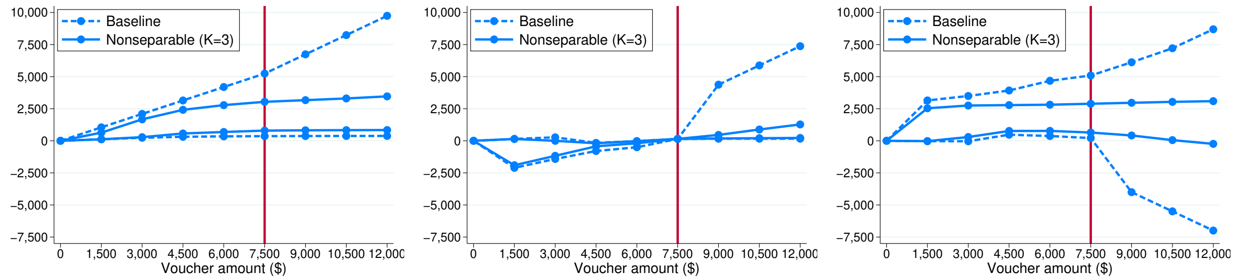
The estimates for $AB\left(\tau_{\mathrm{sq}}\right)$ under the nonparametric specification in Column (1) reveal that the average benefit from the status-quo voucher is between \$364 and \$5,239. Under auxiliary parametric specifications, the bounds can substantially tighten. For example, under the most informative specification in Column (5), the average benefit is between \$1,769 and \$1,873, whereas, under the most flexible specification in Column (10), it is between \$795 and \$3,038.

The estimates for $AC\left(\tau_{\mathrm{sq}}\right)$ reveal that the lower and upper bounds are equal and, in turn, that it is point identified across all specifications. In particular, point identification arises because $AC\left(\tau_{\mathrm{sq}}\right)$ is a function of demand at values of prices at which the demand is exactly observed in the data, namely the prices with and without the status-quo voucher. The point identified value reveals that the average net cost of providing the status-quo voucher is equal to \$150. While this voucher provides an amount of up to \$7,500, the cost is relatively low due to the fact, as highlighted above, that a large proportion of recipients redeem the voucher at low-cost private schools relative to the government-funded schools they would have enrolled in absent the voucher.

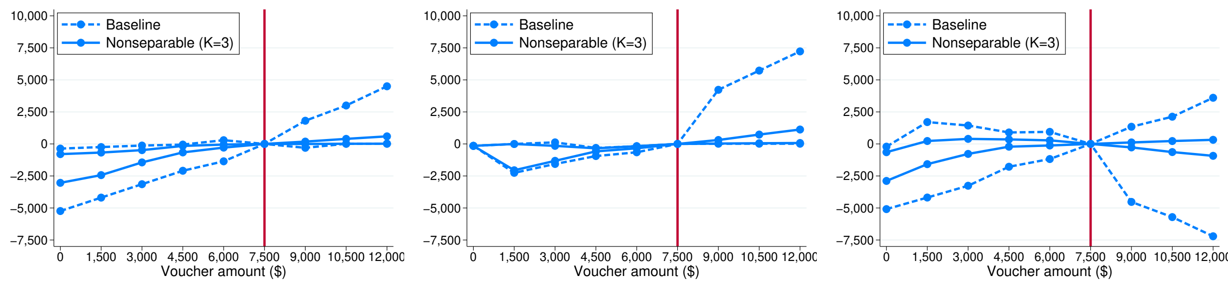
Taking the difference of the average benefit and cost, the estimates for $AS\left(\tau_{\mathrm{sq}}\right)$ reveal that the average benefit net of costs of the status-quo voucher across all specifications is generally positive. In particular, under the nonparametric specification in Column (1), the average surplus is between \$213 and \$5,088 and, under the most flexible parametric specification in Column (10), between \$645 and \$2,887. Intuitively, the positive net benefit arises due to the relatively low net costs of providing the voucher that we highlighted above. Specifically, the voucher recipients have a high welfare benefit from the low-price private schools at which they redeem the voucher relative to the low net costs the government faces to fund the voucher at these schools, which then implies a positive net benefit.

In Appendix S.4.3, we perform several robustness checks on the above conclusion that the provision of the status-quo voucher amount has a positive average surplus. As we noted above, our analysis uses a specific value of c_g for the costs the government faces when a student enrolls in a

Figure 3: Estimated upper and lower bounds on welfare effects for counterfactual voucher amounts



(a) $AB\,(\tau_{\rm c})$ (b) $AC\,(\tau_{\rm c})$ (c) $AS\,(\tau_{\rm c})$



(d) $\Delta AB\,(\tau_{\rm c})$ (e) $\Delta AC\,(\tau_{\rm c})$ (f) $\Delta AS\,(\tau_{\rm c})$

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.

government-funded school, and a value of γ for the administrative costs of providing a voucher. In addition, while the OSP allowed the voucher to be used to offset additional fees and transportation costs, our analysis implicitly assumed that they could be only used to offset tuition. Our robustness analysis measures the sensitivity of our average surplus estimates to taking different values of c_g and γ, and supposing that the voucher could be used to offset an amount δ in addition to the tuition. We find that our conclusions continue to hold for a range of values of c_g, γ and δ.

5.4 Welfare Estimates for Counterfactual Voucher Amounts

Figure 3 next presents the estimates of our various parameters measuring the welfare effects of providing counterfactual voucher amounts. These parameters correspond to those illustrated in Table 2 but for a range of values of $\tau = \tau_{\rm c}$ not necessarily equal to $\tau_{\rm sq}$. We also report the differences with the parameter when $\tau = \tau_{\rm sq}$, as described in (11)-(13). For conservativeness, we present only results under the nonparametric and the most flexible parametric specifications from Table 2, i.e. Columns (1) and (10), respectively. For expositional reasons, since the plots of the estimated bounds and the 95% confidence intervals are close to each other and hence difficult to visually distinguish, we report only the former. The confidence intervals are available in Appendix S.4.4.

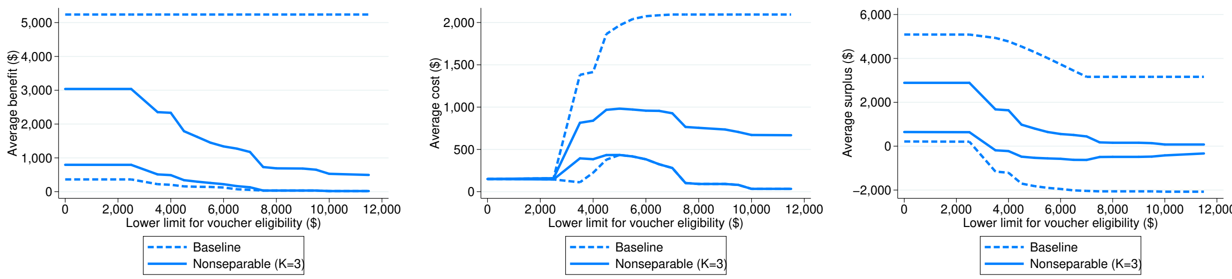
The estimates for $AB\left(\tau_{\mathrm{c}}\right)$ and $\Delta AB\left(\tau_{\mathrm{c}}\right)$ reveal, unsurprisingly, that the average benefit increases with the voucher amount. As in Table 2, the bounds under the parametric specification can be considerably tighter than those under the nonparametric specification. Under the parametric specification, we find that the bounds vary more for lower voucher amounts and are more stable for larger amounts. The estimates for $AC\left(\tau_{\mathrm{c}}\right)$ and $\Delta AC\left(\tau_{\mathrm{c}}\right)$ reveal, in contrast to the status-quo amount in Table 2, that they are generally not point identified but only bounded. This is because, unlike $AC\left(\tau_{\mathrm{sq}}\right)$, these parameters are generally functions of demand at values of prices not observed in the data. Unsurprisingly, the bounds under the nonparametric specification vary non-smoothly and those under the parametric specification vary smoothly given that the latter specification imposes a smooth relationship of how demand varies with prices while the former does not. Similar to the average benefit, the average cost also varies more at lower voucher amounts. For some voucher amounts, the estimates are negative, i.e the government has cost savings. This arises because at these values, as before, recipients continue to redeem the voucher and switch to low-price schools from government-funded schools, but now the government actually saves as the costs of funding the voucher at these schools are significantly lower than that of government-funded schools.

Taking the difference of average benefit and cost, the estimates for $AS\left(\tau_{\mathrm{c}}\right)$ reveal that the provision of counterfactual voucher amounts may have a positive average benefit net of costs. Specifically, under the nonparametric specification, the bounds reveal that we have positive average surplus for voucher amounts below the status-quo, but potentially not above it. This is because the average costs are low relative to the benefit and potentially even negative at voucher amounts below the status-quo, but drastically increase in a non-smooth manner above the status-quo. Under the parametric specification, the smooth relationship of demand with prices allows the pattern of costs below the status-quo voucher to smoothly extend to voucher amounts above it as well, implying a positive average surplus for all voucher amounts. However, not all counterfactual voucher amounts have the same surplus as the status-quo. Comparing the counterfactual to the status-quo surplus at $\tau_c = \$1{,}500$, the bound for $\Delta AS\,(\$1{,}500)$ under the parametric specification is $[-\$1{,}578,\ \$229]$. This suggests that providing vouchers in such low amounts would likely reduce surplus relative to the status-quo amount.

5.5 Role of Low-tuition Schools in the Program

In summary, our welfare estimates reveal that voucher provision has a positive average surplus under both the status-quo and counterfactual voucher amounts. While discussing these results, we specifically noted that they arose in part due to the presence of low-tuition schools in the program that many recipients attend, but that have a small net cost to the government. We conclude our analysis by more directly investigating the importance of these schools in the program when providing the status-quo voucher amount.

Figure 4: Estimated upper and lower bounds on welfare effects of removing schools with tuition at most κ from the program



(a) $AB^{\kappa}\left(\tau_{\text{sq}}\right)$ (b) $AC^{\kappa}\left(\tau_{\text{sq}}\right)$ (c) $AS^{\kappa}\left(\tau_{\text{sq}}\right)$

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.

Specifically, we analyze how our estimates change when we remove schools having prices at most a certain amount from the program. To this end, for a given $\kappa \in \mathbf{R}_{+}$, let $\mathcal{J}^{\kappa} = \{j \in \mathcal{J}_v : p_j(0) \leq \kappa\}$ denote the set of participating private schools with prices no more than κ, and let $j^{\kappa} = \arg\max \mathcal{J}^{\kappa}$ denote the school with the highest price removed from the program. In addition, let $p^{\kappa}(\tau_{\text{sq}}) = (p_1(0), \ldots, p_{j^{\kappa}}(0), p_{j^{\kappa}+1}(\tau_{\text{sq}}), \ldots p_J(\tau_{\text{sq}}))$ denote the prices of the schools in $\mathcal{J}_v$ under the application of the status-quo amount when schools with prices at most κ are removed, i.e. the status-quo voucher amount is applied to only schools with prices above κ. Then, similar to (7), the average benefit of the status-quo voucher amount absent these schools can be defined by

$$AB^{\kappa}\left(\tau_{\text{sq}}\right) = E[B_i^{\kappa}(\tau_{\text{sq}})] \ . \tag{36}$$

where $B_i^{\kappa}(\tau_{\text{sq}})$ is given by the variable that solves (6) when replacing $p_j(\tau)$ with $p_j^{\kappa}(\tau_{\text{sq}})$ for $j \in \mathcal{J}_v$. Similarly, the average cost and benefit net of costs can be defined by

$$AC^{\kappa}\left(\tau_{\text{sq}}\right) = \sum_{j \in \mathcal{J}} c_j^{\kappa}(\tau_{\text{sq}}) \cdot q_j(p^{\kappa}(\tau_{\text{sq}})) - \sum_{j \in \mathcal{J}} c_j(0) \cdot q_j(p(0)) \ , \tag{37}$$

$$AS^{\kappa}\left(\tau_{\text{sq}}\right) = AB^{\kappa}\left(\tau_{\text{sq}}\right) - AC^{\kappa}\left(\tau_{\text{sq}}\right) \ . \tag{38}$$

where $c_j^{\kappa}(\tau_{\text{sq}}) = c_j(\tau_{\text{sq}})$ for $j \in \mathcal{J} \setminus \mathcal{J}^{\kappa}$ and $c_j^{\kappa}(\tau_{\text{sq}}) = 0$ for $j \in \mathcal{J}^{\kappa}$, i.e. we take the same costs as before except with the difference that we take the schools that are removed from the program to have zero costs. In Appendix S.2.5, we describe how we can continue to use the programs in (31) and (34) to learn about these parameters and, in turn, obtain estimates for these parameters using their empirical counterparts as in Table 2 and Figure 3.

Figure 4 presents the results for the above parameters for a range of values of κ and, as in Figure 3, for the nonparametric and most flexible parametric specifications from Table 2. Similar to Figure 3, we only report estimates here and present confidence intervals in Appendix S.4.4.

The bounds under the nonparametric specification are considerably wider than those under the parametric specification and especially so for the average benefit, where the upper bound stays constant across all values of κ. This is because the data do not provide any cross-price variation and, unlike the parametric specification, the nonparametric specification does not impose any cross-price restrictions. Under the parametric specification, Figure 4(a)-(b) reveal that the average benefits and costs first steeply decrease and increase, respectively, from the removal of low-tuition schools from the program and then become more stable when more expensive schools are removed. This highlights that recipients strongly value the presence of low-tuition schools in the program, and when these schools are removed, switch into relatively expensive government-funded schools.

Taking the difference of the average benefit and costs, Figure 4(c) reveals that the removal of low-tuition schools from the program generally results in the reduction of average surplus. Specifically, we find that absent schools with tuition at most \$3,500 in the program we can potentially have a negative surplus. A closer look at Figure 2(a) reveals that nearly 30% of schools in the program have tuition of at most this value. The estimates from Figure 4(c) highlight that the presence of these low-tuition schools in the program play an essential role in explaining the positive net benefit our analysis finds for the provision of the status-quo voucher amount.

6 Conclusion

In this paper, we analyze the welfare effects of voucher provision in the DC Opportunity Scholarship Program (OSP). We do so by developing new tools that show how to generally use data with random allocation of school vouchers to characterize what we can learn about the welfare effects of providing a voucher of a given amount. Applying our tools to the OSP data, our estimates reveal that provision of the status-quo as well as a wide range of counterfactual amounts has a positive net average benefit, and that these positive results arise due to the presence of many low-tuition schools in the program.

We conclude by highlighting some fruitful directions for future research, which are beyond scope of this paper. As noted in Section 2, we do not model the potential general equilibrium effects of vouchers. Our analysis should therefore be interpreted as reflecting short-term, partial equilibrium welfare effects, taking long-term general equilibrium responses as fixed. These responses, however, could have potentially ambiguous effects on the welfare of both voucher recipients and non-recipients (Epple et al., 2017). It would hence be interesting to account for these responses and analyze their consequences on the welfare effects of the voucher. For some recent advancements in this direction, see Bhattacharya et al. (2019), who provides welfare results accounting for certain types of equilibrium responses in a different policy setting.

On a related point, note that our welfare analysis is based on a static choice model. In particular, while the OSP provided vouchers valid for at least five years, we model and analyze school choices collected only in the initial year. In this sense, unless individuals do not change their choices across years, our results should be more appropriately interpreted as the welfare effects of a voucher that is to be used in the same year. As noted in Wolf et al. (2010), there is in fact substantial variation in choice across years. It would hence also be interesting to extend our analysis to encompass related welfare parameters in a dynamic discrete choice model, and study if the positive results we find continue to hold.

Finally, and more generally, it would be interesting to generalize our approach to other types of price variation observed in practice. Specifically, we exploit the fact that in our application there was exogenous, discrete variation in prices due to the random allocation of the voucher. However, in many applications, prices could be endogenously related to the underlying variables of the model. A promising avenue for further research would be to extend our tools to accommodate for such endogeneity, for example through instrumental variable-type assumptions as in Berry and Haile (2009, 2014).

References

Agarwal, N. and Somaini, P. (2020). Revealed preference analysis of school choice models. *Annual Review of Economics*, **12** 471–501.

Allende, C. (2019). Competition under social interactions and the design of education policies. *Job Market Paper*.

Angrist, J., Bettinger, E., Bloom, E., King, E. and Kremer, M. (2002). Vouchers for private schooling in colombia: Evidence from a randomized natural experiment. *American economic review*, **92** 1535–1558.

Arcidiacono, P., Muralidharan, K., Shim, E.-y. and Singleton, J. D. (2016). Valuing school choice: Using a randomized experiment to validate welfare evaluation of private school vouchers.

Berry, S., Gandhi, A. and Haile, P. (2013). Connected substitutes and invertibility of demand. *Econometrica*, **81** 2087–2111.

Berry, S., Levinsohn, J. and Pakes, A. (1995). Automobile prices in market equilibrium. *Econometrica: Journal of the Econometric Society* 841–890.

Berry, S. T. and Haile, P. A. (2009). Nonparametric identification of multinomial choice demand models with heterogeneous consumers. Tech. rep., National Bureau of Economic Research.

Berry, S. T. and Haile, P. A. (2014). Identification in differentiated products markets using market level data. *Econometrica*, **82** 1749–1797.

Bhattacharya, D. (2015). Nonparametric welfare analysis for discrete choice. *Econometrica*, **83** 617–649.

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

Bhattacharya, D. (2019). The empirical content of binary choice models. *arXiv preprint arXiv:1902.11012*.

Bhattacharya, D., Dupas, P. and Kanaya, S. (2019). Demand and welfare analysis in discrete choice models with social interactions. Tech. rep., National Bureau of Economic Research.

Briesch, R. A., Chintagunta, P. K. and Matzkin, R. L. (2010). Nonparametric discrete choice models with unobserved heterogeneity. *Journal of Business & Economic Statistics*, **28** 291–307.

Carneiro, P. M., Das, J. and Reis, H. (2019). The value of private schools: Evidence from pakistan.

Chesher, A., Rosen, A. M. and Smolinski, K. (2013). An instrumental variable model of multiple discrete choice. *Quantitative Economics*, **4** 157–196.

Chiappori, P. and Komunjer, I. (2009). On the nonparametric identification of multiple choice models.

Dynarski, M., Rui, N., Webber, A. and Gutmann, B. (2018). Evaluation of the dc opportunity scholarship program: Impacts two years after students applied. ncee 2018-4010. *National Center for Education Evaluation and Regional Assistance.*

Epple, D., Romano, R. E. and Urquiola, M. (2017). School vouchers: A survey of the economics literature. *Journal of Economic Literature*, **55** 441–92.

Friedman, M. (1962). *Capitalism and freedom.* University of Chicago press.

Gazmuri, A. (2019). School segregation in the presence of student sorting and cream-skimming: Evidence from a school voucher reform.

Howell, W. G., Wolf, P. J., Peterson, P. E. and Campbell, D. E. (2000). Test-score effects of school vouchers in dayton, ohio, new york city, and washington, dc: Evidence from randomized field trials.

Kalouptsidi, M., Kitamura, Y., Lima, L. and Souza-Rodrigues, E. A. (2020). Partial identification and inference for dynamic models and counterfactuals. Tech. rep., National Bureau of Economic Research.

Kamat, V. (2019). Identification with latent choice sets. *arXiv preprint arXiv:1711.02048.*

Kitamura, Y. and Stoye, J. (2018). Nonparametric analysis of random utility models. *Econometrica*, **86** 1883–1909.

Krueger, A. B. and Zhu, P. (2004). Another look at the new york city school voucher experiment. *American Behavioral Scientist*, **47** 658–698.

Levin, H. M. and Driver, C. E. (1997). Cost of an educational voucher system. *Education Economics*, **5** 265–283.

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

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

Mayer, D. P., Peterson, P. E., Myers, D. E., Tuttle, C. C. and Howell, W. G. (2002). School choice in new york city after three years: An evaluation of the school choice scholarships program. final report.

Mills, J. N. and Wolf, P. J. (2017). Vouchers in the bayou: The effects of the louisiana scholarship program on student achievement after 2 years. *Educational Evaluation and Policy Analysis*, **39** 464–484.

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

Muralidharan, K. and Sundararaman, V. (2015). The aggregate effect of school choice: Evidence from a two-stage experiment in india. *The Quarterly Journal of Economics*, **130** 1011–1066.

Neilson, C. (2013). Targeted vouchers, competition among schools, and the academic achievement of poor students.

Sable, J. and Hill, J. (2006). Overview of public elementary and secondary students, staff, schools, school districts, revenues, and expenditures: School year 2004-05 and fiscal year 2004. ed tab. nces 2007-309. *National Center for Education Statistics*.

Tebaldi, P., Torgovitsky, A. and Yang, H. (2019). Nonparametric estimates of demand in the california health insurance exchange. Tech. rep., National Bureau of Economic Research.

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

US Census (2005). Public education finances.

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**.