

# Estimating the Welfare Effects of School Vouchers \*

Vishal Kamat

Toulouse School of Economics  
University of Toulouse Capitole

[vishal.kamat@tse-fr.eu](mailto:vishal.kamat@tse-fr.eu)

Samuel Norris

Harris School of Public Policy  
University of Chicago

[samnorris@uchicago.edu](mailto:samnorris@uchicago.edu)

January 31, 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 vouchers to students. To do so, we develop new discrete choice tools to show how to use data with random allocation of school vouchers to characterize what we can learn about the welfare benefits of providing a voucher of a given amount, as measured by the average willingness to pay for that voucher, and these benefits net of the costs of providing that voucher. A novel feature of our tools is that they allow specifying the relationship of the demand for the various schools with respect to prices to be entirely nonparametric or to be parameterized in a flexible manner, both of which do not necessarily imply that the welfare parameters are point identified. Applying our tools to the OSP data, we find that provision of the status-quo as well as a wide range of counterfactual voucher amounts has a positive net average benefit. We find 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, program evaluation, discrete choice analysis, Opportunity Scholarship Program.

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

---

\*We thank Alex Torgovitsky for useful comments and discussions. Vishal Kamat gratefully acknowledges funding from ANR under grant ANR-17-EURE-0010 (Investissements d'Avenir program).

# 1 Introduction

School vouchers are a topic of active education policy debate across several countries. In its basic form, they are government-funded certificates of a certain amount that parents of 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 they foster school choice that can make voucher recipients better off (Friedman, 1962).

In recent years, a number of studies have empirically investigated this claim by estimating the effects of vouchers on various outcomes using data from programs that randomly allocate vouchers—see, for example, Abdulkadiroğlu et al. (2018), 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) and 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 benefits on outcomes, the data in each of these studies reveals that a non-trivial proportion of recipients choose to use the voucher. By revealed preference arguments, this suggests that recipients may in general value vouchers and be better off across dimensions not easily captured by outcomes.

In this paper, we analyze these potential welfare benefits that recipients may experience under the observed status-quo voucher amount as well as under alternative counterfactual amounts in the DC Opportunity Scholarship Program (OSP). The OSP was the first federally-funded school voucher program in the United States, which was implemented in Washington, DC. It provided a voucher of \$7,500 and, due to oversubscription, randomly allocated the voucher to families. The data from the program reveals that around 70% of the recipients choose to use the voucher, implying that a large proportion of recipients may in fact potentially value the voucher.

Our analysis starts by showing how to generally use data with random allocation of vouchers to characterize what we can 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 had they 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 potential costs the government may face through the provision of the voucher.

To characterize what can be learned about these quantities using the data, we develop new tools. In our model of school choice which is entirely nonparametric, each parameter capturing these various quantities 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

the 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 values revealed by the data. Indeed, this is generally the case unless one solely focuses attention to 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 \$1,030 and \$2,931, whereas, under the nonparametric specification, it is bounded between \$333 and \$5,606. For a wide range of counterfactual voucher amounts, our estimates continue to similarly reveal that provision of the voucher can have a positive net average benefit.

A closer inspection of the data reveals our positive findings on voucher provision arise due to the presence of many 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 low-tuition schools in the OSP. We estimate how the welfare effects of the status-quo voucher amount change when removing such schools from the program. Our estimates reveal the presence of such schools plays an essential role in explaining our positive findings, absent schools with tuition at most \$3,500 in the program can result in a negative net benefit.

We primarily contribute to the literature on the evaluation of school voucher programs. As highlighted above, most papers in this literature estimate the effects of various programs on out-

comes. However, these estimates leave open the question on 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 literature uses differentiated-products demand models to study questions of school choice related to vouchers. These papers analyze, among other things, the changes in school quality induced by a program (Neilson, 2013) and welfare effects of proposed policy changes (Carneiro et al., 2019). Their analyses require exogenous variation in the data 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 the more demanding data and modeling requirements.

In developing our tools, we exploit recent advancements 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 an alternative setting of a treatment effect model.

More broadly, we contribute to the growing literature on non- and semiparametric discrete choice analysis where the underlying model primitive is not necessarily point identified—see, for example, Chesher et al. (2013), Kamat (2019), Kitamura and Stoye (2018), Manski (2007) and Tebaldi et al. (2019). These papers provide various tools, mostly computational, to evaluate different questions such as estimating the effect of different prices and choice sets on demand, characterizing the underlying utility functions, and testing the utility maximization premise of various models. We complement this literature by providing tools to evaluate a question of interest not analyzed in these papers, namely the average willingness to pay for a given decrease in prices.

We organize our analysis in the paper as follows. Section 2 describes our model of school choice and demand specifications we consider. Section 3 defines the parameters we use to measure the welfare benefits as well as the costs of voucher provision. Section 4 characterizes what we can learn about our parameters. Section 5 presents our empirical results on the OSP. Section 6 concludes. Proofs of all results along with additional details pertinent to our analysis are presented in the Supplementary Appendix.

## 2 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  $j$ th school in  $\mathcal{J}_v$ , let  $p_j^* \in \mathbf{R}_+$  denote its original price before applying the voucher and let  $p_j(\tau) \in \mathbf{R}_+$  denote its price after applying a voucher of amount  $\tau \in \mathbf{R}_+$ , where these two prices are related by the relationship

$$p_j(\tau) = \max\{0, p_j^* - \tau\} . \quad (1)$$

Under this notation, the original price and that under the status-quo amount for the  $j$ th 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  $\mathcal{J}_v = \{1, \dots, J\}$ , where the schools in this list are ordered in terms of their original prices, i.e.  $p_1^* \leq \dots \leq p_J^*$ , and we take  $p(\tau) = (p_1(\tau), \dots, p_J(\tau))$  to denote the vector of prices for these schools under a voucher of amount  $\tau$ .

For a given individual, we observe  $Z$  and  $D$ , 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 utility maximization decision. To this end, let  $Y_j$  denote the individual's underlying disposable income under the  $j$ th school in  $\mathcal{J}_g$  or  $\mathcal{J}_n$  and let  $U_j(Y_j)$  denote the corresponding indirect utility under that school. For the 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  $Y_j - p_j$  denote the individual's underlying disposable income under the  $j$ th school in  $\mathcal{J}_v$  had the price of that school been set to  $p_j \in \mathbf{R}_+$ , and let  $U_j(Y_j - p_j)$  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  $p = (p_1, \dots, p_J)$  by

$$D(p) = \begin{cases} \arg \max_{j \in \mathcal{J}_g \cup \mathcal{J}_n} U_j(Y_j) & \text{if } \max_{j \in \mathcal{J}_g \cup \mathcal{J}_n} U_j(Y_j) > \max_{j \in \mathcal{J}_v} U_j(Y_j - p_j) , \\ \arg \max_{j \in \mathcal{J}_v} U_j(Y_j - p_j) & \text{if } \max_{j \in \mathcal{J}_g \cup \mathcal{J}_n} U_j(Y_j) \leq \max_{j \in \mathcal{J}_v} U_j(Y_j - p_j) . \end{cases}$$

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

$$D = D(p(\tau_{\text{sq}})) \cdot Z + D(p(0)) \cdot (1 - Z) . \quad (2)$$

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. 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

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

respectively define the demand for the  $j$ th school in  $\mathcal{J}_v$ , for any school in  $\mathcal{J}_g$  and for any school in  $\mathcal{J}_n$ , conditional on the receipt of the voucher  $Z = z \in \{0, 1\}$ . Analogously, let

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

respectively define the unconditional demand for the  $j$ th school in  $\mathcal{J}_v$ , for any school in  $\mathcal{J}_g$  and for any school in  $\mathcal{J}_n$ . 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  $p_i$  for each  $i \neq j \in \mathcal{J}_v$ .

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  $q \equiv (q_g, q_n, q_1, \dots, q_J)$  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

$$q_j(p) \geq q_j(p') \tag{3}$$

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

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

for each  $j \in \mathcal{J}_v$ , note that it also 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. While the above assumptions impose restrictions directly on the demand functions, note that each of these restrictions follows from restrictions imposed on the underlying variables of the model. For example, Assumption B(i) follows from assuming the voucher to be randomly allocated, i.e.  $Z$  is statistically independent of the remaining variables of the model. On the other hand, Assumption B(ii) follows from assuming  $U_j$  to be weakly or strongly 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  $j \in \mathcal{J}$ ,

$$q_j(p) = \sum_{k=0}^{K_j} \alpha_{jk} \cdot b_{jk}(p) \quad (4)$$

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  $\alpha \equiv (\alpha'_g, \alpha'_n, \alpha'_1, \dots, \alpha'_J)'$ , with  $\alpha_j = (\alpha_{j1}, \dots, \alpha_{jK_j})'$  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

$$q_j(p) = \alpha_{j0} + \alpha_{j1} \cdot p_j \quad \text{for } j \in \mathcal{J}_v, \quad (5)$$

$$q_j(p) = \alpha_{j0} \quad \text{for } j \in \{g, n\}, \quad (6)$$

for some  $\{\alpha_{jk} : j \in \mathcal{J}_v, 0 \leq k \leq 1\}$  and  $\{\alpha_{j0} : j \in \{g, n\}\}$ , i.e. the demand for voucher schools are linear functions of their own prices and the demand for any government or non-voucher school is constant—see Appendix S.2.1 for details on how this specification imposes restrictions similar to those imposed by a logit specification and—like the logit—achieves point identification of the demand functions. However, Assumption A also allows for more flexible specifications with smoother polynomial functions in 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 of  $\tau \in \mathbf{R}_+$ . As mentioned below, this generality, by choosing alternative values of  $\tau$ , 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 a given individual, 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  $\tau$  is defined by the variable  $B(\tau)$  that solves

$$\begin{aligned} & \max \left\{ \max_{j \in \mathcal{J}_g \cup \mathcal{J}_n} U_j(Y_j), \max_{j \in \mathcal{J}_v} U_j(Y_j - p_j(0)) \right\} \\ & \qquad \qquad \qquad = \\ & \max \left\{ \max_{j \in \mathcal{J}_g \cup \mathcal{J}_n} U_j(Y_j - B(\tau)), \max_{j \in \mathcal{J}_v} U_j(Y_j - p_j(t) - B(\tau)) \right\}, \end{aligned} \tag{7}$$

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 that in the absence of the voucher. We then quantify the average benefit of a voucher that provides an amount of  $\tau$  by

$$AB(\tau) = E[B(\tau)], \tag{8}$$

i.e. the average willingness to pay to receive the voucher amount of  $\tau$ .

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 (7) 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  $j$ th 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  $\tau$  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.** Suppose  $U_j$  is continuous and strictly increasing for each  $j \in \mathcal{J}_s$ . Then we have



that  $B(\tau)$  defined in (7) exists and is unique, and that

$$E[B(\tau)] = \sum_{l=0}^{j(\tau)} \int_{a_l(t)}^{a_{l+1}(\tau)} \left( \sum_{j=l+1}^J q_j(p_1(0), \dots, p_l(0), p_{l+1}(\tau) + a, \dots, p_J(\tau) + a) \right) da . \quad (9)$$

While voucher provision can have benefits, it can also be costly to the government who finances the voucher. To benchmark the benefits and perform a cost-benefit analysis, we therefore 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 under 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  $j$ th demand function in  $\mathcal{J}$  under a voucher of amount  $\tau$ . For example, in our baseline empirical analysis, we take

$$c_j(\tau) = \begin{cases} c_g & \text{for } j = \mathcal{J}_g , \\ 0 & \text{for } j = \mathcal{J}_n , \\ \min\{p_j(0), \tau\} & \text{for } j \in \mathcal{J}_v , \end{cases}$$

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. We then measure the average net costs from the provision of a voucher of amount  $\tau$  by

$$AC(\tau) = \sum_{j \in \mathcal{J}} c_j(\tau) \cdot q_j(p(\tau)) - \sum_{j \in \mathcal{J}} c_j(0) \cdot q_j(p(0)) , \quad (10)$$

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 amount of  $\tau$ , let

$$AS(\tau) = AB(\tau) - AC(\tau) \quad (11)$$

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 amount of  $\tau$ . By taking different values of  $\tau$ , 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

$\tau = \tau_c \neq \tau_{sq}$ , we can evaluate these parameters for a counterfactual amount of  $\tau_c$ . In our analysis, we also study the difference of the parameters under these amounts, i.e.

$$\Delta AB(\tau_c) = AB(\tau_c) - AB(\tau_{sq}) , \quad (12)$$

$$\Delta AC(\tau_c) = AC(\tau_c) - AC(\tau_{sq}) , \quad (13)$$

$$\Delta AS(\tau_c) = AS(\tau_c) - AS(\tau_{sq}) , \quad (14)$$

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 data.

### 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  $\theta$  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

$$0 \leq q_j(p) \leq 1 \text{ for each } j \in \mathcal{J} , \quad (15)$$

$$\sum_{j \in \mathcal{J}} q_j(p) = 1 , \quad (16)$$

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 (3). Under our auxiliary specifications, we additionally know that  $q$  satisfies Assumption A, i.e. it satisfies the parametric restrictions stated in (4). Finally, under both specifications, the data also restricts the values that  $q$  can take. Specifically, it follows from (2) and Assumption B(i) that the data reveals

$$q_j(p(0)) = \text{Prob}[D = j | Z = 0] \equiv P_{j|0} , \quad (17)$$

$$q_j(p(\tau_{sq})) = \text{Prob}[D = j | Z = 1] \equiv P_{j|1} \quad (18)$$

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

$$q_j(p(0)) = \text{Prob}[D \in \mathcal{J}_j | Z = 0] \equiv P_{j|0}, \quad (19)$$

$$q_j(p(\tau_{\text{sq}})) = \text{Prob}[D \in \mathcal{J}_j | Z = 1] \equiv P_{j|1} \quad (20)$$

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

$$\mathbf{Q}_B = \{q \in \mathbf{F} : q \text{ satisfies (15) - (16), (3) and (17) - (20)}\} \quad (21)$$

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

$$\mathbf{Q}_A = \{q \in \mathbf{F} : q \text{ satisfies (15) - (16), (3), (4) and (17) - (20)}\} \quad (22)$$

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 value of  $\theta(q)$ . For example, as we noted before, this is the case under the specification described in (5)-(6). 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  $\theta(q)$  may potentially lie 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

$$\theta(\mathbf{Q}) = \{\theta_0 \in \mathbf{R} : \theta(q) = \theta_0 \text{ for some } q \in \mathbf{Q}\} \equiv \Theta, \quad (23)$$

i.e. the image of the set of admissible functions  $\mathbf{Q}$  under the function  $\theta$ . 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.  $\Theta$  in (23) when  $\mathbf{Q} = \mathbf{Q}_B$ ; and then, in Section 4.3, under our auxiliary specification, i.e.  $\Theta$  in (23) 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  $\theta$ . 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  $\theta(\mathbf{Q}_B) = \theta(\mathbf{Q}_B^{\text{fd}})$ . 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 space  $\mathbf{Q}_B^{\text{fd}}$  we consider. In order to do so, we need to first define a collection of sets that plays a key role in the subsequent definition of  $\mathbf{Q}_B^{\text{fd}}$ . To this end, observe that

$$\mathcal{P}_l(\tau) = \{p \in \mathcal{P} : p_j = \min\{p_j(0), p_j(\tau) + a\} \text{ for } a \in [a_l(\tau), a_{l+1}(\tau)] \text{ for each } j \in \mathcal{J}_v\} \quad (24)$$

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

$$\{p(0), p(\tau_{\text{sq}}), p(\tau)\} \quad (25)$$

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 (17)-(20). Note it then follows that

$$\mathcal{P}^* = \bigcup_{l=0}^{j(\tau_{\text{sq}})} \mathcal{P}_l(\tau_{\text{sq}}) \bigcup_{l=0}^{j(\tau_c)} \mathcal{P}_l(\tau_c) \bigcup \{p(0), p(\tau_{\text{sq}}), p(\tau_c)\} \quad (26)$$

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  $\tau_c$  along with the restrictions imposed by the data. Given this set of prices, we define in the following definition the collection of sets  $\mathcal{U}$  that we later use below in the definition of  $\mathbf{Q}_B^{\text{fd}}$ .

**Definition 4.1.** Let  $\mathcal{U} = \{u_1, \dots, u_M\}$  denote a finite partition of the set of prices  $\mathcal{P}^*$  in (26) such that for all  $u \in \mathcal{U}$  we have either

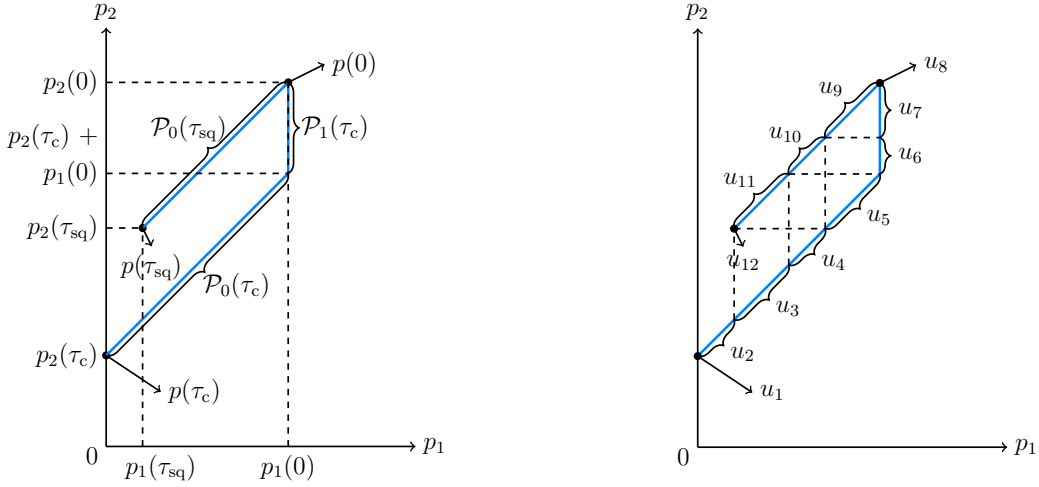
- (i)  $u = \{p \in \mathcal{P} : p_j = \min\{p_j(0), p_j(\tau) + a\} \text{ for } a \in (\underline{a}_u, \bar{a}_u] \text{ or } (\underline{a}_u, \bar{a}_u) \text{ for each } j \in \mathcal{J}_v\}$  where  $\underline{a}_u$  and  $\bar{a}_u$  are such that  $u \subseteq \mathcal{P}_l(t)$  for some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

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

$$u(j) = u'(j) \text{ or } u(j) \cap u'(j) = \emptyset \quad (27)$$

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  $j \in \mathcal{J}_v$ .

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



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

(b) A partition  $\mathcal{U}$  of the union of sets in (a) that satisfies Definition 4.1

Definition 4.1 states that  $\mathcal{U}$  corresponds to a finite partition of  $\mathcal{P}^*$ , where each element of the partition satisfies a specific property. In particular, Definition 4.1(i)-(ii) states that each element is a set that corresponds to a connected subset of that in (24) or (25). In addition, it states in (27) that any pair of sets in this partition are such that they either completely overlap or are disjoint in each price coordinate. Intuitively, this latter property implies that the sets can be ordered across the prices of each voucher school, which will allow the finite-dimensional space we consider to preserve the information provided by the shape restrictions in (3).

To better understand these various set of prices, Figure 1(a) first graphically illustrates the sets of prices in (24) and (25) 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 (27) is satisfied. In Appendix S.2.2, we describe a computational procedure sequentially dividing sets in such a manner that can be used 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  $\mathbf{Q}_B^{\text{fd}}$ . In particular, it is based on a specific parameterization of  $q$  constructed using  $\mathcal{U}$ . To define this parameterization, observe that for each  $j \in \mathcal{J}_v$ , the collection of sets determined by the prices in  $u \in \mathcal{U}$  for the  $j$ th school, i.e.  $\{u(j) : u \in \mathcal{U}\}$ , generates a partition of  $[p_j(\min\{\tau_{\text{sq}}, \tau_c\}), p_j(0)] \subseteq [0, p_j(0)]$ . Given this implies that  $\mathcal{U}_j = \{[0, p_j(\min\{\tau_{\text{sq}}, \tau_c\})]\} \cup \{u(j) : u \in \mathcal{U}\}$  corresponds to a partition of  $[0, p_j(0)]$  for each  $j \in \mathcal{J}_v$ ,

observe that

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

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  $j$ th 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

$$\mathbf{Q}_B^{\text{fd}} = \left\{ q \in \mathbf{Q}_B : q_j(p) = \sum_{w \in \mathcal{W}} 1_w(p) \cdot \beta_j(w) \text{ for some } \{\beta_j(w)\}_{w \in \mathcal{W}} \text{ for each } j \in \mathcal{J} \right\} , \quad (28)$$

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 with  $\mathbf{Q}_B$  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.  $\theta(\mathbf{Q}_B) = \theta(\mathbf{Q}_B^{\text{fd}})$ . In addition, we also show that characterizing  $\theta(\mathbf{Q}_B^{\text{fd}})$ , which is a finite-dimensional problem, can be solved using two finite-dimensional optimization problems. In order to state this result, it useful to first restate  $\theta(\mathbf{Q}_B^{\text{fd}})$  in terms of the variable  $\beta \equiv (\beta'_g, \beta'_n, \beta'_1, \dots, \beta'_J)'$ , where  $\beta_j = (\beta_j(w_1), \dots, \beta_j(w_N))$  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  $\theta$  is continuous in  $q$  and that  $q$  is continuous in  $\beta$ , it follows that  $\theta$  can be written in terms of a continuous function of  $\beta$  in the sense that there exists a continuous function  $\theta_B$  of  $\beta$  such that  $\theta(q) = \theta_B(\beta)$ . Similarly, note that  $\mathbf{Q}_B$  can also be written in terms of  $\beta$  by

$$\mathbf{B} = \left\{ \beta \in \mathbf{R}^{d_\beta} : \left( \sum_{w \in \mathcal{W}} 1_w \cdot \beta_j(w) : j \in \mathcal{J} \right) \in \mathbf{Q}_B \right\} , \quad (29)$$

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

$$\{\theta_0 \in \mathbf{R} : \theta_B(\beta) = \theta_0 \text{ for some } \beta \in \mathbf{B}\} \equiv \Theta_B . \quad (30)$$

In the following proposition, we state the result that the identified set under the baseline specification, i.e.  $\Theta$  in (23) when  $\mathbf{Q} = \mathbf{Q}_B$ , is equal to  $\Theta_B$ . In addition, the proposition also shows that we can characterize  $\Theta_B$  by solving two finite-dimensional optimization problems.

**Proposition 4.1.** Suppose that  $\mathbf{Q} = \mathbf{Q}_B$ . Then, the identified set in (23) is equal to that in (30), i.e.  $\Theta = \Theta_B$ . In addition, if  $\mathbf{B}$  is empty then by definition  $\Theta_B$  is empty; whereas, if  $\mathbf{B}$  is non-empty then  $\Theta_B = [\underline{\theta}_B, \bar{\theta}_B]$ , where

$$\underline{\theta}_B = \min_{\beta \in \mathbf{B}} \theta_B(\beta) \quad \text{and} \quad \bar{\theta}_B = \max_{\beta \in \mathbf{B}} \theta_B(\beta) . \quad (31)$$

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 (31). 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  $\beta$ . In addition, we also explicitly derive  $\theta_B$ , the objectives of these optimization problems, for each of our parameters of interest and observe that they all correspond to linear functions of  $\beta$ . 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, note the same linear programs automatically terminate and, in turn, also automatically indicate when the model is misspecified.

While the optimization problems in (31) are linear programs, they can nonetheless be computationally expensive in cases where the dimension of the optimizing variable  $\beta$  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

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

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

$$\underline{\theta}_B^r = \min_{\beta^r \in \mathbf{B}^r} \theta_B^r(\beta^r) \quad \text{and} \quad \bar{\theta}_B^r = \max_{\beta^r \in \mathbf{B}^r} \theta_B^r(\beta^r) , \quad (32)$$

where  $\mathbf{B}^r$  denotes a set of  $\beta^r$  determined by linear constraints. By an appropriate choice of  $\mathbf{B}^r$ , these alternative linear programs can continue to allow us to learn about our parameters. To see how, observe first that if  $\mathbf{B}^r = \phi(\mathbf{B})$ , we have by construction that these programs are equivalent to those in (31). In turn, by taking  $\mathbf{B}^r$  to be such that  $\phi(\mathbf{B}) \subseteq \mathbf{B}^r$ , it follows that we have  $\underline{\theta}_B^r \leq \underline{\theta}_B$  and  $\bar{\theta}_B^r \geq \bar{\theta}_B$ , and can therefore continue to learn about our parameters by obtaining a set that contains the identified set, i.e.  $\Theta_B \in [\underline{\theta}_B^r, \bar{\theta}_B^r]$ . In Appendix S.2.3, we provide a natural choice of such a  $\mathbf{B}^r$  determined by restrictions on  $\beta^r$  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  $\theta$ .

In order to state the result that shows how to do this, it is useful to first restate the identified set in terms of the variable  $\alpha$  that parameterizes a given  $q \in \mathbf{Q}_A$  through (4). Given each parameter  $\theta$  is continuous in  $q$  and that  $q$  is continuous in  $\alpha$ , note that it follows that  $\theta$  can be written in terms of a continuous function of  $\alpha$  in the sense that there exists a continuous function  $\theta_A$  of  $\alpha$  such that  $\theta(q) = \theta_A(\alpha)$ . Similarly, note that  $\mathbf{Q}_A$  can also be written in terms of  $\alpha$  by

$$\mathbf{A} = \left\{ \alpha \in \mathbf{R}^{d_\alpha} : \left( \sum_{k=0}^{K_j} \alpha_{jk} \cdot b_{jk} : j \in \mathcal{J} \right) \in \mathbf{Q}_B \right\} . \quad (33)$$

where  $d_\alpha$  denotes the dimension of  $\alpha$ , i.e. the set of values of  $\alpha$  that ensure that the corresponding  $q$  is in  $\mathbf{Q}_B$ . Then, the identified set under the auxiliary specification, i.e.  $\Theta$  in (23) when  $\mathbf{Q} = \mathbf{Q}_A$ , can equivalently be given by

$$\theta_A(\mathbf{A}) = \{ \theta_0 \in \mathbf{R} : \theta_A(\alpha) = \theta_0 \text{ for some } \alpha \in \mathbf{A} \} \equiv \Theta_A , \quad (34)$$

i.e. the image of the set  $\mathbf{A}$  under the function  $\theta_A$ . 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  $\Theta_A$  is empty; whereas, if  $\mathbf{A}$  is connected and non-empty, then the closure of  $\Theta_A$  is given by  $[\underline{\theta}_A, \bar{\theta}_A]$ , where

$$\underline{\theta}_A = \inf_{\alpha \in \mathbf{A}} \theta_A(\alpha) \quad \text{and} \quad \bar{\theta}_A = \sup_{\alpha \in \mathbf{A}} \theta_A(\alpha) . \quad (35)$$

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 (5)-(6) 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  $j \in \mathcal{J}_v$ ,

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

for some  $\{ \alpha_{jk} : 0 \leq k \leq K \}$ , and for each  $j \in \{g, n\}$ ,  $q_j(p) = \alpha_{j0}$  for some  $\alpha_{j0}$ .



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

$$q_j(p) = \sum_{i=1}^J \sum_{k=0}^K \alpha_{jik} \cdot p_i^k$$

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

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

$$q_j(p) = \sum_{\substack{i=1 \\ i \neq j}}^J \sum_{k=0}^K \sum_{l=0}^K \alpha_{jikl} \cdot p_j^k \cdot p_i^l$$

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

$$q_j(p) = \sum_{i=1}^J \sum_{k=0}^K \alpha_{jik} \cdot p_i^k$$

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

Assumption **O** states that the demand for each  $j \in \mathcal{J}_v$  is a function of only its own price, where this function is a polynomial of degree  $K$ , and that the demand for each  $j \in \{g, n\}$  is a constant function. When  $K$  is equal one, note that this corresponds to the linear specification in (5)-(6). However, for larger values of  $K$ , it allows for more smooth and flexible patterns in prices. Nonetheless, while more flexible, it can still be viewed as restrictive as it assumes the demand for a given school to be constant across changes in 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  $i \in \mathcal{J}_v \setminus \{j\}$ , where these bivariate functions are bivariate polynomials of degree  $K$ .

As we illustrate in Appendix **S.2.4**, each of these specifications can be implemented through Proposition **4.2** by rewriting them using the Bernstein polynomial basis. In particular, under this basis, we can explicitly write **A** in a straightforward manner in terms of a system of linear equality and inequality restrictions on  $\alpha$ . Moreover, we can also explicitly write  $\theta_A$  for each of our parameters in terms of a linear function of  $\alpha$ . As a result, similar to (31), since the optimization problems in (35) have linear objectives as well a constraint set determined by a linear system of equations, it follows that they correspond to linear programming problems.

## 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 a congressional act in January 2004, and started accepting students for the 2004-2005 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 dollars). 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 amount of \$7,500 that could be used to offset tuition, fees, and transportation to any private school of their choice participating in the program.

The congressional act 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 oversubscribed, i.e. the number of applicants would exceed the number of available slots in participating private schools. As a result, it required vouchers be allocated randomly 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 outcome 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 that there is 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 had significant effects on improving 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

Table 1: Enrollment shares across school type by voucher receipt

	With voucher	Without voucher	Difference
Government-funded	0.296	0.910	-0.614
Non-participating private	0.006	0.011	-0.005
Participating private	0.698	0.079	0.619
Observations	1,090	730	

*Notes:* Observations rounded to the nearest 10.

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

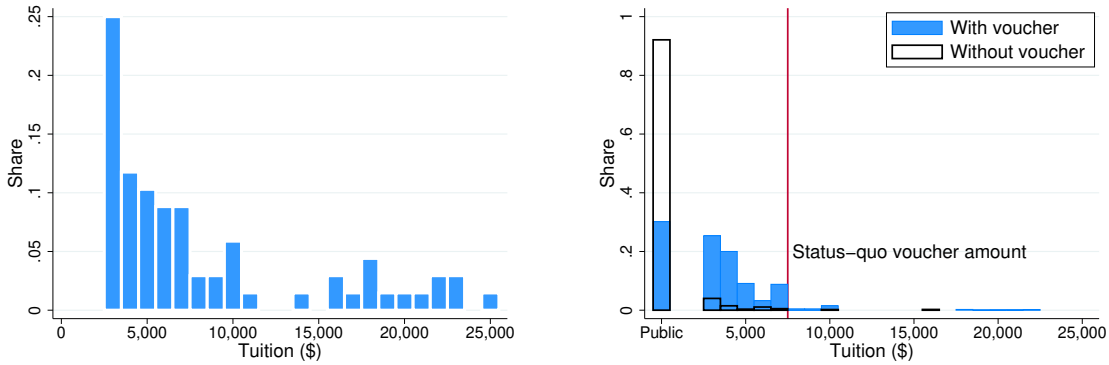
that recipients may value receiving the voucher. Our analysis below estimates these potential welfare benefits using data collected by the OSP.

## 5.2 Data and Summary Statistics

The OSP collected detailed data for the first two years of the program, namely 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 statistics of the school setting in the various years. In our analysis, we focus on the second year of the program, i.e. 2005, as it corresponded to the largest year of the evaluation, 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 discuss in [Section 6](#), this avoids complications that arise from the dynamics of the setup. In [Appendices S.3.1-S.3.2](#), 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 from the data our analysis exploits, namely the enrollments shares and the prices as measured by the tuition of private schools participating in the program.

[Table 1](#) presents the empirical enrollment shares across any of the three types of schools, i.e. government-funded schools (which includes charter schools) and private schools participating and not participating in the program, by voucher receipt. Observe the proportion of voucher recipients who use the voucher, corresponding to those enrolled in participating private schools, is relatively large (69.8%). As noted above, this implies, by revealed preference, that recipients value the voucher. In addition, observe the large difference in the proportion enrolling in participating private schools with and without the voucher (61.9%), which suggests that prices play an important role in inducing private school enrollment. Finally, observe the difference in the proportion enrolled in government-funded schools with and without the voucher (-61.4%), which reveals that a large proportion of those induced into participating private schools by the voucher (91.2% out of 61.9%)

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



(a) Histogram of prices across participating private schools

(b) Histogram of enrolment shares across participating private school 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.

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). 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% increase in the number of students attending a participating private school, a full 59% (97%) of which 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 our analysis estimates.

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.3.2 presents characteristics of these schools along with government-funded schools, those where the majority of recipients would have enrolled absent the voucher. These sets of schools differ across several attributes. The private schools tend to be more religious and specifically Catholic, have lower school sizes, have more students tracked by ability, and lower learning difficulties program. This suggests recipients may value these attributes and, hence, the voucher that makes these schools more 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—see Neilson (2013) and Carneiro et al. (2019) for examples of such an analysis.

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. 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. As a sensitivity analysis, we also present results for a range of other values. For our baseline cost value 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, as our estimates below more precisely capture, that the government can potentially experience small net costs or even savings by the provision of a voucher.

### 5.3 Welfare Estimates for the Status-quo Voucher Amount

We now present the results of our analysis that estimates the welfare effects of providing the status-quo voucher amount as well as alternative counterfactual amounts. [Table 2](#) first presents the estimates for the status-quo voucher amount. Each row of the table corresponds to a parameter from [\(8\)](#), [\(10\)](#) or [\(11\)](#) taking  $\tau = \tau_{sq} \equiv \$7,500$ , whereas 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 [\(32\)](#) with the choice of  $\mathbf{B}^f$  described in [Appendix S.2.3](#) and those under the parametric specifications are computed using the optimization problems in [\(35\)](#), where in both cases the enrollment shares in the restrictions in [\(17\)-\(20\)](#) are replaced by their empirical counterparts.

The empty sets reveal that some of the specifications may be misspecified. Specifically, the specification in [\(5\)-\(6\)](#) 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  $q_g(p)$  and  $q_n(p)$  to be constant for all values of  $p$ , which then implies given [\(19\)](#) and [\(20\)](#) that the 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. More generally, the rejection arises from the fact that while the specification imposes that the demand for a given school is not affected by prices of other schools, the data reveals this requirement is too restrictive. In contrast, for the more flexible parametric specifications as well as the nonparametric specification, the results in [Columns \(1\) and \(5\)-\(10\)](#) imply that they are not misspecified. 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

Table 2: Average benefit, cost and surplus estimates for the status-quo voucher amount

	Baseline	Own-price			Additively separable			Nonseparable		
		$K$			$K$			$K$		
		1	2	3	1	2	3	1	2	3
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
$AB(\tau_{sq})$	599	$\emptyset$	$\emptyset$	$\emptyset$	2,086	1,562	1,311	2,086	1,562	1,311
	6,029	$\emptyset$	$\emptyset$	$\emptyset$	2,252	2,779	3,050	2,324	2,903	3,210
$AC(\tau_{sq})$	207	$\emptyset$	$\emptyset$	$\emptyset$	207	207	207	207	207	207
	207	$\emptyset$	$\emptyset$	$\emptyset$	207	207	207	207	207	207
$AS(\tau_{sq})$	392	$\emptyset$	$\emptyset$	$\emptyset$	1,879	1,355	1,104	1,879	1,355	1,104
	5,822	$\emptyset$	$\emptyset$	$\emptyset$	2,045	2,572	2,843	2,117	2,696	3,004

*Notes:* For each estimate, the upper and lower value for each panel correspond to the lower and upper bound, respectively. All amounts measured in US dollars.

*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.

below, these bounds are quite tight and allow us to reach information conclusions.

The estimates for  $AB(\tau_{sq})$  under the nonparametric specification in Column (1) reveal that the average benefit from the status-quo voucher is between \$599 and \$6,029. Under auxiliary parametric specifications, the bounds can substantially tighten. For example, under the most informative specification in Column (5), the average benefit can be between \$2,086 and \$2,252, whereas, under the most flexible specification in Column (10), it can be between \$1,311 and \$3,210. The estimates for  $AC(\tau_{sq})$  reveal that the lower and upper bound are equal and, in turn, that it is point identified across all specifications. In particular, point identification arises because  $AC(\tau_{sq})$  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 \$207. While this voucher provides an amount of upto \$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  $AS(\tau_{sq})$  reveal that the average benefit net of costs of the status-quo voucher across all specification is generally positive. In particular, under the nonparametric specification in Column (1), the bounds reveal that the average surplus is between \$392 and \$5,822 and, under the most flexible parameteric specification in Column (10), the bounds reveal that it is between \$1,104 and \$3,004. 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.

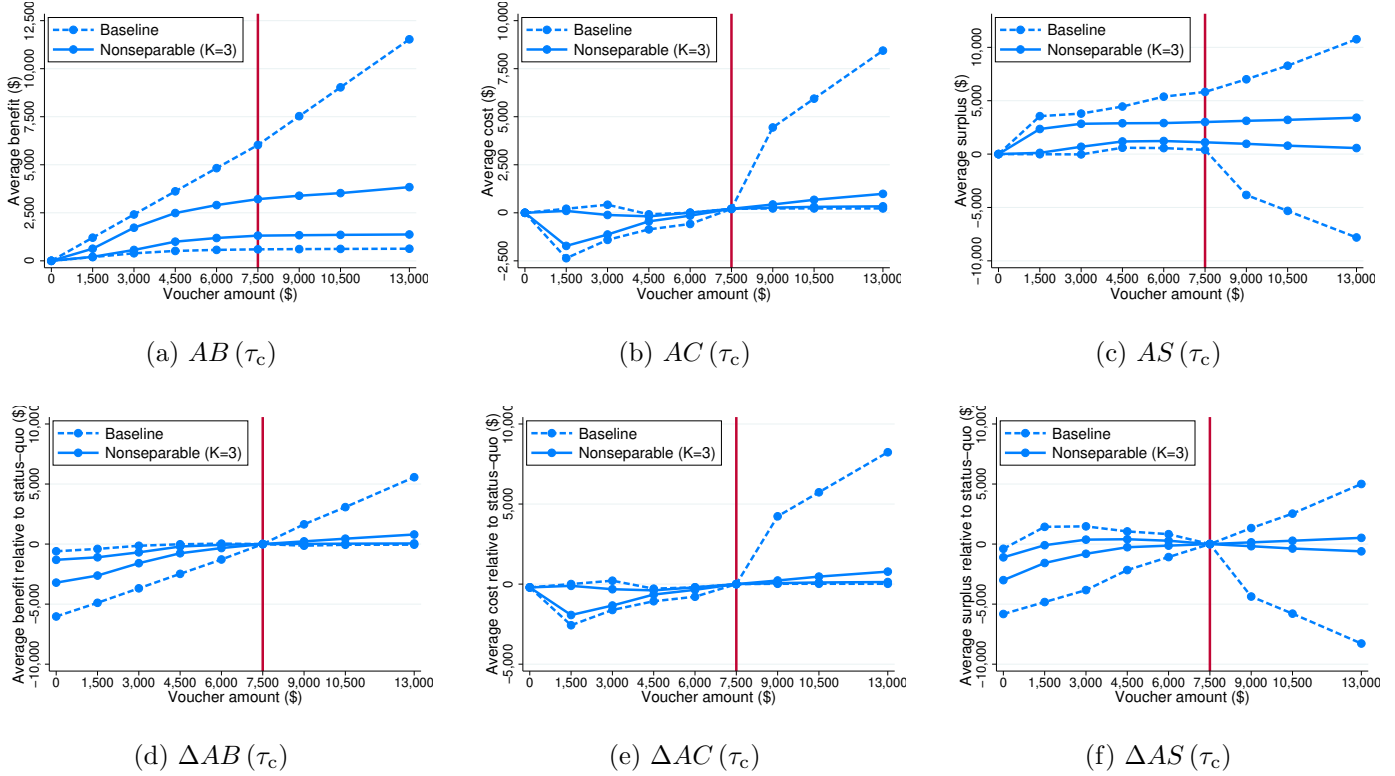
In Appendix S.3.3, we perform several robustness checks on the above conclusion that the provision of the status-quo voucher amount has a positive net average benefit. As we noted above, our analysis chose a specific value of  $c_g$  for the costs the government faces when a student enrolls in a government-funded school. In addition, while the OSP allowed the voucher to be used to offset additional fees and transportation costs, our analysis implicitly presumed 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  $c_g$  as well as supposing that the voucher could be used to offset an amount  $\delta$  in addition to the tuition. We find that our conclusion continues to hold for a range of values of  $c_g$  and  $\delta$ .

#### 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  $\tau = \tau_c$  not necessarily equal to  $\tau_{sq}$  as well as their differences with the parameter when  $\tau = \tau_{sq}$  as described in (12)-(14). For conservativeness, we present only results under the nonparametric and the most flexible parameteric specifications from Table 2, i.e. Columns (1) and (10), respectively. As in Table 2, the estimates are obtained using the corresponding optimization problems with the empirical enrollment shares.

The estimates for  $AB(\tau_c)$  and  $\Delta AB(\tau_c)$  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  $AC(\tau_c)$  and  $\Delta AC(\tau_c)$  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  $AC(\tau_{sq})$ , 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 reveal that we can in fact have also a negative cost, i.e. 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 school, 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.

Figure 3: Estimates for counterfactual voucher amounts



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

Taking the difference of average benefit and cost, the estimates for  $AS(\tau_c)$  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 about it as well implying a positive average surplus for all voucher amounts. Comparing these values to the status-quo surplus, the estimates for  $\Delta AS(\tau_c)$  under the parametric specification reveal that the bounds for voucher amounts below around \$1,500 are strictly negatively and, in turn, that providing these low voucher amounts can be strictly worse off than the status-quo amount.



## 5.5 Role of Low-tuition Schools in the Program

In summary, our welfare estimates under both the status-quo and counterfactual voucher amounts reveal that voucher provision can have a positive net average benefit. While discussing these results, we specifically noted that they arose due to the presence of low-tuition schools in the program that induce a high welfare benefit for recipients relative to the net costs the government faces when the voucher is redeemed at these schools. We conclude our analysis by investigating the importance of these schools in the program when providing the status-quo voucher amount.

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 private participating schools with prices at most an amount  $\kappa$  and let  $j^\kappa = \arg \max \mathcal{J}^\kappa$  denote the school with the highest price removed from the program. In addition, let  $p^\kappa(\tau) = (p_1(0), \dots, p_{j^\kappa}(0), p_{j^\kappa+1}(\tau), \dots, p_J(\tau))$  denote the prices of the schools in  $\mathcal{J}_v$  under the application of the status-quo amount when schools with prices at most  $\kappa$  are removed, i.e. the status-quo voucher amount is applied to only schools with prices above  $\kappa$ . Then, similar to (8), the average benefit of the status-quo voucher amount absent these schools can be defined by

$$AB^\kappa(\tau_{\text{sq}}) = E[B^\kappa(\tau_{\text{sq}})] . \quad (36)$$

where  $B^\kappa(\tau_{\text{sq}})$  is given by the variable that solves

$$\begin{aligned} & \max \left\{ \max_{j \in \mathcal{J}_g \cup \mathcal{J}_n} U_j(Y_j), \max_{j \in \mathcal{J}} U_j(Y_j - p_j(0)) \right\} \\ & = \\ & \max \left\{ \max_{j \in \mathcal{J}_g \cup \mathcal{J}_n} U_j(Y_j - B^\kappa(\tau_{\text{sq}})), \max_{j \in \mathcal{J}} U_j(Y_j - p_j^\kappa(\tau_{\text{sq}}) - B^\kappa(\tau_{\text{sq}})) \right\} , \end{aligned}$$

Similarly, the average cost and benefit net of costs can be defined by

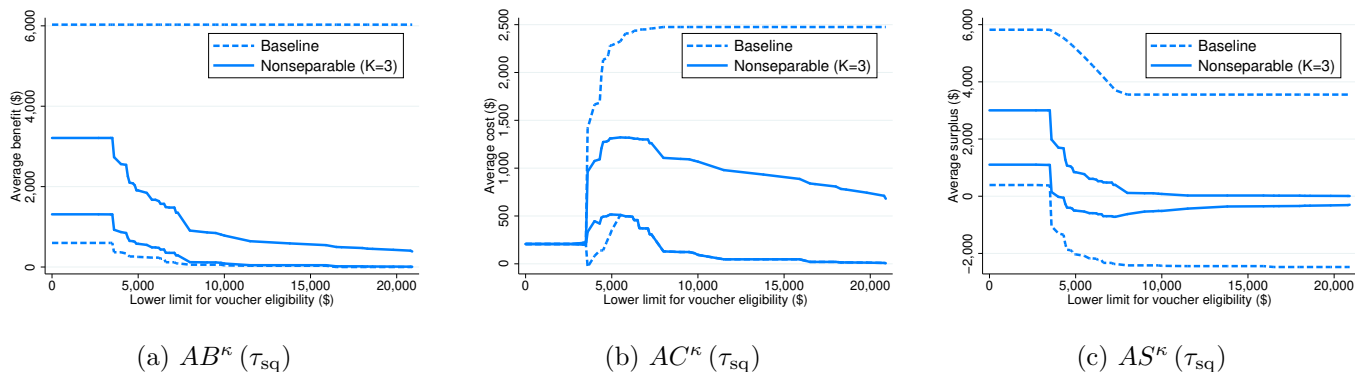
$$AC^\kappa(\tau_{\text{sq}}) = \sum_{j \in \mathcal{J}} c_j^\kappa(\tau_{\text{sq}}) \cdot q_j(p^\kappa(\tau_{\text{sq}})) - \sum_{j \in \mathcal{J}} c_j(0) \cdot q_j(p(0)) , \quad (37)$$

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

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.3.4, we describe how we can continue to use the programs in (32) and (35) 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 present the results for the above parameters for a range of values of  $\kappa$  and, as in Figure 3, for the nonparametric and most flexible parametric specifications from Table 2. The bounds under the nonparametric specification are considerably wider than those under the parametric

Figure 4: Estimates from removing schools with tuition at most  $\kappa$  from the program



(a)  $AB^\kappa (\tau_{sq})$

(b)  $AC^\kappa (\tau_{sq})$

(c)  $AS^\kappa (\tau_{sq})$

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.

specification and especially so for the average benefit, where the upper bound stays constant across all values of  $\kappa$ . This is because the data does 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 absent these schools are switching to the relatively more 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(b) reveals that about 25% of schools in the program are concentrated with tuition 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 analyzed the welfare effects of voucher provision in the DC Opportunity Scholarship Program (OSP). We did so by developing new tools that showed 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 revealed 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.

To conclude, we note while the OSP provided vouchers valid for at least five years, our analysis focused only on school choices collected in the initial year. As a result, unless recipients do not change their choices across years, it is more appropriate to interpret our estimates in some sense as the welfare effects of a voucher that is to be used in the same year. As noted in [Wolf et al. \(2010\)](#), the data reveals there is in fact substantial variation in choices across years. It would be interesting to estimate the welfare effects of the voucher across these years and analyze if the positive results we find continue to hold. To do so, one would potentially need to considerably extend our analysis to introduce and characterize related welfare parameters in some version of a dynamic discrete choice model. We leave such extensions for future research.

## References

- ABDULKADIROĞLU, A., PATHAK, P. A. and WALTERS, C. R. (2018). Free to choose: Can school choice reduce student achievement? *American Economic Journal: Applied Economics*, **10** 175–206.
- 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.
- 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*.
- 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.
- 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.
- 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.
- 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.

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