Online Appendix to "The Demand for Effective Charter Schools"

Christopher R. Walters

UC Berkeley and NBER

August 2017

Contents

Appendix A: Semi-Parametric Identification

This appendix considers semi-parametric identification of a model with one charter school and no heterogeneity in unobserved application costs. I first derive students' optimal application and attendance rules in this version of the model. I then then show that these rules imply a representation equivalent to the single-spell discrete duration model analyzed by Heckman and Navarro (2007). Finally, I apply Theorem 2 in Heckman and Navarro (2007) to give conditions under which the model is semi-parametrically identified.

A.1 Application and Attendance Choices

In a model with one charter school, the expected utility associated with receiving a charter school offer is given by

$$w(1|X_i, D_i, \psi_{i1}) = E\left[\max\left\{v_1(X_i, D_{i1}, D_{i0}) + \psi_{i1} + \xi_{i1}, 0\right\}|X_i, D_i, \psi_{i1}\right],$$

while the expected utility of receiving no offer equals zero. The probability of an offer is π_1, and the cost of applying is $c(X_i, \eta_i)$. The optimal application rule is therefore

$$A_{i1} = 1\left\{\pi w(1|X_i, D_i, \psi_i) - c(X_i, \eta_i) > 0\right\}$$

$$= 1\left\{\pi_1 h\left(v_1(X_i, D_{i1}, D_{i0}) + \psi_{i1}\right) > c(X_i, \eta_i)\right\},$$

where

$$h(v) \equiv E\left[\max\left\{v + \xi_{i1}, 0\right\}\right].$$

We can rewrite the function $h(v)$ as

$$h(v) = F_{-\xi_1}(v) \times [v - K_{-\xi_1}(v)],$$

where $F_{-\xi_1}(v)$ is the CDF of $-\xi_{i1}$ evaluated at v and $K_{-\xi_1}(v)$ is the conditional expectation of $-\xi_{i1}$ truncated from above at v. When ξ_{i1} has full support on the real line, it is straightforward to show that this function satisfies $h(v) > 0\ \forall v$, $\lim_{t\to-\infty} h(v) = 0$, and $h'(v) = F_{-\xi_1}(v) > 0\ \forall v$.

Assuming $c(X_i, \eta_i)$ is always positive, the application rule can then be rewritten

$$A_{i1} = 1\left\{v_1(X_i, D_{i1}, D_{i0}) - h^{-1}\left(c(X_i, \eta_i)/\pi_1\right) + \psi_{i1} > 0\right\}.$$

Due to the non-linearity of the $h^{-1}(\cdot)$ function the left-hand side of this inequality is not additively separable in the observables (X_i, D_{i1}, D_{i0}) and the unobservables (ψ_{i1}, η_i). This is true even when the cost function itself is separable as $c(x, \eta) = c_1(x) + c_2(\eta)$.

To obtain an additively separable representation I consider a special case where there is no cost heterogeneity, so $c(x, \eta) = c(x)$. In this case the application rule can be rewritten as a separable threshold crossing model:

$$A_i = 1\left\{\Omega(X_i, D_{i1}, D_{i0}) < \psi_{i1}\right\},$$

with $\Omega(x, d_1, d_0) = h^{-1}(c(x)/\pi_1) - v_1(x, d_1, d_0)$.

A student attends the charter school if she applies, receives a lottery offer, and the final utility for the charter school exceeds the final utility for public school. Then we can write

$$S_i = A_{i1} \times Z_{i1} \times 1\{v_1(X_i, D_{i1}, D_{i0}) + \psi_{i1} + \xi_{i1} > 0\}.$$

$$= A_{i1} \times Z_{i1} \times 1\{-v_1(X_i, D_{i1}, D_{i0}) < \zeta_{i1}\},$$

where $\zeta_{i1} = \psi_{i1} + \xi_{i1}$.

A.2 Identification

To analyze identification of the model, it is useful to show it is a special case of the discrete duration model analyzed by Heckman and Navarro (2007). Define

$$T_i = 1 + A_{i1} + Z_{i1} + S_i.$$

T_i may be viewed as the number of periods that student i participates in the charter enrollment process. If she chooses not to apply, then $T_i = 1$. If she applies but loses the lottery, then $T_i = 2$. If she applies, wins the lottery, but turns down the charter offer, then $T_i = 3$. If she applies, wins the lottery, and accepts the offer, then $T_i = 4$, which is the maximum duration.

Let B_{it} denote an indicator equal to one if student i decides to stop at step t, observed only when $T_i \geq t$. Each step in this process obeys a separable threshold-crossing model. We have $B_{i1} = 1\{\Omega(X_i, D_{i1}, D_{i0}) \geq \psi_{i1}\}$. At $t = 2$, the student exits when she loses the lottery, which occurs with probability $1 - \pi_1$ at this stage. Hence we can write $B_{i2} = 1\{1 - \pi_1 \geq \nu_{i1}\}$, where we have normalized $\nu_{i1} \sim U(0,1)$ and ν_{i1} is independent of all other variables since the lottery is random. At $t = 3$, we have $B_{i3} = 1\{-v_1(X_i, D_{i1}, D_{i0}) \geq \zeta_{i1}\}$. We define $B_{i0} = 0$ and $B_{i4} = 1 - B_{i3}$. Finally, potential outcomes for $T_i \in \{1, 2, 3\}$ equal Y_{i0}, while the potential outcome for $T_i = 4$ is Y_{i1}.

This argument shows that the one-charter choice model is a special case of the model in Heckman and Navarro (2007). Their Theorem 2 gives identification of the joint distribution of latent utilities and each potential outcome in this model. The following is a restatement of this theorem, slightly adapted to the problem at hand.

Theorem A1: *Suppose charter application, lottery offer, and attendance rules are given by B_{i1}, B_{i2} and B_{i3} as defined above, and that $\Omega(X_i, D_{i1}, D_{i0})$ and $v_1(X_i, D_{i1}, D_{i0})$ are elements of the Matzkin (1992) class of functions (see Appendix A of Heckman and Navarro, 2007 for a definition of this class). Write potential outcomes as $Y_{ij} = \mu_j(X_i) + \epsilon_{ij}$ with $E[\epsilon_{ij}|X_i] = 0$ for $j \in \{0, 1\}$. Suppose a random sample of data on $(X_i, D_{i1}, D_{i0}, A_{i1}, Z_{i1}, S_i, Y_i)$ is available. Assume:*

1. *$(\epsilon_{i1}, \epsilon_{i0}, \psi_{i1}, \zeta_{i1})$ are continuous mean-zero random variables with finite variances and supports with upper limits $(\bar{\epsilon}_1, \bar{\epsilon}_0, \bar{\psi}_1, \bar{\zeta}_1)$ and lower limits $(\underline{\epsilon}_1, \underline{\epsilon}_0, \underline{\psi}_1, \underline{\zeta}_1)$. These conditions also hold for each component of each subvector.*

2. $(\epsilon_{i1}, \epsilon_{i0}, \psi_{i1}, \zeta_{i1}) \perp\!\!\!\perp (X_i, D_{i1}, D_{i0})$.

3. $Supp(\mu_j(X_i), \Omega(X_i, D_{i1}, D_{i0}), v_1(X_i, D_{i1}, D_{i0})) = Supp(\mu_j(X_i)) \times Supp(\Omega(X_i, D_{i1}, D_{i0})) \times Supp(v_1(X_i, D_{i1}, D_{i0}))$.

4. $Supp\,(\Omega(X_i, D_{i1}, D_{i0}), -v_1(X_i, D_{i1}, D_{i0})) \supseteq Supp\,(\psi_{i1}, \zeta_{i1})$.

5. $\nu_{i1} \perp\!\!\!\perp (\epsilon_{i1}, \epsilon_{i0}, \psi_{i1}, \zeta_{i1}, X_i, D_{i1}, D_{i0})$.

Then we can identify $\mu_1(x)$, $\mu_0(x)$, $\Omega(x, d_1, d_0)$, $v_1(x, d_1, d_0)$, and the joint distribution functions $F_{\psi_1\zeta_1\epsilon_1}(\psi_1, \zeta_1, \epsilon_1)$ and $F_{\psi_1\zeta_1\epsilon_0}(\psi_1, \zeta_1, \epsilon_0)$ up to scale if the Matzkin class is specified up to scale, and exactly if a specific normalization is used.

This theorem follows exactly from the argument for Theorem 2 in Heckman and Navarro (2007). The only slight twist is that there is not full support for the lottery offer "choice" index at $t = 2$, but this is irrelevant since by condition 5 ν_{i1} is independent of all other variables in the model. The remaining primitives of the model are then identified by using the joint distribution of (ψ_{i1}, ζ_{i1}) to recover the marginal distribution of $\xi_{i1} = \zeta_{i1} - \psi_{i1}$. The probability π_1 is identified by the offer rate among lottery applicants, and the function $h(\cdot)$ is determined by the distribution of ξ_{i1}. We can then recover $c(x) = \pi_1 h\left(-v_1(x, d_1, d_0) - \Omega(x, d_1, d_0)\right)$ for any (d_1, d_0).

A final observation is that though the assumptions of Theorem A1 are sufficient for identification, they are stronger than necessary for this special case: the model is overidentified. To see this, note that Heckman and Navarro (2007) establish identification of separate potential outcome distributions corresponding to each node in the duration model. Since charter applications and lottery offers are assumed to have no direct effect on outcomes, however, the potential outcomes corresponding to $T_i = 1$, $T_i = 2$ and $T_i = 3$ are known to be the same in this case. With full support of $\Omega(X_i, D_{i1}, D_{i0})$ and $v_1(X_i, D_{i1}, D_{i0})$, there are multiple ways to identify $\mu_0(x)$: we could look at individuals in a limit set with zero probability of applying to charter schools, or we could look at individuals in a limit set with an application probability equal to one and a conditional attendance probability equal to one who are lotteried back into traditional public schools. In principle one could use this fact to weaken the support conditions without sacrificing identification.

Appendix B: Model Fit

The model estimated here fits the data well. This can be seen in Table A4 and Figure A1, which compare observed and model-predicted patterns of heterogeneity across choices, outcomes, and schools. Panel A of Figure A1 splits the sample into deciles based on the model-predicted probability of applying to at least one charter school as a function of observed characteristics and distance. The horizontal axis plots mean predicted application probabilities in these cells, while the vertical axis displays empirical application probabilities. These points lie mostly along the 45 degree line, indicating that the model accurately reproduces differences in application probabilities across groups; the predicted probabilities range from near zero to 0.35, implying that the model captures a substantial amount of heterogeneity in preferences explained by observables. There is slight visual evidence of nonlinearity and an F-test marginally rejects the null hypothesis that all points lie exactly on the line ($p = 0.04$), but in general the model appears to provide a relatively good fit.

To assess whether the model captures heterogeneity in outcomes, Panel B of Figure A1 compares model-predicted and observed mean test scores in deciles of model predictions, separately for charter and non-charter students. Model-predicted outcomes are expected eighth grade math scores conditional on a student's observed characteristics and choices, which implicitly incorporates heterogeneity on both observed and unobserved dimensions. Most points lie close to the 45-degree line and an F-test does not reject the hypothesis that the model fits all moments. Predicted scores exhibit substantial dispersion and there is significant overlap between predictions for charter and non-charter students.

Finally, Table A4 explores the model's capacity to match cross-school heterogeneity in choices and treatment effects. Panel A reports model-based and observed application probabilities for each school while Panel B displays differences in outcomes for lottery winners and losers by school. The model slightly over-predicts application rates and underpredicts the offer takeup rate. Predicted patterns of heterogeneity across schools appear to accurately reflect the observed differences. As in Panel A of Figure A1 the hypothesis that the model fits all moments perfectly is rejected, but overall the model appears to generate a reasonably accurate description of heterogeneity along many dimensions.

Appendix C: 2SLS Weights

This appendix derives the estimand in 2SLS models of the type estimated by Abdulkadiroğlu et al. (2011) and other lottery-based studies of school choice programs. Consider the system

$$Y_i = \beta C_i + \sum_a \gamma_a 1\{A_i = a\} + \epsilon_i,$$

$$C_i = \pi Z_i^{max} + \sum_a \lambda_a 1\{A_i = a\} + \eta_i,$$

where Y_i is a test score, C_i is a charter attendance dummy, Z_i^{max} is a lottery offer dummy, and the sample is assumed to be restricted to lottery applicants. The reduced form corresponding to this system is

$$Y_i = \rho Z_i{}^{max} + \sum \tau_a 1\{A_i = a\} + u_i.$$

The reduced form and first stage are OLS regressions of test scores and charter attendance on the lottery offer with saturated portfolio controls. These equations therefore generate inverse-variance weighted averages of within-portfolio mean differences (Angrist, 1998). Specifically, we have

$$\rho = \sum_a \left(\frac{w_a}{\sum_{a'} w_{a'}}\right) \rho_a,$$

$$\pi = \sum_a \left(\frac{w_a}{\sum_{a'} w_{a'}}\right) \pi_a,$$

where ρ_a and π_a denote coefficients from regressions of Y_i and C_i on Z_i^{max} within lottery portfolio a, and

$$w_a = Pr\left[A_i = a\right] \times Var(Z_i^{max}|A_i = a).$$

Since the 2SLS model is just-identified, the 2SLS estimand β is equal to the ratio of reduced form and first stage coefficients. This implies:

$$\beta = \frac{\rho}{\pi}$$

$$= \frac{\sum_a w_a \rho_a}{\sum_a w_a \pi_a}$$

$$= \frac{\sum_a (w_a \pi_a)(\rho_a/\pi_a)}{\sum_a w_a \pi_a}$$

$$= \sum_a \left(\frac{\omega_a}{\sum_{a'} \omega_{a'}}\right) IV_a,$$

where $IV_a = (\rho_a/\pi_a)$ is a portfolio-specific IV coefficient and

$$\omega_a = w_a \pi_a$$

$$= Pr\left[A_i = a\right] \times Var\left(Z_i^{max}|A_i = a\right) \times \left(Pr\left[C_i = 1|Z_i^{max} = 1, A_i = a\right] - Pr\left[C_i = 1|Z_i^{max} = 0, A_i = a\right]\right).$$

This argument shows that 2SLS estimation with application portfolio fixed effects generates a weighted average of portfolio-specific IV coefficients with weights proportional to the product of sample size, the variance of the offer, and the first stage shift in charter attendance resulting from an offer. These weights are similar to the weights derived in Angrist and Imbens (1995) for 2SLS models with saturated instrument-covariate interactions in the first stage. A saturated first stage generates weights proportional to $w_a \pi_a^2$ rather than $w_a \pi_a$.

Appendix D: Equilibrium Admission Probabilities

D.1 Description of the Game

This appendix describes the determination of equilibrium admission probabilities for use in counterfactual simulations. These probabilities are determined in a Subgame Perfect Nash Equilibrium in which students make utility-maximizing choices as described in Section 3, and schools set admission probabilities to maximize enrollment subject to capacity constraints.

The time of the game follows Figure 1. Strategies in each stage of the game are as follows:

1. Students choose applications.

2. Schools observe students' application choices, and choose their admission probabilities.

3. Offers are randomly assigned among applicants.

4. Students observe their offers and make school choices.

To simplify the game, I assume that the distribution of students is atomless, so schools do not change their admission probabilities in the second stage in response to the application decisions of individual students in the first stage. Students therefore act as "probability takers" in the first stage, in the sense that they do not expect schools to react to their application choices when setting admission probabilities. This implies that the game can be analyzed as if applications and admission probabilities are chosen simultaneously. I analyze the static Nash equilibria of this simultaneous-move game, which are equivalent to Subgame Perfect equilibria of the dynamic game described above.

D.2 Definition of Equilibrium

An equilibrium of the game requires an application rule for each student, a vector of admission probabilities π^*, and a rule for assigning school choices that satisfy the following conditions:

1. The probability that student i chooses application bundle a is given by $q(a|X_i, D_i, \theta_i, \tau_i, \eta_i; \pi^*)$, where q_a is defined as in Section 5 and now explicitly depends on the vector of admission probabilities students expect to face in each lottery.

2. For each school j, π_j^* is chosen to maximize enrollment subject to school j's capacity constraint, taking student application rules as given and assuming that other school choose π_{-j}^*, which denotes the elements of π^* excluding the jth.

3. After receiving the offer vector z, student i chooses school j with probability $p(j|z, X_i, D_i, \theta_i, \tau_i)$, as defined in Section 5.

D.3 School Problem

I begin by deriving a school's optimal admission probability as a function of students' expected admission probabilities and the actions of other schools. Let Λ_j denote the capacity of school j, which is the maximum share of students that can attend school j. Suppose that students anticipate the admission probability vector π^e when making application decisions in the first stage of the model. Their application decisions are described by $q(a|X_i, D_i, \theta_i, \tau_i, \eta_i; \pi^e)$. In addition, suppose that schools other than j admit students with probability π_{-j}. If school j admits students with probability π_j in the second stage, its enrollment is given by

$$e_j(\pi_j, \pi_{-j}, \pi^e) = E\left[\sum_{a\in\{0,1\}^J}\sum_{z\in\{0,1\}^J} q(a|X_i, D_i, \theta_i, \tau_i, \eta_i; \pi^e) f(z|a; \pi_j, \pi_{-j}) p(j|z, X_i, D_i, \theta_i, \tau_i)\right],$$

where $f(z|a; \pi_j, \pi_{-j})$ is the probability mass function for offers, now explicitly written as a function of admission probabilities. School j choose π_j to solve

$$\max_{\pi_j\in[0,1]} e_j(\pi_j, \pi_{-j}, \pi^e) \quad s.t. \quad e_j(\pi_j, \pi_{-j}, \pi^e) \le \Lambda_j. \tag{1}$$

The best response function $\pi_j^{BR}(\pi_{-j}, \pi^e)$ is the solution to problem (1). The optimal admission probability sets school j's enrollment equal to its capacity if possible. The following equation implicitly defines π_j^{BR} at interior solutions:

$$E\left[\sum_a\sum_z q(a|X_i, D_i, \theta_i, \tau_i, \eta_i; \pi^e) f(z|a; \pi_j^{BR}, \pi_{-j}) p(j|z, X_i, D_i, \theta_i, \tau_i)\right] = \Lambda_j.$$

Noting that $p(j|z, x, d, \theta, \tau) = 0$ when $z_j = 0$ and setting the probability mass function for the offer at school j to $a_j\pi_j$, this equation can be rewritten

$$E\left[\sum_{a:a_j=1}\sum_{z:z_j=1} q(a|X_i, D_i, \theta_i, \tau_i, \eta_i; \pi^e) f_{-j}(z_{-j}|a_{-j}; \pi_{-j}) \pi_j^{BR} p(j|X_i, D_i, \theta_i, \tau_i)\right] = \Lambda_j,$$

where z_{-j}, a_{-j}, and f_{-j} are z, a and f excluding the jth elements. An interior solution for π_j^{BR} therefore satisfies

$$\pi_j^{BR} = \frac{\Lambda_j}{E\left[\sum_{a:a_j=1}\sum_{z:z_j=1} q(a|X_i, D_i, \theta_i, \tau_i, \eta_i; \pi^e) f_{-j}(z_{-j}|a_{-j}; \pi_{-j}) p(j|X_i, D_i, \theta_i, \tau_i)\right]}$$

$$\equiv \Gamma_j(\pi_{-j}, \pi^e).$$

If the denominator of Γ_j is sufficiently small, it may exceed one, in which case school j cannot fill its capacity. In this case, the optimal action is to set $\pi_j = 1$ and fill as many seats as possible. This implies that the best response function is given by

$$\pi_j^{BR}(\pi_{-j}, \pi^e) = \min\{\Gamma_j(\pi_{-j}, \pi^e), 1\}.$$

D.4 Existence of Equilibrium

Let $\pi^{BR} : [0,1]^J \to [0,1]^J$ denote the following vector-valued function:

$$\pi^{BR}(\pi) \equiv \left(\pi_1^{BR}(\pi_{-1}, \pi), \ldots, \pi_J^{BR}(\pi_{-J}, \pi)\right).$$

A vector of admission probabilities supports a Nash equilibrium if and only if it is a fixed point of $\pi^{BR}(\pi)$. The following theorem shows that an equilibrium of the game always exists.

Theorem D1: *There exists a $\pi^* \in [0,1]^J$ such that $\pi^{BR}(\pi^*) = \pi^*$.*

Proof: Note that $q(a|X_i, D_i, \theta_i, \tau_i, \eta_i; \pi)$ is continuous in π and strictly positive, $p(j|X_i, D_i, \theta_i, \tau_i)$ is strictly positive when $z_j = 1$, and $f_{-j}(z_{-j}|a_{-j}; \pi_{-j})$ is continuous in π_{-j} and sums to one for each a_{-j}, so the denominator of Γ_j is always non-zero and continuous in π. π_j^{BR} is therefore a composition of continuous functions, and is continuous. Then π^{BR} is a continuous function that maps the compact, convex set $[0,1]^J$ to itself. Brouwer's Fixed Point Theorem immediately applies and π^{BR} has at least one fixed point in $[0,1]^J$.

D.5 Uniqueness of Equilibrium

I next give conditions under which the equilibrium is unique. Define the functions

$$\ell_j(\pi) \equiv \pi_j - \min\{\Gamma_j(\pi_{-j}, \pi), 1\}$$

and let $\ell(\pi) = (\ell_1(\pi), \ldots, \ell_J(\pi))$. A vector π^* supporting an equilibrium satisfies $\ell(\pi^*) = 0$. A sufficient condition for a unique zero of this function and therefore a unique equilibrium is that the Jacobian of $\ell(\pi)$ is a positive dominant diagonal matrix. This requires the following two conditions to hold at every value of $\pi \in [0,1]^J$:

$$\text{1a. } \frac{\partial \ell_j}{\partial \pi_j} > 0 \quad \forall j.$$

$$\text{2a. } \left|\frac{\partial \ell_j}{\partial \pi_j}\right| \geq \sum_{k \neq j} \left|\frac{\partial \ell_j}{\partial \pi_k}\right| \quad \forall j.$$

To gain intuition for when a unique equilibrium is more likely, note that in any equilibrium, admission probabilities must be strictly positive for all schools; an admission rate of zero guarantees zero enrollment, while expected enrollment is positive and less than Λ_j for a sufficiently small positive π_j. When $\pi_j > 0$, we can write Γ_j as

$$\Gamma_j(\pi_{-j}, \pi) = \frac{\Lambda_j \pi_j}{e_j(\pi_j, \pi_{-j}, \pi)}$$

It follows that conditions 1a and 2a are equivalent to the following conditions on the model's enrollment elasticities:

$$\text{1b. } \frac{\partial \log e_j}{\partial \log \pi_j} > \left(\frac{\Lambda_j - e_j}{\Lambda_j} \right) \quad \forall j$$

$$\text{2b. } \frac{\partial \log e_j}{\partial \log \pi_j} \geq \sum_{k \neq j} \left(\frac{\pi_j}{\pi_k} \right) \times \left| \frac{\partial \log e_j}{\partial \log \pi_k} \right| + \left(\frac{\Lambda_j - e_j}{\Lambda_j} \right) \quad \forall j$$

Condition 1b is more likely to hold throughout the parameter space when demand for charter schools is strong, so that $e_j(\pi_j, \pi_{-j}, \pi) > \Lambda_k$ at most values of π. Condition 2b is also more likely to hold in these circumstances, and when the cross elasticities of enrollment at school j with respect to other schools' admission probabilities are small. This occurs when charter demand is more segmented. If preferences for distance are strong enough, for example, each student will consider only the closest charter school, and the cross elasticities are zero, leading to a unique equilibrium. To compute equilibria in the counterfactual simulations, I numerically solved for fixed points of the best response vector $\pi^{BR}(\pi)$. Experimenting with starting values never produced more than one equilibrium in any counterfactual.

Appendix References

Abdulkadiroğlu, A., J. D. Angrist, S. Dynarski, T. J. Kane, and P. A. Pathak (2011): "Accountability and flexibility in public schools: evidence from Boston's charters and pilots," *Quarterly Journal of Economics*, 126(2), 699–748.

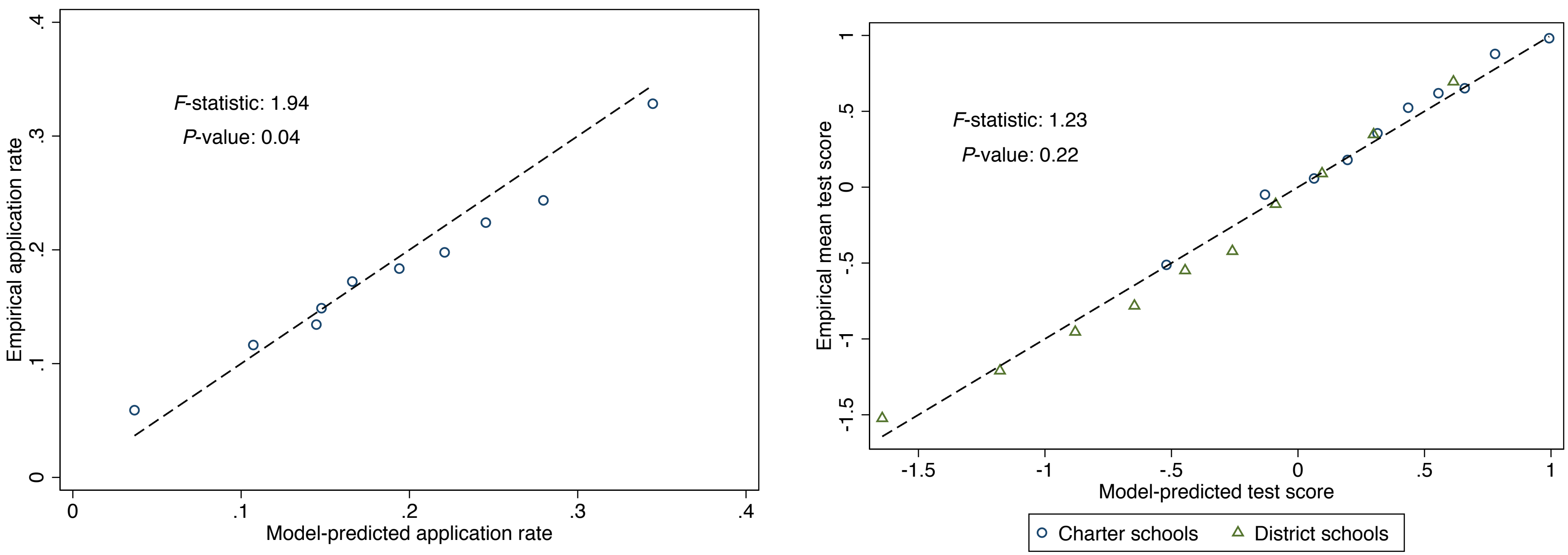
Angrist, J. D. (1998): "Estimating the labor market impact of voluntary military service using Social Security data on military applicants," *Econometrica*, 66, 249–288.

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

Heckman, J. J. and S. Navarro (2007): "Dynamic discrete choice and dynamic treatment effects," *Journal of Econometrics*, 136, 341–396.

Matzkin, R. L. (1992): "Nonparametric and distribution-free estimation of the binary threshold crossing and the binary choice models," *Econometrica*, 60, 239–270.

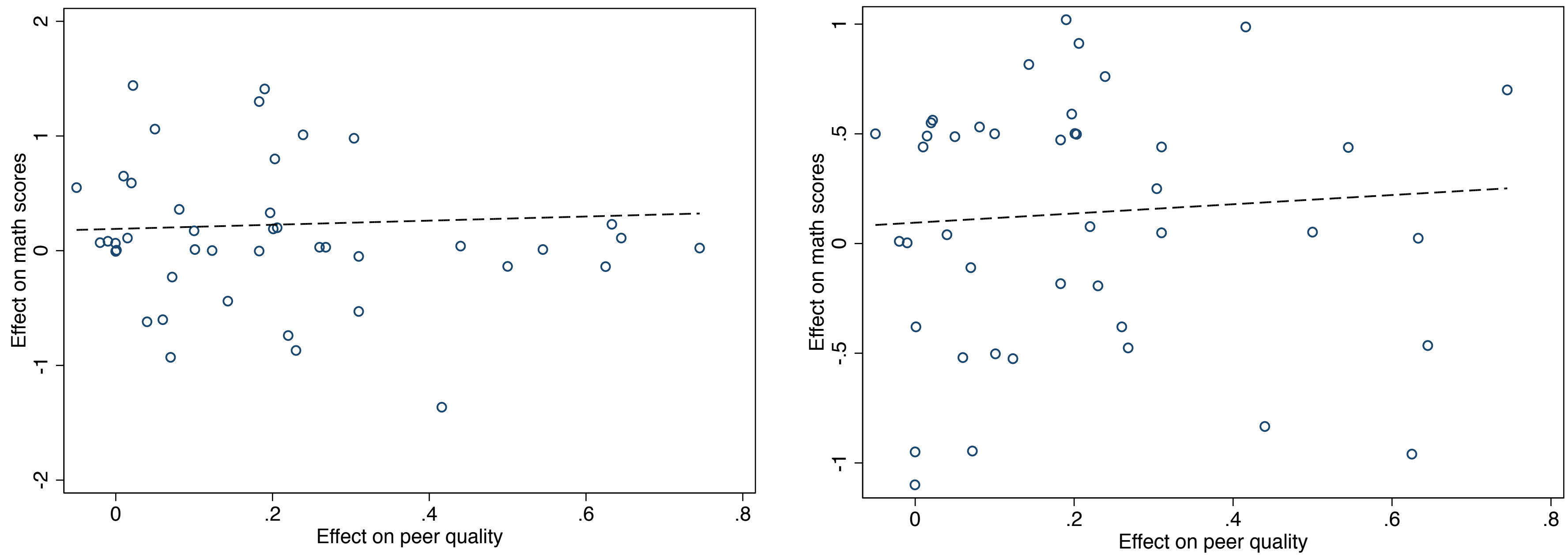
Figure A1: Model fit



A. Application probability *B. Eighth grade math scores*

Notes: This table compares charter application rates and mean test scores to predictions from the two-mass mixture model. Panel A splits the sample into deciles of the model-predicted probability of applying to at least one charter school. Points on the vertical axis are mean observed application rates in these bins, while points on the horizontal axis are means of model-predicted rates. Panel B computes mean model-predicted eighth grade math scores conditional on each student's observed school choice. The sample is split into deciles of this predicted score separately for charter and traditional public schools. Circles plot mean observed scores against mean model predictions for charter schools, while triangles plot corresponding observed and predicted means for traditional public schools. Predictions are averages over 1,000,000 simulations of the two-mass mixture model in column (3) of Table 4, with covariates and spatial locations drawn with replacement from the empirical joint distribution. Dashed lines show the 45-degree line. F-statistics and p-values are from tests of the hypothesis that these lines fit all points perfectly up to sampling error, treating the model predictions as fixed.

Figure A2: Relationship between charter lottery effects on test scores and peer quality



A. Math *B. Reading*

Notes: This figure plots coefficients from regressions of sixth grade test scores on charter lottery offers against coefficients from regressions of peer quality on offers, lottery by lottery. Lotteries are defined as combinations of application cohorts and schools applied. Peer quality for a given student is defined as the average fourth grade test score of the students with whom he or she attends sixth grade. The lines come from OLS regressions of test score effects on peer quality effects, weighting by lottery sample size. The slopes are 0.19 (s.e. = 0.42) for math and 0.21 (s.e. = 0.31) for reading.

Table A1: Boston charter middle schools

	Grade coverage (1)	Years open (2)	Records available (3)	Oversubscribed cohorts (4)
Academy of the Pacific Rim	5-12	1997-	Yes	2006-2009
Boston Collegiate	5-12	1998-	Yes	2006-2009
Boston Preparatory	6-12	2004-	Yes	2006-2009
Edward Brooke	K-8 (with 5th entry)	2002-	Yes	2007-2009
Excel Academy	5-8	2003-	Yes	2008-2009
MATCH Middle School	6-8	2008-	Yes	2007-2009
Smith Leadership Academy	6-8	2003-	No	-
Roxbury Preparatory	6-8	1999-	Yes	2006-2009
Uphams Corner	5-8	2002-2009	No	-

Notes: This table lists charter middle schools in Boston, Massachusetts. Schools are included if they served traditional student populations, accept students in fifth or sixth grade, and operated for cohorts attending fourth grade between 2006 and 2009. Column (2) lists the opening and (where relevant) closing year for each school. Column (3) indicates whether applicant records were available for cohorts attending fourth grade between 2006 and 2009, and column (4) lists the cohorts for which lotteries were held during this period.

Table A2: Covariate balance

	Differential (1)
Female	-0.017
	(0.035)
Black	-0.006
	(0.033)
Hispanic	0.029
	(0.031)
Subsidized lunch	-0.003
	(0.031)
Special education	-0.001
	(0.027)
Limited English proficiency	0.001
	(0.022)
Value-added of public schools in zip code	-0.005
	(0.012)
Fourth grade math score	-0.060
	(0.069)
Fourth grade reading score	0.026
	(0.073)
Miles to closest charter school	-0.045
	(0.076)
Miles to closest district school	0.013
	(0.021)
Joint *p* -value	0.846
N	1601

Notes: This table reports coefficients from regressions of pre-lottery characteristics on a charter lottery offer dummy, controlling for lottery portfolio indicators. The *p* -value is from a test that the coefficients in all regressions are zero.

Table A3: Attrition

	Full sample (1)	Lottery applicants (2)
Followup rate	0.848	0.806
Difference by predicted score	0.055 (0.021)	0.062 (0.047)
Difference by lottery win/loss	-	-0.016 (0.046)
Interaction between win/loss and predicted score	-	0.015 (0.053)
N	10797	1986

Notes: This table reports the fraction of follow-up test scores observed in eighth grade for students attending fourth grade in Boston between 2006 and 2009. Column (1) shows the follow-up rate for the full sample as well as the difference in followup rates between students with above-median and below-median predicted eighth grade math scores. Predicted scores are fitted values from regressions of eighth grade math scores on the baseline variables from Table 1. Column (2) shows the followup rate for lottery applicants along with coefficients from a regression of a followup indicator on the lottery offer, an indicator for an above-median predicted score, and the interaction of the two, controlling for lottery portfolio indicators.

Table A4: Model fit

Charter applications and attendance			Eighth grade math scores				
				Lottery winners		Lottery losers	
	Model	Data		Model	Data	Model	Data
	(1)	(2)		(3)	(4)	(5)	(6)
Apply to charter	0.213	0.175	Charter applicants	0.181	0.236	-0.191	-0.151
Apply to more than one	0.068	0.046	Non-applicants	-	-	-0.459	-0.453
Offer takeup rate	0.774	0.836					
Apply to:			Applicants to:				
Charter 1	0.033	0.031	Charter 1	0.209	0.331	-0.055	0.050
Charter 2	0.046	0.036	Charter 2	0.193	0.405	-0.049	0.110
Charter 3	0.057	0.048	Charter 3	0.106	0.131	-0.051	0.013
Charter 4	0.052	0.039	Charter 4	0.234	0.401	-0.198	-0.207
Charter 5	0.062	0.051	Charter 5	0.117	0.017	-0.072	-0.145
Charter 6	0.019	0.017	Charter 6	0.234	0.241	0.009	-0.073
Charter 7	0.015	0.016	Charter 7	0.471	0.443	0.011	0.088
			χ^2 statistic (d.f.):	40.2 (27)			
			P-value:	0.049			

Notes: This table compares choices and outcomes to predictions from the two-mass mixture model in column (3) of Table 4. Model predictions are averages over 1,000,000 simulations of the model, with covariates and spatial locations drawn with replacement from the empirical joint distribution. The χ^2 statistic and p-value come from a Wald test of the hypothesis that the model fits all observed moments up to sampling error, treating the model predictions as fixed.

Table A5: Estimates of school-specific parameters

	Admisson probability (1)	Mean utility (2)	Math effect (3)	Reading effect (4)
Charter school 1	0.516	-0.736	0.579	0.515
	(0.064)	(0.096)	(0.104)	(0.108)
Charter school 2	0.390	-0.931	0.687	0.610
	(0.057)	(0.103)	(0.105)	(0.110)
Charter school 3	0.653	-0.763	0.619	0.455
	(0.039)	(0.092)	(0.101)	(0.106)
Charter school 4	0.706	-1.264	0.818	0.499
	(0.051)	(0.099)	(0.104)	(0.109)
Charter school 5	0.394	-0.259	0.560	0.234
	(0.074)	(0.095)	(0.099)	(0.103)
Charter school 6	0.824	-2.072	0.772	0.649
	(0.055)	(0.121)	(0.128)	(0.134)
Charter school 7	0.875	-1.670	1.078	0.942
	(0.039)	(0.120)	(0.120)	(0.126)
P-values: no heterogeneity	0.000	0.000	0.000	0.000

Notes: This table reports estimates of school-specific lottery admission probabilities, mean utilities, and test score effects. Estimates come from the two-mass mixture model in column (3) of Table 4. Column (1) shows each charter school's admission probability averaged across applicant cohorts. Column (2) displays mean utility estimates from the two-mass mixture model. Columns (3) and (4) show estimates of average causal effects on eighth grade math and reading scores. *P*-values come from Wald tests of the hypothesis that all parameters in a column are equal.

Table A6: Selection-corrected estimates of charter school effects on sixth and seventh grade test scores

	Sixth grade				Seventh grade			
	Math scores		Reading scores		Math scores		Reading scores	
	Public school outcome	Charter effect	Public school outcome	Charter effect	Public school outcome	Charter effect	Public school outcome	Charter effect
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Constant/main effect	-0.492	0.650	-0.542	0.244	-0.429	0.627	-0.513	0.4307
	(0.014)	(0.085)	(0.015)	(0.090)	(0.014)	(0.088)	(0.015)	(0.089)
Female	-0.001	0.006	0.157	-0.059	0.006	0.130	0.217	-0.031
	(0.014)	(0.043)	(0.015)	(0.045)	(0.015)	(0.044)	(0.015)	(0.045)
Black	-0.208	0.163	-0.155	0.150	-0.207	0.237	-0.098	0.161
	(0.023)	(0.068)	(0.024)	(0.073)	(0.024)	(0.071)	(0.024)	(0.072)
Hispanic	-0.104	0.236	-0.094	0.128	-0.102	0.201	-0.030	0.192
	(0.024)	(0.072)	(0.025)	(0.077)	(0.025)	(0.075)	(0.025)	(0.076)
Subsidized lunch	-0.146	0.163	-0.143	0.061	-0.146	0.174	-0.132	0.170
	(0.020)	(0.053)	(0.021)	(0.056)	(0.021)	(0.054)	(0.021)	(0.055)
Special education	-0.342	-0.015	-0.326	0.006	-0.353	-0.023	-0.399	0.006
	(0.018)	(0.059)	(0.019)	(0.063)	(0.019)	(0.061)	(0.019)	(0.062)
Limited English proficiency	0.041	-0.129	-0.046	-0.005	0.081	-0.073	-0.017	-0.028
	(0.019)	(0.065)	(0.020)	(0.069)	(0.020)	(0.066)	(0.020)	(0.067)
Value-added of closest district school	0.096	-0.015	0.116	-0.115	0.099	-0.062	-0.018	0.108
	(0.046)	(0.130)	(0.049)	(0.138)	(0.047)	(0.133)	(0.048)	(0.135)
Fourth grade math score	0.569	-0.163	0.175	-0.083	0.487	-0.103	0.170	-0.064
	(0.010)	(0.031)	(0.010)	(0.033)	(0.010)	(0.032)	(0.010)	(0.032)
Fourth grade reading score	0.095	0.036	0.460	0.021	0.093	-0.039	0.377	-0.054
	(0.010)	(0.032)	(0.010)	(0.034)	(0.010)	(0.033)	(0.010)	(0.033)
Charter school preference, θ_i	0.055	-0.091	0.023	-0.011	0.046	-0.071	0.039	-0.057
	(0.015)	(0.044)	(0.016)	(0.047)	(0.016)	(0.046)	(0.016)	(0.046)
Idiosyncratic preference, τ_{ij}	-	-0.025	-	0.061	-	0.034	-	0.072
		(0.049)		(0.052)		(0.050)		(0.051)
P-value: No selection on unobservables	0.000		0.438		0.021		0.137	
N	10,122				9,731			

Notes: This table reports selection-corrected estimates of the effects of charter school attendance on sixth and seventh grade test scores. Each pair of columns shows results from a regression of test scores on indicators for attendance at traditional public and charter schools, covariates and their interactions with charter attendance, and control functions correcting for selection on unobservables. The control functions are posterior means from the two-mass mixture model in column (3) of Table 4. Columns (1)-(4) show estimates for sixth grade, while columns (5)-(8) show estimates for seventh grade. P-values are from tests of the hypothesis that the control function coefficients equal zero. Standard errors are adjusted for estimation of the control functions.

Table A7: Selection-corrected estimates of charter school effects on eighth grade test scores for alternative preference models

	Homogeneous charter schools				Heterogeneous charter schools, single normal distribution			
	Math scores		Reading scores		Math scores		Reading scores	
	Public school outcome	Charter effect	Public school outcome	Charter effect	Public school outcome	Charter effect	Public school outcome	Charter effect
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Constant/main effect	-0.433	0.674	-0.510	0.4436	-0.436	0.716	-0.510	0.5807
	(0.008)	(0.077)	(0.008)	(0.081)	(0.008)	(0.067)	(0.008)	(0.070)
Female	-0.025	0.058	0.184	-0.019	-0.025	0.058	0.184	-0.018
	(0.015)	(0.046)	(0.016)	(0.048)	(0.015)	(0.046)	(0.016)	(0.048)
Black	-0.196	0.255	-0.087	0.200	-0.197	0.253	-0.087	0.197
	(0.025)	(0.073)	(0.026)	(0.077)	(0.025)	(0.073)	(0.026)	(0.077)
Hispanic	-0.102	0.257	-0.041	0.242	-0.103	0.257	-0.041	0.245
	(0.025)	(0.077)	(0.027)	(0.081)	(0.026)	(0.077)	(0.027)	(0.081)
Subsidized lunch	-0.129	0.180	-0.127	0.139	-0.127	0.185	-0.127	0.159
	(0.022)	(0.056)	(0.023)	(0.058)	(0.022)	(0.056)	(0.023)	(0.058)
Special education	-0.372	0.095	-0.397	0.140	-0.370	0.093	-0.397	0.136
	(0.020)	(0.065)	(0.021)	(0.068)	(0.020)	(0.065)	(0.021)	(0.068)
Limited English proficiency	0.077	-0.100	0.044	-0.080	0.077	-0.095	0.044	-0.067
	(0.020)	(0.069)	(0.021)	(0.072)	(0.020)	(0.069)	(0.021)	(0.072)
Value-added of closest district school	0.136	-0.009	0.112	-0.045	0.140	-0.007	0.113	-0.033
	(0.049)	(0.138)	(0.051)	(0.144)	(0.049)	(0.138)	(0.051)	(0.144)
Fourth grade math score	0.477	-0.120	0.165	-0.043	0.476	-0.121	0.165	-0.045
	(0.011)	(0.033)	(0.011)	(0.035)	(0.011)	(0.033)	(0.011)	(0.035)
Fourth grade reading score	0.065	-0.014	0.366	-0.075	0.065	-0.015	0.366	-0.080
	(0.011)	(0.034)	(0.011)	(0.035)	(0.011)	(0.034)	(0.011)	(0.035)
Charter school preference, θ_i	0.079	-0.099	0.036	-0.068	0.067	-0.084	0.060	-0.068
	(0.026)	(0.055)	(0.028)	(0.058)	(0.029)	(0.045)	(0.031)	(0.050)
Idiosyncratic preference, τ_{ij}	-	-	-	-	-	0.022	-	0.001
						(0.116)		(0.121)
P-values: No selection on unobservables	0.007		0.366		0.032		0.169	

Notes: This table reports selection-corrected estimates of the effects of charter school attendance on eighth grade test scores using control functions based on the preference models in columns (1) and (2) of Table 4. Each pair of columns shows results from a regression of test scores on indicators for attendance at traditional public and charter schools, covariates and their interactions with charter attendance, and control functions correcting for selection on unobservables. Control functions in columns (1)-(4) are posterior means from the model in column (1) of Table 4. Control functions in columns (5)-(8) are posterior means from the model in column (2) of Table 4. Columns (1), (3), (5) and (7) display public school coefficients, while columns (2), (4), (6), and (8) display interactions with charter attendance. Main effects in columns (6) and (8) are enrollment-weighted averages of effects for the seven schools. P-values are from tests of the hypothesis that the control function coefficients equal zero. Standard errors are adjusted for estimation of the control functions.

Table A8: Selection-corrected estimates of charter school effects on eighth grade math scores for alternative test score transformations

	Percentile rank		Change in percentile rank		Log percentile rank	
	Public school outcome	Charter effect	Public school outcome	Charter effect	Public school outcome	Charter effect
	(1)	(2)	(3)	(4)	(5)	(6)
Constant/main effect	0.489	0.211	-0.005	0.209	-1.027	0.5726
	(0.005)	(0.028)	(0.005)	(0.027)	(0.017)	(0.103)
Female	-0.007	0.019	-0.007	0.017	-0.017	0.045
	(0.005)	(0.014)	(0.005)	(0.014)	(0.017)	(0.051)
Black	-0.059	0.075	-0.051	0.071	-0.111	0.139
	(0.007)	(0.022)	(0.007)	(0.022)	(0.028)	(0.082)
Hispanic	-0.032	0.077	-0.026	0.076	-0.018	0.100
	(0.008)	(0.023)	(0.008)	(0.023)	(0.028)	(0.087)
Subsidized lunch	-0.039	0.057	-0.031	0.049	-0.076	0.126
	(0.007)	(0.017)	(0.006)	(0.017)	(0.024)	(0.063)
Special education	-0.106	0.023	-0.111	0.016	-0.471	0.245
	(0.006)	(0.020)	(0.006)	(0.020)	(0.022)	(0.073)
Limited English proficiency	0.022	-0.028	0.022	-0.027	0.070	-0.047
	(0.006)	(0.021)	(0.006)	(0.021)	(0.023)	(0.077)
Value-added of closest district school	0.041	0.001	0.044	-0.008	0.078	0.055
	(0.015)	(0.042)	(0.015)	(0.041)	(0.055)	(0.155)
Fourth grade math score	0.143	-0.034	-0.125	-0.051	0.437	-0.237
	(0.003)	(0.010)	(0.003)	(0.010)	(0.012)	(0.037)
Fourth grade reading score	0.019	-0.005	0.022	-0.005	0.081	-0.047
	(0.003)	(0.010)	(0.003)	(0.010)	(0.012)	(0.038)
Charter school preference, θ_i	0.018	-0.028	0.017	-0.027	0.052	-0.089
	(0.005)	(0.014)	(0.005)	(0.014)	(0.018)	(0.053)
Idiosyncratic preference, τ_{ij}	-	-0.005	-	-0.003	-	0.008
		(0.016)		(0.016)		(0.059)
P-value: No selection on unobservables	0.001		0.001		0.001	

Notes: This table reports selection-corrected estimates of the effects of charter school attendance on eighth grade math scores for several transformations of test scores. The control functions are posterior means from the two-mass mixture model in column (3) of Table 4. Columns (1) and (2) report results for test scores measured as percentile ranks, columns (3) and (4) show results using the change in percentile rank between fourth and eighth grade, and columns (5) and (6) display results using the log of the percentile rank. Standard errors are adjusted for estimation of the control functions.

Table A9: Selection-corrected estimates for eighth grade test scores with alternative control functions

	Math scores		Reading scores	
	Public school outcome	Charter effect	Public school outcome	Charter effect
	(1)	(2)	(3)	(4)
Constant/main effect	-0.458	0.838	-0.512	0.5328
	(0.010)	(0.135)	(0.010)	(0.141)
Female	-0.024	0.061	0.184	-0.018
	(0.015)	(0.046)	(0.016)	(0.048)
Black	-0.193	0.250	-0.087	0.199
	(0.025)	(0.073)	(0.026)	(0.077)
Hispanic	-0.100	0.260	-0.041	0.243
	(0.025)	(0.078)	(0.027)	(0.081)
Subsidized lunch	-0.128	0.194	-0.126	0.151
	(0.022)	(0.056)	(0.023)	(0.059)
Special education	-0.370	0.098	-0.397	0.135
	(0.020)	(0.065)	(0.021)	(0.068)
Limited English proficiency	0.075	-0.090	0.044	-0.073
	(0.020)	(0.069)	(0.021)	(0.072)
Value-added of closest district school	0.136	0.005	0.113	-0.039
	(0.049)	(0.138)	(0.051)	(0.145)
Fourth grade math score	0.476	-0.122	0.165	-0.044
	(0.011)	(0.033)	(0.011)	(0.035)
Fourth grade reading score	0.066	-0.019	0.366	-0.078
	(0.011)	(0.034)	(0.011)	(0.036)
Type one (high θ_i)	0.164	-0.292	0.145	-0.114
	(0.046)	(0.134)	(0.048)	(0.140)
Idiosyncratic preference, τ_{ij}	-	-0.019	-	0.008
		(0.052)		(0.055)
P-value: No selection on unobservables	0.001		0.044	

Notes: This table reports selection-corrected estimates of the effects of charter school attendance on eighth grade test scores using the posterior type probability from the two-mass mixture model as a control function. Standard errors are adjusted for estimation of the control functions.