Appendix to accompany: Family Networks and School Choice

Andrew Dustan*
Department of Economics
Vanderbilt University

March 2018

*E-mail: andrew.dustan@vanderbilt.edu.

A. Model of school choice

This appendix extends a model of school choice from Hastings, Kane, and Staiger (2009) by incorporating incomplete information, risk aversion, and learning from peers. In my model, the utility from attending each school is uncertain because of incomplete information about student-school match quality. Risk-averse students revise their beliefs about utilities by receiving informative signals about match quality from peers. The setup is similar to models of consumer demand for experience goods, in particular Roberts and Urban (1988) and Erdem and Keane (1996), where consumers are uncertain about product quality and revise their beliefs due to word-of-mouth or informative advertising.[1]

This model produces testable hypotheses about how students react to new information about specific schools. First, the model predicts that the average impact of new information on same-school expected utility is positive. This is a prediction about the average effect of new information over all students and schools in the population, not a prediction that the average effect will be positive for each school. Second, the model predicts that the impact of new information depends on how positive or negative the signal was. Finally, these effects are predicted to apply, to a lesser degree, to other schools that are observably similar to the school about which the information was received.

General setup

The student's problem is to maximize expected utility by choosing one school to attend from his choice set. Here I abstract from the problem of portfolio construction and focus on the first choice. This is reasonable if one thinks that the first listed option is the student's most-preferred school, a modest assumption given the large number of options that a student is allowed to list in order to diversify and choose safety schools.

Student i's utility from school $j \in J$ is a function of student-school match quality:

$$U_{ij} = U\left(\boldsymbol{X_{ij}}\boldsymbol{\beta_i}\right) = U\left(\left(\boldsymbol{\bar{X}_j} + \boldsymbol{\widetilde{X}_{ij}}\right)\boldsymbol{\beta_i}\right)$$

where match quality is expressed as the sum of student-school attributes in the vector $\boldsymbol{X_{ij}}$ weighted by the student-specific vector of preference parameters $\boldsymbol{\beta_i}$. The attribute vector is decomposed into two terms: $\boldsymbol{\bar{X}_j}$ is the average level in the population and $\boldsymbol{\widetilde{X}_{ij}}$ is the student-specific deviation from this level. An example of a student-school attribute is academic fit, which is on average higher at some schools than others, but also has a student-specific component that depends on how well the school caters to the student's particular learning style and ability level.

[1] Students may also gain productive knowledge about schools from their peers, which allows them to obtain higher utility from attending the peer's school. I omit this channel from the model since its effect on choice is obvious and the focus is on the role of information for risk-averse decision-makers.

The student knows the relative weights $\boldsymbol{\beta_i}$ he puts on each attribute. If he also knows $\boldsymbol{X_{ij}}$, and if he is risk-neutral with respect to match quality, so that $U(\boldsymbol{X_{ij}\beta_i}) = \boldsymbol{X_{ij}\beta_i}$, this model is nearly identical to the one in Hastings, Kane, and Staiger (2009). In that case, the student chooses school j if it provides the highest match quality out of all schools in the choice set: $\boldsymbol{X_{ij}\beta_i} > \boldsymbol{X_{ik}\beta_i}\ \forall k \neq j \in J$.[2]

Incomplete information about match quality

Incomplete information about match quality is modeled by making it so that the student imperfectly observes student-school attributes. He does not observe $\boldsymbol{\bar{X}_j}$ or $\boldsymbol{\widetilde{X}_{ij}}$, but he knows the distributions from which each is drawn:

$$\boldsymbol{\bar{X}_j} \sim \mathcal{N}\left(\boldsymbol{\bar{X}_j^0}, \boldsymbol{\Sigma}_{\boldsymbol{\bar{X}_j}}\right), \quad \boldsymbol{\widetilde{X}_{ij}} \sim \mathcal{N}\left(\boldsymbol{\widetilde{X}_{ij}^0}, \boldsymbol{\Sigma}_{\boldsymbol{\widetilde{X}_{ij}}}\right).$$

For simplicity of exposition, the covariance matrices $\boldsymbol{\Sigma}_{\boldsymbol{\bar{X}_j}}$ and $\boldsymbol{\Sigma}_{\boldsymbol{\widetilde{X}_{ij}}}$ are assumed to be diagonal, and $\boldsymbol{\bar{X}_j}$ and $\boldsymbol{\widetilde{X}_{ij}}$ are assumed to be mean independent. Thus $\boldsymbol{X_{ij}}$ is distributed normally with mean $\boldsymbol{X_{ij}^0} = \boldsymbol{\bar{X}_j^0} + \boldsymbol{\widetilde{X}_{ij}^0}$ and diagonal covariance matrix with $(\ell, \ell)^{\text{th}}$ entry $1/\tau_{\ell ij}^0$.[3]

Because $\boldsymbol{X_{ij}}$ is unknown, a risk-neutral student chooses j if it maximizes *expected* match quality: $\mathrm{E}_0[\boldsymbol{X_{ij}\beta_i}] > \mathrm{E}_0[\boldsymbol{X_{ik}\beta_i}]\ \forall k \neq j \in J$, where the 0 subscript indicates that the expectation is formed solely on the basis of the match quality distributions. Incomplete information about match quality (in particular, about mean quality $\boldsymbol{\bar{X}_j}$) is sufficient to predict the results from Hastings and Weinstein (2008), where giving information about school-level average test scores to students increased the weight that students placed on test scores when choosing schools.[4]

Risk aversion

Allowing the student to be risk-averse will address a troubling result from the risk-neutral model. Risk neutrality implies that the relative precision with which match quality is known does not affect choice. That is, presented with a choice between two schools of equal expected match quality but where one's match is known with complete certainty and the other with uncertainty, the student will be indifferent between them. A risk-averse student will prefer the school where match quality is known with certainty.

To model risk aversion, I allow utility to be concave in match quality. Following Roberts

[2]Hastings, Kane, and Staiger (2009) do not explicitly model uncertainty, but they do say that uncertainty about an attribute would lead to a lower effective weight being placed on it.

[3]I assume that for any two schools j and k, $\boldsymbol{X_{ij}}$ and $\boldsymbol{X_{ik}}$ are mean independent.

[4]Intuitively, students were choosing on the basis of both signal and noise about test scores, and the information intervention allowed students to choose on the basis of a stronger signal.

and Urban (1988), I use exponential utility:

$$U_{ij} = -exp\left(-\rho \boldsymbol{X_{ij}}\boldsymbol{\beta_i}\right)$$

where ρ, the coefficient of risk aversion, is assumed to be positive. Due to exponential utility and the joint normal distribution of $\boldsymbol{X_{ij}}$, expected utility from school j in the absence of additional information can be written in terms of the mean and variance (or precision) of the prior distribution of match quality:[5]

$$\begin{aligned} U^*_{0ij} &= \mathrm{E}_0\left[\boldsymbol{X_{ij}}\boldsymbol{\beta_i}\right] - \frac{\rho}{2} Var\left(\boldsymbol{X_{ij}}\boldsymbol{\beta_i}\right) \\ &= \boldsymbol{X^0_{ij}}\boldsymbol{\beta_i} - \frac{\rho}{2}\sum_{\ell}\frac{\beta^2_{\ell i}}{\tau^0_{\ell ij}}. \end{aligned} \tag{1}$$

where $\beta^2_{\ell i}/\tau^0_{\ell ij}$ is the variance of the distribution of match quality from attribute ℓ. The student optimizes with respect to both the mean and variance of match quality, so schools are now "penalized" when beliefs about them are noisier. He chooses the school j that provides the highest expected utility of all available schools: $U^*_{0ij} > U^*_{0ik}\ \forall k \neq j \in J$.

Effect of peer information

When student i's peer attends school j, the student improves on his prior belief about match quality by receiving informative signals about student-school attributes $\boldsymbol{X_{ij}}$. This information comes in the form of an unbiased, noisy signal about each attribute:

$$\boldsymbol{P_{ij}} = \boldsymbol{X_{ij}} + \boldsymbol{\varepsilon_{ij}}, \quad \boldsymbol{\varepsilon_{ij}} \sim \mathcal{N}\left(\boldsymbol{0}, \boldsymbol{\Sigma_{P_{ij}}}\right),$$

where $\boldsymbol{\Sigma_{P_{ij}}}$ is diagonal with entries $1/\tau^P_{\ell ij}$. The signals received are about student-school attributes for student i, not the peer.[6] The idea is that social interactions with the peer allow i to learn more about the school and infer something about how much he will benefit from different aspects of it.

The student uses this new information to update his expected utility from attending school j. Because the prior and signal are both distributed normally and because the covariance matrix for each is diagonal, the form of the posterior distribution of each student-school attribute is simple:

[5] The full expression for expected utility is $\mathrm{E}_0\left[U_{ij}\right] = -exp\left\{-\rho\left(\boldsymbol{X^0_{ij}}\boldsymbol{\beta_i} - \frac{\rho}{2}\sum_\ell \frac{\beta^2_{\ell i}}{\tau^0_{\ell ij}}\right)\right\}$, but since this is strictly monotonically increasing in the terms in braces, this is equivalent to optimizing with respect to equation 1.

[6] This is in contrast with Roberts and Urban (1988), in which only quality for the peer is observed.

$$X^1_{\ell ij} \sim \mathcal{N}\left(\frac{\tau^0_{\ell ij}X^0_{\ell ij} + \tau^P_{\ell ij}P_{\ell ij}}{\tau^0_{\ell ij} + \tau^P_{\ell ij}},\ \frac{1}{\tau^0_{\ell ij} + \tau^P_{\ell ij}}\right)$$

The posterior distribution of each attribute is a precision-weighted average of the prior and signal. The expected utility from j is now

$$U^*_{1ij} = \widehat{\boldsymbol{X}^{\boldsymbol{1}}_{\boldsymbol{ij}}}\boldsymbol{\beta_i} - \frac{\rho}{2}\sum_{\ell}\frac{\beta^2_{\ell i}}{\left(\tau^0_{\ell ij} + \tau^P_{\ell ij}\right)} \tag{2}$$

where $\widehat{\boldsymbol{X}^{\boldsymbol{1}}_{\boldsymbol{ij}}}$ is the mean of the posterior distribution of $\boldsymbol{X}^{\boldsymbol{1}}_{\boldsymbol{ij}}$. To see how the peer signals affected expected utility, compare equations 1 and 2:

$$U^*_{1ij} - U^*_{0ij} = \left(\widehat{\boldsymbol{X}^{\boldsymbol{1}}_{\boldsymbol{ij}}} - \boldsymbol{X}^{\boldsymbol{0}}_{\boldsymbol{ij}}\right)\boldsymbol{\beta_i} + \frac{\rho}{2}\sum_{\ell}\frac{\beta^2_{\ell i}\tau^P_{\ell ij}}{\tau^0_{\ell ij}\left(\tau^0_{\ell ij} + \tau^P_{\ell ij}\right)}. \tag{3}$$

The change in expected utility comes from two sources. The first term is the change in expected match quality. This quantity may be positive or negative depending on the content of the peer signal. Students may learn that the school is a better or worse match for them than they had guessed. The second term is the change in expected utility resulting from the lower variance in the posterior distribution of match quality. This quantity is unambiguously positive. The increased knowledge about match quality works in the school's favor because the risk-averse student is now more certain about how good the match is.

This gives rise to two results, derived at the end of this appendix:

Result 1: *The expected effect of peer information on U^*_{ij}, taken over all students i and schools j, is positive:* $\mathrm{E}_{ij}\left[U^*_{1ij} - U^*_{0ij}\right] > 0$.

This is the key testable hypothesis of the model that distinguishes it from models without channels through which information strictly increases expected utility. It says that, on average, receiving peer information about a school increases the expected utility from attending there. Intuitively, the signal is sometimes better than the student's prior belief and sometimes it is worse, but the average effect on expected match quality is zero. On the other hand, the reduction in uncertainty about match quality always works in the school's favor. Note that the expected effect may be positive for certain schools and negative for others, because mean quality $\bar{\boldsymbol{X}}_{\boldsymbol{j}}$ is drawn from a random distribution. This hypothesis is about the expected effect over all schools.

Result 2: *All else equal, the change in expected utility from j depends positively on how favorable the peer signal about match quality from j was:* $\frac{\partial\left(U^*_{1ij} - U^*_{0ij}\right)}{\partial \boldsymbol{P_{ij}}\boldsymbol{\beta_i}} > 0$.

This hypothesis simply says that when the student receives a relatively good (i.e. high) signal about the match quality from a school, he is more likely to choose that school than if he had received a relatively bad (low) signal.

Shared attributes across schools

Students may know that the level of an attribute is shared across schools. In the empirical setting studied here, schools are divided into subsystems that share important attributes such as curriculum and vocational orientation. In this case, learning about one school in the subsystem also yields useful information about all other schools in the same subsystem. In order to model the shared attributes in a simple way, we can maintain all prior assumptions of the model and additionally assume that for school j in subsystem s, match quality is expressed as $\boldsymbol{X_{ijs}\beta_i} + \mu_{is}$, where $\mu_{is} \equiv \bar{\mu}_s + \widetilde{\mu}_{is}$. The average component of subsystem match quality is distributed $\bar{\mu}_s \sim \mathcal{N}\left(\bar{\mu}_s^0, \sigma_s^2\right)$, the student-specific component is distributed $\widetilde{\mu}_{is} \sim \mathcal{N}\left(\widetilde{\mu}_{is}^0, \eta_{is}^2\right)$, and $1/\tau_{is}^{\mu} \equiv \sigma_s^2 + \eta_{is}^2$. In addition to the signal $\boldsymbol{P_{ij}}$ about unshared attributes, the student receives a signal about the shared attribute:

$$q_{is} = \mu_{is} + \xi_{is}, \quad \xi_{is} \sim \mathcal{N}\left(0, 1/\tau_{is}^{q}\right).$$

When the student receives a signal about school j in subsystem s, he can update his expected utility from a different school k in the same subsystem:

$$U_{1iks}^* - U_{0iks}^* = \left(\widehat{\mu}_{is}^1 - \mu_{is}^0\right) + \frac{\rho}{2}\frac{\tau_{is}^{q}}{\tau_{is}^{\mu}\left(\tau_{is}^{\mu} + \tau_{is}^{q}\right)} \tag{4}$$

where $\widehat{\mu}_{is}^1$ is the mean of the posterior distribution of the shared attribute and μ_{is}^0 is the mean of the prior. This assumption of a shared attribute produces two additional results, derived at the end of the appendix:

Result 3: *The expected effect of peer information on the expected utility from any other school in the same subsystem is positive: indexing the peer's school by j and fixing another school k_j in j's subsystem s_j, $\mathrm{E}_{ij}\left[U_{1ik_js_j}^* - U_{0ik_js_j}^*\right] > 0$.*

On average, receiving a signal about a school increases the expected utility from attending other schools in the same subsystem. The intuition is the same as for Hypothesis 1. Surprises about the match quality from j's subsystem are also surprises about the match quality for all other schools in the subsystem. The surprises cancel out when we average across all schools and students. There is always a reduction in uncertainty about match quality from j's subsystem, which increases expected utility from attending schools in the subsystem.

Result 4: *Suppose the student receives a peer signal about school j in subsystem s. All else* *equal, the change in expected utility from school k in subsystem s depends positively on how favorable the peer signal about subsystem match quality was:* $\frac{\partial\left(U^*_{1iks}-U^*_{0iks}\right)}{\partial q_{is}} > 0.$

The more positive a surprise to the match quality for j's subsystem, the larger is the increase in expected utility from other schools in the same subsystem.

Proofs

Result 1: $\mathrm{E}_{ij}\left[U^*_{1ij} - U^*_{0ij}\right] > 0.$

Proof: Equation 3 gives the expected change, over all students and schools, in expected utilities when a signal is received. This expectation is:

$$\begin{aligned}\mathrm{E}_{ij}\left[U^*_{1ij} - U^*_{0ij}\right] &= \mathrm{E}_{ij}\left[\left(\widehat{\boldsymbol{X}^{\mathbf{1}}_{\boldsymbol{ij}}} - \boldsymbol{X}^{\mathbf{0}}_{\boldsymbol{ij}}\right)\boldsymbol{\beta_i} + \frac{\rho}{2}\sum_{\ell}\frac{\beta^2_{\ell i}\tau^P_{\ell ij}}{\tau^0_{\ell ij}\left(\tau^0_{\ell ij}+\tau^P_{\ell ij}\right)}\right] \\ &= \mathrm{E}_{ij}\left[\widehat{\boldsymbol{X}^{\mathbf{1}}_{\boldsymbol{ij}}}\boldsymbol{\beta_i}\right] - \mathrm{E}_{ij}\left[\boldsymbol{X}^{\mathbf{0}}_{\boldsymbol{ij}}\boldsymbol{\beta_i}\right] + \frac{\rho}{2}\sum_{\ell}\mathrm{E}_{ij}\left[\frac{\tau^P_{\ell ij}}{\tau^0_{\ell ij}\left(\tau^0_{\ell ij}+\tau^P_{\ell ij}\right)}\right].\end{aligned}$$

From the definition of $\widehat{\boldsymbol{X}^{\mathbf{1}}_{\boldsymbol{ij}}}$:

$$\begin{aligned}\mathrm{E}_{ij}\left[\widehat{\boldsymbol{X}^{\mathbf{1}}_{\boldsymbol{ij}}}\boldsymbol{\beta_i}\right] &= \mathrm{E}_{ij}\left[\sum_{\ell}\beta_{\ell i}\frac{\tau^0_{\ell ij}X^0_{\ell ij}+\tau^P_{\ell ij}P_{\ell ij}}{\tau^0_{\ell ij}+\tau^P_{\ell ij}}\right] \\ &= \sum_{\ell}\left\{\mathrm{E}_{ij}\left[\frac{\tau^0_{\ell ij}}{\tau^0_{\ell ij}+\tau^P_{\ell ij}}\beta_{\ell i}X^0_{\ell ij}\right] + \mathrm{E}_{ij}\left[\frac{\tau^P_{\ell ij}}{\tau^0_{\ell ij}+\tau^P_{\ell ij}}\beta_{\ell i}P_{\ell ij}\right]\right\} \\ &= \sum_{\ell}\left\{\mathrm{E}_{ij}\left[\frac{\tau^0_{\ell ij}}{\tau^0_{\ell ij}+\tau^P_{\ell ij}}\beta_{\ell i}X^0_{\ell ij}\right] + \mathrm{E}_{ij}\left[\frac{\tau^P_{\ell ij}}{\tau^0_{\ell ij}+\tau^P_{\ell ij}}\beta_{\ell i}\left(X_{\ell ij}+\varepsilon_{\ell ij}\right)\right]\right\} \\ &= \sum_{\ell}\left\{\mathrm{E}_{ij}\left[\frac{\tau^0_{\ell ij}}{\tau^0_{\ell ij}+\tau^P_{\ell ij}}\beta_{\ell i}X^0_{\ell ij}\right] + \mathrm{E}_{ij}\left[\frac{\tau^P_{\ell ij}}{\tau^0_{\ell ij}+\tau^P_{\ell ij}}\beta_{\ell i}\left(X_{\ell ij}\right)\right]\right\} \\ &= \sum_{\ell}\left\{\mathrm{E}_{ij}\left[\frac{\tau^0_{\ell ij}}{\tau^0_{\ell ij}+\tau^P_{\ell ij}}\beta_{\ell i}X^0_{\ell ij}\right] + \mathrm{E}_{ij}\left[\frac{\tau^P_{\ell ij}}{\tau^0_{\ell ij}+\tau^P_{\ell ij}}\beta_{\ell i}X^0_{\ell ij}\right]\right\} \\ &= \sum_{\ell}\mathrm{E}_{ij}\left[\beta_{\ell i}X^0_{\ell ij}\right] \\ &= \mathrm{E}_{ij}\left[\boldsymbol{X}^{\mathbf{0}}_{\boldsymbol{ij}}\boldsymbol{\beta_i}\right].\end{aligned}$$

Substituting this result back into the original equation, we have:

$$\begin{aligned}\mathrm{E}_{ij}\left[U^*_{1ij}-U^*_{0ij}\right] &= \mathrm{E}_{ij}\left[\widehat{\boldsymbol{X}^{\mathbf{1}}_{\boldsymbol{ij}}}\boldsymbol{\beta_i}\right]-\mathrm{E}_{ij}\left[\boldsymbol{X^0_{ij}\beta_i}\right]+\frac{\rho}{2}\sum_{\ell}\mathrm{E}_{ij}\left[\frac{\tau^P_{\ell ij}}{\tau^0_{\ell ij}\left(\tau^0_{\ell ij}+\tau^P_{\ell ij}\right)}\right]\\ &= \mathrm{E}_{ij}\left[\boldsymbol{X^0_{ij}\beta_i}\right]-\mathrm{E}_{ij}\left[\boldsymbol{X^0_{ij}\beta_i}\right]+\frac{\rho}{2}\sum_{\ell}\mathrm{E}_{ij}\left[\frac{\tau^P_{\ell ij}}{\tau^0_{\ell ij}\left(\tau^0_{\ell ij}+\tau^P_{\ell ij}\right)}\right]\\ &= \frac{\rho}{2}\sum_{\ell}\mathrm{E}_{ij}\left[\frac{\tau^P_{\ell ij}}{\tau^0_{\ell ij}\left(\tau^0_{\ell ij}+\tau^P_{\ell ij}\right)}\right]>0\end{aligned}$$

where the inequality holds because the τ and ρ terms are all positive by definition. ■

Result 2: All else equal, $\frac{\partial\left(U^*_{1ij}-U^*_{0ij}\right)}{\partial\boldsymbol{P_{ij}\beta_i}}>0$.

Proof: Treating $\boldsymbol{\beta_i}$ and $\boldsymbol{X^0_{ij}}$ as fixed:

$$\begin{aligned}\frac{\partial\left(U^*_{1ij}-U^*_{0ij}\right)}{\partial\boldsymbol{P_{ij}\beta_i}} &= \sum_{\ell}\frac{\partial\left(U^*_{1ij}-U^*_{0ij}\right)}{\partial P_{\ell ij}\beta_{\ell i}}=\sum_{\ell}\frac{\partial U^*_{1ij}}{\partial P_{\ell ij}\beta_{\ell i}}\\ =\sum_{\ell}\frac{\partial}{\partial P_{\ell ij}\beta_{\ell i}}\left(\beta_{\ell i}\frac{\tau^0_{\ell ij}X^0_{\ell ij}+\tau^P_{\ell ij}P_{\ell ij}}{\tau^0_{\ell ij}+\tau^P_{\ell ij}}\right) &= \sum_{\ell}\frac{\tau^P_{\ell ij}}{\tau^0_{\ell ij}+\tau^P_{\ell ij}}>0\end{aligned}$$

where the inequality holds because the τ terms are positive. ■

Result 3: indexing the peer's school by j and fixing another school k_j in j's subsystem s_j, $\mathrm{E}_{ij}\left[U^*_{1ik_js_j}-U^*_{0ik_js_j}\right]>0$.

Proof: This is almost identical to the proof for Result 1, except that the student is only receiving information about the shared attribute μ_{is}. From equation 4, again excluding the effect of productive knowledge, the expectation of the change in expected utility from any other school in the same subsystem is:

$$\mathrm{E}_{ij}\left[U^*_{1ik_js_j}-U^*_{0ik_js_j}\right]=\mathrm{E}_{ij}\left[\left(\widehat{\mu}^1_{is}-\mu^0_{is}\right)\right]+\mathrm{E}_{ij}\left[\frac{\rho}{2}\frac{\tau^q_{is}}{\tau^\mu_{is}\left(\tau^\mu_{is}+\tau^q_{is}\right)}\right].$$

Using the steps from the proof of Hypothesis 1, we have that $\mathrm{E}_{ij}\left[\widehat{\mu}_{is}^{1}\right] = \mathrm{E}_{ij}\left[\mu_{is}^{0}\right]$. So:

$$\begin{aligned}
&\mathrm{E}_{ij}\left[\left(\widehat{\mu}_{is}^{1} - \mu_{is}^{0}\right)\right] + \mathrm{E}_{ij}\left[\frac{\rho}{2}\frac{\tau_{is}^{q}}{\tau_{is}^{\mu}\left(\tau_{is}^{\mu} + \tau_{is}^{q}\right)}\right] \\
&= \mathrm{E}_{ij}\left[\left(\mu_{is}^{0} - \mu_{is}^{0}\right)\right] + \mathrm{E}_{ij}\left[\frac{\rho}{2}\frac{\tau_{is}^{q}}{\tau_{is}^{\mu}\left(\tau_{is}^{\mu} + \tau_{is}^{q}\right)}\right] \\
&= \mathrm{E}_{ij}\left[\frac{\rho}{2}\frac{\tau_{is}^{q}}{\tau_{is}^{\mu}\left(\tau_{is}^{\mu} + \tau_{is}^{q}\right)}\right] > 0
\end{aligned}$$

where the inequality holds because the τ and ρ terms are all positive. ■

Result 4: Suppose that schools j and k are in the same subsystem s. Then all else equal, $\frac{\partial\left(U_{1iks}^{*} - U_{0iks}^{*}\right)}{\partial q_{is}} > 0$.

Proof: Treating μ_{is}^{0} as fixed:

$$\frac{\partial\left(U_{1iks}^{*} - U_{0iks}^{*}\right)}{\partial q_{is}} = \frac{\partial U_{1iks}^{*}}{\partial q_{is}} = \frac{\partial \widehat{\mu}_{is}^{1}}{\partial q_{is}} = \frac{\partial}{\partial q_{is}}\left(\frac{\tau_{is}^{\mu}\mu_{is}^{0} + \tau_{is}^{q} q_{is}}{\tau_{is}^{\mu} + \tau_{is}^{q}}\right) = \frac{\tau_{is}^{q}}{\tau_{is}^{\mu} + \tau_{is}^{q}} > 0$$

where the inequality holds because the τ terms are positive. ■

References

Erdem, Tulin, and Michael P. Keane. 1996. "Decision-Making under Uncertainty: Capturing Dynamic Brand Choice Processes in Turbulent Consumer Goods Markets." *Marketing Science* 15 (1): 1–20.

Hastings, Justine, and Jeffrey Weinstein. 2008. "Information, School Choice, and Academic Achievement: Evidence from Two Experiments." *Quarterly Journal of Economics* 123 (4): 1373–1414.

Hastings, Justine S., Thomas J. Kane, and Douglas O. Staiger. 2009. "Heterogeneous Preferences and the Efficacy of Public School Choice." NBER Working Paper No. 12145.

Roberts, John H., and Glen L. Urban. 1988. "Modeling Multiattribute Utility, Risk, and Belief Dynamics for New Consumer Durable Brand Choice." *Management Science* 34 (2): 167–185.

B. Additional tables and figures

Table B.1: Effects of older sibling admission on younger sibling COMIPEMS exam score

	(1) Full	(2) Older siblings at margin of subsystem admission
Score $\geq$ cutoff	-0.113	-0.134
	(0.1158)	(0.1493)
Observations	387255	238096
Adjusted R^2	0.162	0.131
Mean of DV 1 pt below cutoff	62.288	60.821
Bandwidth	8.6	10.7

Note: Regressions include cutoff school-year fixed effects and piecewise-linear polynomial terms in older sibling's centered exam score. Observations are weighted using the edge kernel. Standard errors accounting for clustering at the older sibling level are in parentheses. Stars for statistical significance are based on t-tests using (G-1) degrees of freedom, where G is the number of points of support of the centered score. * p<0.10, ** p<0.05, *** p<0.01.

Table B.2: Predictors of choosing elite school as first choice

	(1) Elite first choice
Parental education (years)	0.011***
	(0.0004)
Male	−0.005**
	(0.0023)
Proportion of older sibling's MS with elite first choice	0.463***
	(0.0075)
log(Older sibling's MS cohort size)	0.014***
	(0.0018)
Distance from closest elite school (km)	−0.002***
	(0.0003)
Observations	129366
Mean of dependent variable	0.755

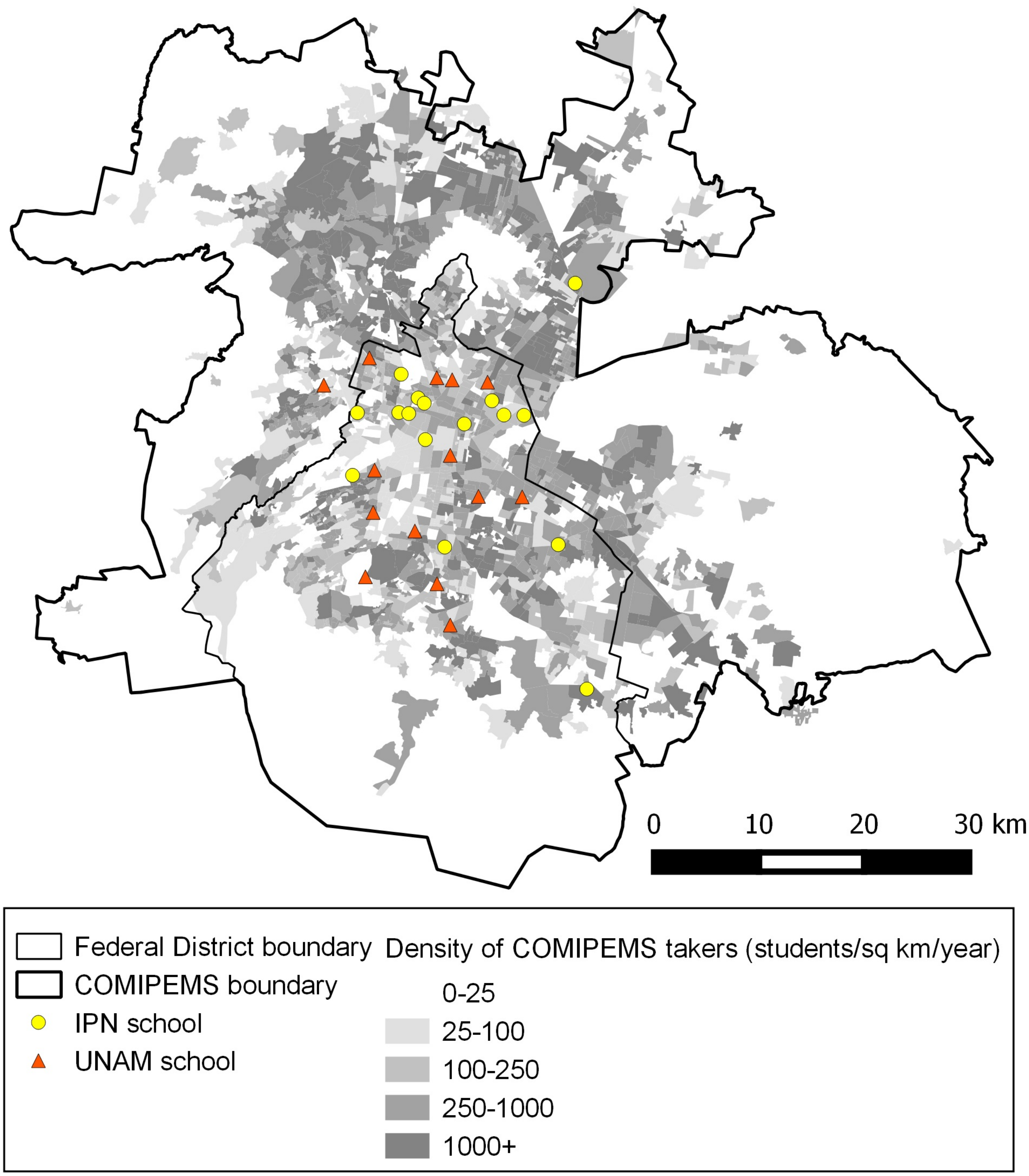
Note: Estimates are average marginal effects from a probit regression. Sample consists of students whose older siblings were below the cutoff of an elite school and were assigned to a non-elite school as a result. Specification also includes younger sibling exam year fixed effects. Standard errors accounting for clustering at the older sibling level are in parentheses. * p<0.10, ** p<0.05, *** p<0.01.

Table B.3: Conditional logit estimates of school choice model

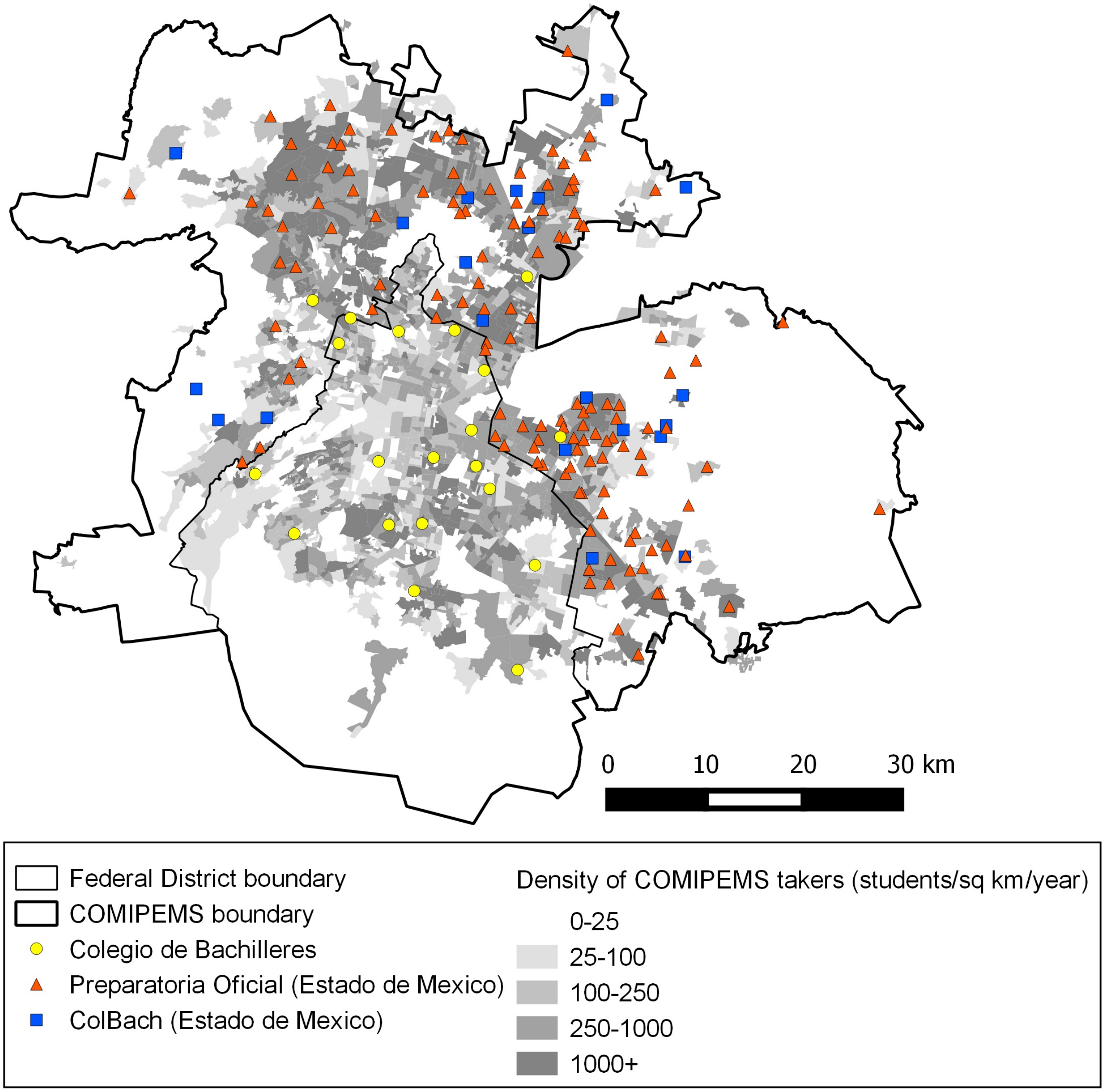
		(1)	(2)	(3)	(4)	(5)	(6)
Sample		Full		Older siblings at margin of subsystem admission			
Model		Conditional logit	Exploded logit	Conditional logit	Exploded logit	Conditional logit	Exploded logit
School characteristic	Interacted with						
School above cutoff	Constant	1.294***	1.305***	0.786***	0.685***	0.438***	0.314***
		(0.023)	(0.016)	(0.037)	(0.022)	(0.039)	(0.024)
	Score ≥ cutoff	0.712***	0.606***	0.600***	0.515***	0.497***	0.391***
		(0.029)	(0.020)	(0.045)	(0.027)	(0.048)	(0.029)
	Centered score	0.006	0.011***	0.001	−0.004	0.002	−0.002
		(0.005)	(0.003)	(0.007)	(0.004)	(0.008)	(0.005)
	Centered score × Score ≥ cutoff	−0.031***	−0.041***	−0.036***	−0.031***	−0.035***	−0.030***
		(0.006)	(0.004)	(0.009)	(0.006)	(0.010)	(0.006)
School below cutoff	Constant	1.845***	1.937***	1.557***	1.386***	1.265***	1.063***
		(0.028)	(0.016)	(0.049)	(0.026)	(0.051)	(0.027)
	Score ≥ cutoff	−1.173***	−0.679***	−0.642***	−0.421***	−0.524***	−0.319***
		(0.043)	(0.021)	(0.071)	(0.036)	(0.072)	(0.037)
	Centered score	0.068***	0.051***	0.029***	0.027***	0.027***	0.025***
		(0.006)	(0.003)	(0.010)	(0.005)	(0.010)	(0.005)
	Centered score × Score ≥ cutoff	−0.074***	−0.058***	−0.040***	−0.025***	−0.045***	−0.027***
		(0.009)	(0.004)	(0.015)	(0.007)	(0.015)	(0.008)
Other school belonging to subsystem above cutoff	Constant			0.625***	0.674***	0.506***	0.568***
				(0.027)	(0.021)	(0.026)	(0.020)
	Score ≥ cutoff			0.237***	0.286***	0.210***	0.261***
				(0.034)	(0.026)	(0.033)	(0.025)
	Centered score			0.008	0.009**	0.009*	0.010**
				(0.005)	(0.004)	(0.005)	(0.004)
	Centered score × Score ≥ cutoff			−0.000	−0.008	−0.001	−0.009*
				(0.007)	(0.005)	(0.007)	(0.005)
Other school belonging to subsystem below cutoff	Constant			0.443***	0.545***	0.363***	0.467***
				(0.036)	(0.024)	(0.036)	(0.024)
	Score ≥ cutoff			−0.203***	−0.189***	−0.177***	−0.170***
				(0.047)	(0.031)	(0.047)	(0.031)
	Centered score			−0.000	−0.003	−0.002	−0.004
				(0.007)	(0.005)	(0.007)	(0.005)
	Centered score × Score ≥ cutoff			−0.001	−0.006	−0.002	−0.007
				(0.009)	(0.006)	(0.009)	(0.006)
Distance from school above cutoff (km)	Constant					−0.049***	−0.050***
						(0.002)	(0.002)
	Score ≥ cutoff					−0.014***	−0.016***
						(0.003)	(0.002)
	Centered score					0.000	0.000
						(0.000)	(0.000)
	Centered score × Score ≥ cutoff					0.000	0.000
						(0.001)	(0.000)
Distance from school below cutoff (km)	Constant					−0.040***	−0.045***
						(0.002)	(0.002)
	Score ≥ cutoff					0.015***	0.015***
						(0.003)	(0.002)
	Centered score					−0.000	−0.000
						(0.000)	(0.000)
	Centered score × Score ≥ cutoff					−0.001	−0.000
						(0.001)	(0.000)
Mean COMIPEMS score of students admitted to school	Constant	0.058***	0.044***	0.052***	0.040***	0.051***	0.040***
		(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Proportion of older sib's MS cohort choosing as 1st choice	Constant	4.338***	3.052***	4.506***	3.343***	4.305***	3.139***
		(0.028)	(0.020)	(0.033)	(0.024)	(0.033)	(0.024)
Distance from student's home	Constant	−0.202***	−0.205***	−0.205***	−0.211***	−0.168***	−0.168***
		(0.001)	(0.000)	(0.001)	(0.001)	(0.001)	(0.001)
Observations		363191	363191	181650	181650	181650	181650

Note: Coefficient estimates are from a conditional logit model for students within 10 points of an admission cutoff. Observations are weighted with respect to centered COMIPEMS score, using the edge kernel. Exploded logit estimates are for the student's top three choices. All specifications include school subsystem fixed effects. Standard errors accounting for clustering at the older sibling level are in parentheses. * p<0.10, ** p<0.05, *** p<0.01.

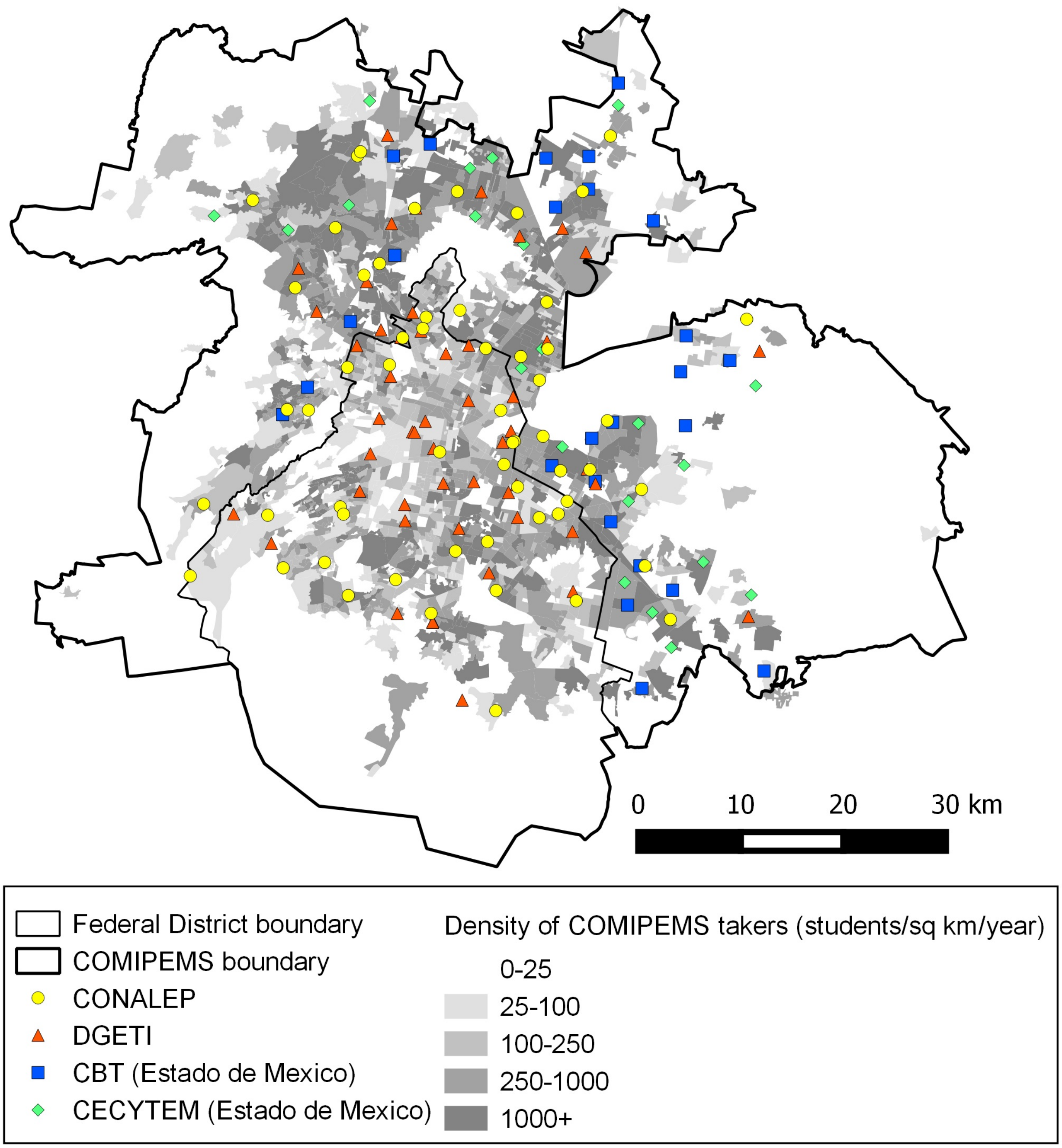
Figure B.1: Map of COMIPEMS zone of Mexico City
A. Elite high schools



B. Non-elite academic high schools



C. Non-elite technical and vocational high schools



C. Sibling matching and related robustness checks

This appendix provides detail on the rate at which older siblings are successfully matched with younger siblings using the matching algorithm described in the paper, as well as how the algorithm's performance may affect the estimated admission effects. The analysis proceeds in three steps. First, among students who are likely to be older siblings due to their observable characteristics, I show correlates of successfully matching with a younger sibling. Second, assuming that locating a match is a function of observable older sibling characteristics, I re-estimate key results of the paper using inverse probability weighting and show that the average admission effects are only minimally altered. Finally, for likely older siblings near an admissions cutoff, I use the RD design to estimate the effect of marginal admission on the probability of finding a matched younger sibling. The evidence points to null or very small admission effects on match rates. Taken together, the evidence suggests that the results in the paper are not driven by the performance of the matching algorithm or its differential success with respect to older siblings' admissions outcomes.

In order to estimate the success rate of the matching algorithm and how it varies with respect to student characteristics, the sample is first restricted to students who completed both the birth order and number of siblings questions on the COMIPEMS demographic questionnaire, and through their responses indicated that they have at least one younger sibling (e.g. birth order is firstborn, number of siblings is one). The sample is further limited to students in 2008 COMIPEMS cohorts or earlier, since later cohorts had two or fewer years within which to locate a match. For the resulting sample of 994,941 likely older siblings, the algorithm's success rate in matching a younger sibling is 39%. Recalling the algorithm's requirements, this is the rate at which students are matched to their younger sibling who is adjacent in birth order, rather than finding any younger sibling.

To explore correlates of matching success, a binary variable for locating a match is regressed on student observables. The results are in Table C.1. Column 1 shows that higher-performing students are more likely to be matched. For example, a one standard deviation increase in COMIPEMS exam score predicts an increase of 5.0 percentage points in match probability, while a one standard deviation increase in middle school GPA predicts a 3.9 percentage point increase. These findings are consistent with low-performing students being more difficult to match, which is expected given that matching requires either a shared phone number or combination of postal code and middle school. Low-performing students are likely to come from households that change addresses and phone numbers more frequently, with associated changes in middle school attended. Column 2 confirms that importance of family background, finding that higher parental education predicts a higher matching rate. In addition, males and firstborns have higher match probabilities, conditional on other covariates.

Table C.1: Correlates of finding a matched younger sibling among likely older siblings

	(1) Matched younger sibling	(2) Matched younger sibling
COMIPEMS examination score	0.003***	0.002***
	(0.0000)	(0.0000)
Middle school GPA	0.047***	0.054***
	(0.0007)	(0.0007)
Male		0.021***
		(0.0010)
Parental education		0.009***
		(0.0001)
Firstborn		0.073***
		(0.0010)
Observations	994941	967222
Adjusted R^2	0.032	0.043
Mean of DV	0.386	0.388

Note: Sample consists of students from 1998-2008 who report birth order and number of siblings such that they appear to have a younger sibling. Robust standard errors in parentheses. * p<0.10, ** p<0.05, *** p<0.01.

Because the matching algorithm disproportionately locates younger siblings from more advantaged backgrounds, it is important to evaluate whether the resulting sample composition significantly affects the estimated effects of marginal admission found in the paper. To do this, I first use the regression results from column 2 of Table C.1 to estimate inverse probability weights for each likely older sibling. Then, in Table C.2, RD admission effects on several key outcomes from the paper are compared between unweighted (Panel A) and weighted (Panel B) specifications. Multiple findings emerge. First, the unweighted estimates are very similar to those found in the paper, even when restricting the sample to likely older siblings, which reduces the sample compared to the analysis in the paper (due to the paper allowing for matches where one or more siblings did not report birth order or number of siblings). Second, inverse probability weighting attenuates the estimated admission effects, but the magnitude of these changes is small. In no case does the magnitude of the coefficient fall by 10% or more as a proportion of the unweighted coefficient. Thus, to the extent that observables are a good proxy for the factors causing differential match rates, the results suggest that the average effects in the main paper are minimally affected by sample composition.

Table C.2: Unweighted and inverse probability weighted regression discontinuity estimates of selected older sibling admission treatment effects

Panel A. No probability weighting	(1) Cutoff school is first choice	(2) Cutoff school is any choice	(3) School below cutoff is first choice	(4) School below cutoff is any choice	(5) Schools chosen in subsystem of cutoff school	(6) Schools chosen in subsystem of school below cutoff	(7) Assigned to cutoff school	(8) Assigned to school below cutoff	(9) Assigned to school in subsystem of cutoff school	(10) Assigned to school in subsystem below cutoff
Score $\geq$ cutoff	0.072***	0.103***	−0.052***	−0.127***	0.205***	−0.156***	0.046***	−0.050***	0.045***	−0.047***
	(0.0026)	(0.0032)	(0.0021)	(0.0042)	(0.0153)	(0.0140)	(0.0015)	(0.0022)	(0.0032)	(0.0037)
Observations	344249	344249	273077	237743	295114	295114	645404	344487	281098	236811
Adjusted R^2	0.159	0.145	0.033	0.113	0.207	0.058	0.027	0.016	0.034	0.018
Mean of DV 1 pt below cutoff	0.146	0.619	0.077	0.633	1.900	1.436	0.072	0.119	0.181	0.229
Bandwidth	9.8	9.7	7.2	6.9	17.1	17.0	19.7	9.5	16.4	13.1

Panel B. Inverse probability weighting	(1) Cutoff school is first choice	(2) Cutoff school is any choice	(3) School below cutoff is first choice	(4) School below cutoff is any choice	(5) Schools chosen in subsystem of cutoff school	(6) Schools chosen in subsystem of school below cutoff	(7) Assigned to cutoff school	(8) Assigned to school below cutoff	(9) Assigned to school in subsystem of cutoff school	(10) Assigned to school in subsystem below cutoff
Score $\geq$ cutoff	0.067***	0.102***	−0.047***	−0.123***	0.188***	−0.147***	0.043***	−0.048***	0.042***	−0.044***
	(0.0025)	(0.0034)	(0.0020)	(0.0044)	(0.0150)	(0.0140)	(0.0015)	(0.0022)	(0.0032)	(0.0037)
Observations	344249	344249	273077	237743	295114	295114	645404	344487	281098	236811
Adjusted R^2	0.154	0.147	0.034	0.115	0.210	0.060	0.028	0.018	0.036	0.021
Mean of DV 1 pt below cutoff	0.146	0.619	0.077	0.633	1.900	1.436	0.072	0.119	0.181	0.229
Bandwidth	9.8	9.7	7.2	6.9	17.1	17.0	19.7	9.5	16.4	13.1

Note: Sample limited to older siblings who report birth order and number of siblings such that they appear to have a younger sibling. Local linear regressions include cutoff school-year fixed effects. Standard errors accounting for clustering at the older sibling level are in parentheses. Stars for statistical significance are based on t-tests using (G-1) degrees of freedom, where G is the number of points of support of the centered score. * p<0.10, ** p<0.05, *** p<0.01.

A further concern is that, for students in the RD sample, admission to the cutoff school may affect the probability of matching a younger sibling. For example, if admission to the cutoff school on average causes students to attend schools that are more or less commute-accessible, it might also change the probability that the family moves and changes its phone number. This could lead to a sharp change in sample composition across the admission cutoff, which may bias the estimated admission effects. The paper provides evidence that is is unlikely: the density of observations is nearly constant across the cutoff and covariates are balanced around the cutoff. The matching rate among likely older siblings near an admissions cutoff provides further evidence that marginal admission is not an important determinant of matching success. Table C.3 presents RD results using the local linear specification from the paper. Column 1 shows no evidence that marginal admission affects the match rate, with a point estimate of 0.3 percentage points compared to the mean match rate of 40% among students one point below the cutoff. Columns 2 through 6 demonstrate that there is only weak evidence for heterogeneous effects of admission with respect to student characteristics. The strongest evidence for heterogeneity is with respect to middle school grade point average, where above-average students experience a 0.8 percentage point higher admission effect on match rate than below-average students. These small point estimates, combined with the density and covariate balance demonstrated in the paper, suggest that differential match rates due to admission are unlikely to be driving the observed older sibling effects on younger sibling school choice.

Table C.3: Effect of cutoff school admission on probability of finding a matching younger sibling among likely older siblings

	(1)	(2)	(3)	(4)	(5)	(6)
Interaction variable	No interaction	Parent has high school education or above	Middle school GPA > median	Male	Firstborn	Cutoff score above sample median
Score ≥ cutoff	0.003	0.003	−0.001	0.004*	0.004*	0.005*
	(0.0017)	(0.0026)	(0.0024)	(0.0023)	(0.0025)	(0.0023)
(Score ≥ cutoff) × (Interaction)		0.002	0.008**	−0.004	−0.003	−0.007*
		(0.0038)	(0.0035)	(0.0036)	(0.0035)	(0.0038)
Observations	1590284	1334789	1590284	1590284	1590284	1590284
Adjusted R^2	0.028	0.031	0.031	0.028	0.036	0.028
Mean of DV 1 pt below cutoff	0.398	0.399	0.398	0.398	0.398	0.398
Bandwidth	18.1	16.0	18.1	18.1	18.1	18.1

Note: Sample consists of students from 1998-2008 who report birth order and number of siblings such that they appear to have a younger sibling. Regressions include cutoff school-year fixed effects and polynomials in student's centered exam score. Observations are weighted using the edge kernel. Standard errors accounting for clustering at the older sibling level are in parentheses. Stars for statistical significance are based on t-tests using (G-1) degrees of freedom, where G is the number of points of support of the centered score. * p<0.10, ** p<0.05, *** p<0.01.

D. Parametric RD estimates

The following tables replicate the analysis in the paper with parametric RD estimators. All use the uniform kernel and a fixed bandwidth of either 5 (Panel A) or 10 (Panels B and C). Panels A and B include piecewise-linear terms in the running variable, while Panel C uses a piecewise-quadratic fit. The analysis is otherwise identical to the table with the corresponding number in the paper. Table D.1 is intentionally omitted to obtain this numbering correspondence.

Table D.2: Effect of older sibling admission on younger sibling choice

Panel A. Linear, BW=5	(1) Cutoff school is first choice	(2) Cutoff school is any choice	(3) School below cutoff is first choice	(4) School below cutoff is any choice
Score $\geq$ cutoff	0.073***	0.106***	−0.051***	−0.124***
	(0.0030)	(0.0038)	(0.0020)	(0.0040)
Observations	235207	235207	233774	233774
Adjusted R^2	0.153	0.143	0.027	0.109
Mean of DV 1 pt below cutoff	0.142	0.606	0.078	0.618
Bandwidth	5.0	5.0	5.0	5.0

Panel B. Linear, BW=10	(1) Cutoff school is first choice	(2) Cutoff school is any choice	(3) School below cutoff is first choice	(4) School below cutoff is any choice
Score $\geq$ cutoff	0.072***	0.105***	−0.049***	−0.118***
	(0.0021)	(0.0026)	(0.0013)	(0.0028)
Observations	459187	459187	456391	456391
Adjusted R^2	0.150	0.135	0.027	0.104
Mean of DV 1 pt below cutoff	0.142	0.606	0.077	0.618
Bandwidth	10.0	10.0	10.0	10.0

Panel C. Quadratic, BW=10	(1) Cutoff school is first choice	(2) Cutoff school is any choice	(3) School below cutoff is first choice	(4) School below cutoff is any choice
Score $\geq$ cutoff	0.073***	0.105***	−0.054***	−0.128***
	(0.0032)	(0.0041)	(0.0022)	(0.0044)
Observations	459187	459187	456391	456391
Adjusted R^2	0.150	0.135	0.027	0.104
Mean of DV 1 pt below cutoff	0.142	0.606	0.077	0.618
Bandwidth	10.0	10.0	10.0	10.0

Table D.3: Effect of older sibling admission on younger sibling's preference for same school, heterogeneity by grade year difference of siblings

Panel A. Linear, BW=5	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Cutoff school is first choice	Cutoff school is first choice	Cutoff school is any choice	Cutoff school is any choice	School below cutoff is first choice	School below cutoff is first choice	School below cutoff is any choice	School below cutoff is any choice
Score $\geq$ cutoff	0.080***	(One per	0.107***	(One per	-0.060***	(One per	-0.122***	(One per
	(0.0043)	cutoff)	(0.0053)	cutoff)	(0.0029)	cutoff)	(0.0056)	cutoff)
(Score $\geq$ cutoff) $\times$ (Sibs 3+ years apart)	-0.014**	-0.016**	-0.005	-0.006	0.017***	0.018***	-0.006	-0.007
	(0.0060)	(0.0060)	(0.0076)	(0.0076)	(0.0040)	(0.0040)	(0.0081)	(0.0081)
Observations	235207	235172	235207	235172	233774	233738	233774	233738
Adjusted R^2	0.150	0.153	0.149	0.151	0.026	0.029	0.112	0.112
Mean of DV 1 pt below cutoff	0.142	0.142	0.606	0.606	0.078	0.078	0.618	0.618
Bandwidth	5.0	5.0	5.0	5.0	5.0	5.0	5.0	5.0

Panel B. Linear, BW=10	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Cutoff school is first choice	Cutoff school is first choice	Cutoff school is any choice	Cutoff school is any choice	School below cutoff is first choice	School below cutoff is first choice	School below cutoff is any choice	School below cutoff is any choice
Score $\geq$ cutoff	0.077***	(One per	0.106***	(One per	-0.054***	(One per	-0.115***	(One per
	(0.0029)	cutoff)	(0.0036)	cutoff)	(0.0019)	cutoff)	(0.0039)	cutoff)
(Score $\geq$ cutoff) $\times$ (Sibs 3+ years apart)	-0.011**	-0.012***	-0.002	-0.004	0.011***	0.011***	-0.007	-0.007
	(0.0041)	(0.0042)	(0.0052)	(0.0052)	(0.0026)	(0.0027)	(0.0056)	(0.0056)
Observations	459187	459173	459187	459173	456391	456375	456391	456375
Adjusted R^2	0.150	0.153	0.139	0.141	0.026	0.029	0.106	0.107
Mean of DV 1 pt below cutoff	0.142	0.142	0.606	0.606	0.077	0.077	0.618	0.618
Bandwidth	10.0	10.0	10.0	10.0	10.0	10.0	10.0	10.0

Panel C. Quadratic, BW=10	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Cutoff school is first choice	Cutoff school is first choice	Cutoff school is any choice	Cutoff school is any choice	School below cutoff is first choice	School below cutoff is first choice	School below cutoff is any choice	School below cutoff is any choice
Score $\geq$ cutoff	0.078***	(One per	0.106***	(One per	-0.060***	(One per	-0.124***	(One per
	(0.0038)	cutoff)	(0.0048)	cutoff)	(0.0026)	cutoff)	(0.0051)	cutoff)
(Score $\geq$ cutoff) $\times$ (Sibs 3+ years apart)	-0.011**	-0.012***	-0.002	-0.004	0.011***	0.011***	-0.007	-0.007
	(0.0041)	(0.0042)	(0.0052)	(0.0052)	(0.0026)	(0.0027)	(0.0056)	(0.0056)
Observations	459187	459173	459187	459173	456391	456375	456391	456375
Adjusted R^2	0.150	0.153	0.139	0.141	0.026	0.029	0.106	0.107
Mean of DV 1 pt below cutoff	0.142	0.142	0.606	0.606	0.077	0.077	0.618	0.618
Bandwidth	10.0	10.0	10.0	10.0	10.0	10.0	10.0	10.0

Table D.4: Effect of older sibling admission on number of other schools chosen in cutoff subsystem

Panel A. Linear, BW=5	(1) Schools chosen in subsystem of cutoff school	(2) Schools chosen in subsystem of cutoff school	(3) Schools chosen in subsystem of cutoff school	(4) Schools chosen in subsystem of school below cutoff	(5) Schools chosen in subsystem of school below cutoff	(6) Schools chosen in subsystem of school below cutoff
Score $\geq$ cutoff	0.188***	0.164***	(One per	−0.155***	−0.180***	(One per
	(0.0233)	(0.0322)	cutoff)	(0.0216)	(0.0311)	cutoff)
(Score $\geq$ cutoff) $\times$ (Sibs 3+ years apart)		0.053	0.055		0.041	0.045
		(0.0473)	(0.0476)		(0.0439)	(0.0443)
Observations	118399	118399	118331	118399	118399	118331
Adjusted R^2	0.199	0.201	0.201	0.051	0.049	0.049
Mean of DV 1 pt below cutoff	1.868	1.868	1.868	1.420	1.420	1.419
Bandwidth	5.0	5.0	5.0	5.0	5.0	5.0

Panel B. Linear, BW=10	(1) Schools chosen in subsystem of cutoff school	(2) Schools chosen in subsystem of cutoff school	(3) Schools chosen in subsystem of cutoff school	(4) Schools chosen in subsystem of school below cutoff	(5) Schools chosen in subsystem of school below cutoff	(6) Schools chosen in subsystem of school below cutoff
Score $\geq$ cutoff	0.203***	0.191***	(One per	−0.154***	−0.151***	(One per
	(0.0159)	(0.0217)	cutoff)	(0.0146)	(0.0208)	cutoff)
(Score $\geq$ cutoff) $\times$ (Sibs 3+ years apart)		0.023	0.024		−0.018	−0.017
		(0.0319)	(0.0320)		(0.0294)	(0.0295)
Observations	231571	231571	231538	231571	231571	231538
Adjusted R^2	0.199	0.202	0.204	0.051	0.052	0.052
Mean of DV 1 pt below cutoff	1.867	1.867	1.867	1.420	1.420	1.420
Bandwidth	10.0	10.0	10.0	10.0	10.0	10.0

Panel C. Quadratic, BW=10	(1) Schools chosen in subsystem of cutoff school	(2) Schools chosen in subsystem of cutoff school	(3) Schools chosen in subsystem of cutoff school	(4) Schools chosen in subsystem of school below cutoff	(5) Schools chosen in subsystem of school below cutoff	(6) Schools chosen in subsystem of school below cutoff
Score $\geq$ cutoff	0.192***	0.178***	(One per	−0.156***	−0.153***	(One per
	(0.0251)	(0.0292)	cutoff)	(0.0233)	(0.0277)	cutoff)
(Score $\geq$ cutoff) $\times$ (Sibs 3+ years apart)		0.023	0.024		−0.018	−0.017
		(0.0319)	(0.0320)		(0.0294)	(0.0295)
Observations	231571	231571	231538	231571	231571	231538
Adjusted R^2	0.199	0.202	0.204	0.051	0.052	0.052
Mean of DV 1 pt below cutoff	1.867	1.867	1.867	1.420	1.420	1.420
Bandwidth	10.0	10.0	10.0	10.0	10.0	10.0

Table D.5: Effect of older sibling admission on younger sibling's choices, heterogeneity by differences in schools above and below cutoff

Panel A. Linear, BW=5	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Sample	Cutoff school and school below cutoff in same subsystem				Cutoff school elite; school below cutoff in different subsystem			
Dependent variable	Cutoff school is first choice	Cutoff school is first choice	Cutoff school is any choice	Cutoff school is any choice	First choice belongs to cutoff school subsystem	First choice belongs to cutoff school subsystem	Schools chosen in cutoff school subsystem	Schools chosen in cutoff school subsystem
Score $\geq$ cutoff	0.075***	(One per	0.077***	(One per	0.159***	(One per	0.473***	(One per
	(0.0062)	cutoff)	(0.0076)	cutoff)	(0.0171)	cutoff)	(0.0801)	cutoff)
(Score $\geq$ cutoff) $\times$ (Dif. in cutoff scores of	0.002	0.001	−0.004	0.014				
schools above and below cutoff $\geq$ median)	(0.0089)	(0.0095)	(0.0107)	(0.0116)				
(Score $\geq$ cutoff) $\times$ (School below cutoff					−0.054**	−0.046**	−0.191*	−0.201*
belongs to non-elite subsystem)					(0.0190)	(0.0193)	(0.0898)	(0.0914)
Observations	116252	116183	116252	116183	59748	59748	59748	59748
Adjusted R^2	0.187	0.190	0.171	0.173	0.165	0.167	0.120	0.121
Mean of DV 1 pt below cutoff	0.163	0.163	0.644	0.644	0.563	0.563	2.284	2.284
Bandwidth	5.0	5.0	5.0	5.0	5.0	5.0	5.0	5.0
Panel B. Linear, BW=10	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Sample	Cutoff school and school below cutoff in same subsystem				Cutoff school elite; school below cutoff in different subsystem			
Dependent variable	Cutoff school is first choice	Cutoff school is first choice	Cutoff school is any choice	Cutoff school is any choice	First choice belongs to cutoff school subsystem	First choice belongs to cutoff school subsystem	Schools chosen in cutoff school subsystem	Schools chosen in cutoff school subsystem
Score $\geq$ cutoff	0.076***	(One per	0.076***	(One per	0.154***	(One per	0.443***	(One per
	(0.0042)	cutoff)	(0.0053)	cutoff)	(0.0117)	cutoff)	(0.0554)	cutoff)
(Score $\geq$ cutoff) $\times$ (Dif. in cutoff scores of	0.002	−0.006	−0.006	0.011				
schools above and below cutoff $\geq$ median)	(0.0060)	(0.0064)	(0.0074)	(0.0080)				
(Score $\geq$ cutoff) $\times$ (School below cutoff					−0.050***	−0.041***	−0.145**	−0.163**
belongs to non-elite subsystem)					(0.0130)	(0.0132)	(0.0621)	(0.0634)
Observations	227450	227406	227450	227406	117326	117326	117326	117326
Adjusted R^2	0.191	0.195	0.162	0.163	0.166	0.167	0.119	0.120
Mean of DV 1 pt below cutoff	0.163	0.163	0.643	0.643	0.563	0.563	2.284	2.284
Bandwidth	10.0	10.0	10.0	10.0	10.0	10.0	10.0	10.0
Panel C. Quadratic, BW=10	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Sample	Cutoff school and school below cutoff in same subsystem				Cutoff school elite; school below cutoff in different subsystem			
Dependent variable	Cutoff school is first choice	Cutoff school is first choice	Cutoff school is any choice	Cutoff school is any choice	First choice belongs to cutoff school subsystem	First choice belongs to cutoff school subsystem	Schools chosen in cutoff school subsystem	Schools chosen in cutoff school subsystem
Score $\geq$ cutoff	0.075***	(One per	0.075***	(One per	0.154***	(One per	0.439***	(One per
	(0.0056)	cutoff)	(0.0068)	cutoff)	(0.0133)	cutoff)	(0.0628)	cutoff)
(Score $\geq$ cutoff) $\times$ (Dif. in cutoff scores of	0.002	−0.006	−0.006	0.011				
schools above and below cutoff $\geq$ median)	(0.0061)	(0.0064)	(0.0074)	(0.0080)				
(Score $\geq$ cutoff) $\times$ (School below cutoff					−0.050***	−0.041***	−0.145**	−0.163**
belongs to non-elite subsystem)					(0.0130)	(0.0133)	(0.0621)	(0.0634)
Observations	227450	227406	227450	227406	117326	117326	117326	117326
Adjusted R^2	0.191	0.195	0.162	0.163	0.166	0.167	0.119	0.120
Mean of DV 1 pt below cutoff	0.163	0.163	0.643	0.643	0.563	0.563	2.284	2.284
Bandwidth	10.0	10.0	10.0	10.0	10.0	10.0	10.0	10.0

Table D.6: Separating effects by sibling pair characteristics, heterogeneity by sibling pair characteristics and differences in schools above and below cutoff

Panel A. Linear, BW=5	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Sample	Same-sex sibling pair				Opposite-sex sibling pair			
Subsample	Cutoff school and school below cutoff in same subsystem		Cutoff school elite; school below cutoff in different subsystem		Cutoff school and school below cutoff in same subsystem		Cutoff school elite; school below cutoff in different subsystem	
Dependent variable	Cutoff school is first choice	Cutoff school is first choice	Cutoff school is any choice	Cutoff school is any choice	First choice belongs to cutoff school subsystem	First choice belongs to cutoff school subsystem	Schools chosen in cutoff school subsystem	Schools chosen in cutoff school subsystem
Score $\geq$ cutoff	(One per cutoff)	(One per cutoff)	(One per cutoff)	(One per cutoff)	(One per cutoff)	(One per cutoff)	(One per cutoff)	(One per cutoff)
(Score $\geq$ cutoff) $\times$ (Dif. in cutoff scores of schools above and below cutoff $\geq$ median)	0.001	−0.005			−0.002	0.027		
	(0.0141)	(0.0164)			(0.0141)	(0.0173)		
(Score $\geq$ cutoff) $\times$ (School below cutoff belongs to non-elite subsystem)			−0.052*	−0.086			−0.040	−0.293*
			(0.0274)	(0.1277)			(0.0276)	(0.1333)
Observations	59154	59154	30530	30530	56020	56020	29215	29215
Adjusted R^2	0.184	0.173	0.171	0.125	0.173	0.189	0.168	0.120
Mean of DV 1 pt below cutoff	0.173	0.665	0.571	2.291	0.154	0.622	0.554	2.275
Bandwidth	5.0	5.0	5.0	5.0	5.0	5.0	5.0	5.0
	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)
Sample	Absolute difference in siblings' middle school GPA < median				Absolute difference in siblings' middle school GPA $\geq$ median			
Subsample	Cutoff school and school below cutoff in same subsystem		Cutoff school elite; school below cutoff in different subsystem		Cutoff school and school below cutoff in same subsystem		Cutoff school elite; school below cutoff in different subsystem	
Dependent variable	Cutoff school is first choice	Cutoff school is first choice	Cutoff school is any choice	Cutoff school is any choice	First choice belongs to cutoff school subsystem	First choice belongs to cutoff school subsystem	Schools chosen in cutoff school subsystem	Schools chosen in cutoff school subsystem
Score $\geq$ cutoff	(One per cutoff)	(One per cutoff)	(One per cutoff)	(One per cutoff)	(One per cutoff)	(One per cutoff)	(One per cutoff)	(One per cutoff)
(Score $\geq$ cutoff) $\times$ (Dif. in cutoff scores of schools above and below cutoff $\geq$ median)	0.019	0.026			−0.007	−0.003		
	(0.0155)	(0.0184)			(0.0145)	(0.0169)		
(Score $\geq$ cutoff) $\times$ (School below cutoff belongs to non-elite subsystem)			−0.048	−0.507***			−0.032	0.022
			(0.0293)	(0.1415)			(0.0279)	(0.1336)
Observations	49168	49168	25885	25885	54342	54342	28105	28105
Adjusted R^2	0.173	0.175	0.164	0.115	0.201	0.199	0.181	0.134
Mean of DV 1 pt below cutoff	0.171	0.647	0.554	2.330	0.165	0.658	0.588	2.356
Bandwidth	5.0	5.0	5.0	5.0	5.0	5.0	5.0	5.0
Panel B. Linear, BW=10	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Sample	Same-sex sibling pair				Opposite-sex sibling pair			
Subsample	Cutoff school and school below cutoff in same subsystem		Cutoff school elite; school below cutoff in different subsystem		Cutoff school and school below cutoff in same subsystem		Cutoff school elite; school below cutoff in different subsystem	
Dependent variable	Cutoff school is first choice	Cutoff school is first choice	Cutoff school is any choice	Cutoff school is any choice	First choice belongs to cutoff school subsystem	First choice belongs to cutoff school subsystem	Schools chosen in cutoff school subsystem	Schools chosen in cutoff school subsystem
Score $\geq$ cutoff	(One per cutoff)	(One per cutoff)	(One per cutoff)	(One per cutoff)	(One per cutoff)	(One per cutoff)	(One per cutoff)	(One per cutoff)
(Score $\geq$ cutoff) $\times$ (Dif. in cutoff scores of schools above and below cutoff $\geq$ median)	−0.003	0.011			−0.008	0.008		
	(0.0095)	(0.0112)			(0.0093)	(0.0118)		
(Score $\geq$ cutoff) $\times$ (School below cutoff belongs to non-elite subsystem)			−0.057***	−0.118			−0.021	−0.214**
			(0.0187)	(0.0875)			(0.0189)	(0.0916)
Observations	116309	116309	60118	60118	110383	110383	57204	57204
Adjusted R^2	0.190	0.158	0.170	0.124	0.184	0.178	0.166	0.119
Mean of DV 1 pt below cutoff	0.172	0.664	0.571	2.291	0.153	0.620	0.554	2.275
Bandwidth	10.0	10.0	10.0	10.0	10.0	10.0	10.0	10.0
	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)
Sample	Absolute difference in siblings' middle school GPA < median				Absolute difference in siblings' middle school GPA $\geq$ median			
Subsample	Cutoff school and school below cutoff in same subsystem		Cutoff school elite; school below cutoff in different subsystem		Cutoff school and school below cutoff in same subsystem		Cutoff school elite; school below cutoff in different subsystem	
Dependent variable	Cutoff school is first choice	Cutoff school is first choice	Cutoff school is any choice	Cutoff school is any choice	First choice belongs to cutoff school subsystem	First choice belongs to cutoff school subsystem	Schools chosen in cutoff school subsystem	Schools chosen in cutoff school subsystem
Score $\geq$ cutoff	(One per cutoff)	(One per cutoff)	(One per cutoff)	(One per cutoff)	(One per cutoff)	(One per cutoff)	(One per cutoff)	(One per cutoff)
(Score $\geq$ cutoff) $\times$ (Dif. in cutoff scores of schools above and below cutoff $\geq$ median)	0.003	0.006			−0.001	0.007		
	(0.0103)	(0.0125)			(0.0096)	(0.0115)		
(Score $\geq$ cutoff) $\times$ (School below cutoff belongs to non-elite subsystem)			−0.049**	−0.256**			−0.023	−0.142
			(0.0201)	(0.0976)			(0.0191)	(0.0921)
Observations	96573	96573	50648	50648	107072	107072	55129	55129
Adjusted R^2	0.181	0.165	0.167	0.114	0.208	0.183	0.179	0.132
Mean of DV 1 pt below cutoff	0.170	0.645	0.554	2.330	0.164	0.656	0.588	2.356
Bandwidth	10.0	10.0	10.0	10.0	10.0	10.0	10.0	10.0

Table D.6 (continued)

Panel C. Quadratic, BW=10	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Sample	Same-sex sibling pair				Opposite-sex sibling pair			
Subsample	Cutoff school and school below cutoff in same subsystem		Cutoff school elite; school below cutoff in different subsystem		Cutoff school and school below cutoff in same subsystem		Cutoff school elite; school below cutoff in different subsystem	
Dependent variable	Cutoff school is first choice	Cutoff school is first choice	Cutoff school is any choice	Cutoff school is any choice	First choice belongs to cutoff school subsystem	First choice belongs to cutoff school subsystem	Schools chosen in cutoff school subsystem	Schools chosen in cutoff school subsystem
Score $\geq$ cutoff	(One per cutoff)	(One per cutoff)	(One per cutoff)	(One per cutoff)	(One per cutoff)	(One per cutoff)	(One per cutoff)	(One per cutoff)
(Score $\geq$ cutoff) $\times$ (Dif. in cutoff scores of schools above and below cutoff $\geq$ median)	-0.003	0.011			-0.008	0.008		
	(0.0095)	(0.0112)			(0.0093)	(0.0118)		
(Score $\geq$ cutoff) $\times$ (School below cutoff belongs to non-elite subsystem)			-0.057***	-0.118			-0.021	-0.215**
			(0.0187)	(0.0875)			(0.0189)	(0.0916)
Observations	116309	116309	60118	60118	110383	110383	57204	57204
Adjusted R^2	0.189	0.158	0.170	0.124	0.184	0.178	0.166	0.119
Mean of DV 1 pt below cutoff	0.172	0.664	0.571	2.291	0.153	0.620	0.554	2.275
Bandwidth	10.0	10.0	10.0	10.0	10.0	10.0	10.0	10.0
	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)
Sample	Absolute difference in siblings' middle school GPA < median				Absolute difference in siblings' middle school GPA $\geq$ median			
Subsample	Cutoff school and school below cutoff in same subsystem		Cutoff school elite; school below cutoff in different subsystem		Cutoff school and school below cutoff in same subsystem		Cutoff school elite; school below cutoff in different subsystem	
Dependent variable	Cutoff school is first choice	Cutoff school is first choice	Cutoff school is any choice	Cutoff school is any choice	First choice belongs to cutoff school subsystem	First choice belongs to cutoff school subsystem	Schools chosen in cutoff school subsystem	Schools chosen in cutoff school subsystem
Score $\geq$ cutoff	(One per cutoff)	(One per cutoff)	(One per cutoff)	(One per cutoff)	(One per cutoff)	(One per cutoff)	(One per cutoff)	(One per cutoff)
(Score $\geq$ cutoff) $\times$ (Dif. in cutoff scores of schools above and below cutoff $\geq$ median)	0.003	0.006			-0.001	0.007		
	(0.0103)	(0.0125)			(0.0096)	(0.0115)		
(Score $\geq$ cutoff) $\times$ (School below cutoff belongs to non-elite subsystem)			-0.049**	-0.256**			-0.023	-0.142
			(0.0201)	(0.0977)			(0.0191)	(0.0921)
Observations	96573	96573	50648	50648	107072	107072	55129	55129
Adjusted R^2	0.181	0.165	0.167	0.114	0.208	0.183	0.179	0.132
Mean of DV 1 pt below cutoff	0.170	0.645	0.554	2.330	0.164	0.656	0.588	2.356
Bandwidth	10.0	10.0	10.0	10.0	10.0	10.0	10.0	10.0

Table D.7: Differential effect of older sibling admission on school choice, by student and sibling pair characteristics

Panel A. Linear, BW=5	(1)	(2)	(3)	(4)	(5)	(6)
Interaction variable	Older sibling is firstborn	Parent has high school education or above	Younger sibling has middle school GPA > median	Younger sibling has hours studied > median	Younger sib MS average COMIPEMS score > median	Distance from home to cutoff school < 2km
Dependent variable	Cutoff school is first choice	Cutoff school is first choice	Cutoff school is first choice	Cutoff school is first choice	Cutoff school is first choice	Cutoff school is first choice
(Score $\geq$ cutoff) $\times$ (Interaction)	0.023***	0.001	0.016**	0.012	-0.004	0.005
	(0.0069)	(0.0067)	(0.0064)	(0.0066)	(0.0062)	(0.0109)
Observations	191557	201888	210961	202290	230977	222060
Adjusted R^2	0.149	0.152	0.160	0.151	0.154	0.163
Mean of DV 1 pt below cutoff	0.144	0.145	0.144	0.145	0.143	0.141
Bandwidth	5.0	5.0	5.0	5.0	5.0	5.0
Panel B. Linear, BW=10	(1)	(2)	(3)	(4)	(5)	(6)
Interaction variable	Older sibling is firstborn	Parent has high school education or above	Younger sibling has middle school GPA > median	Younger sibling has hours studied > median	Younger sib MS average COMIPEMS score > median	Distance from home to cutoff school < 2km
Dependent variable	Cutoff school is first choice	Cutoff school is first choice	Cutoff school is first choice	Cutoff school is first choice	Cutoff school is first choice	Cutoff school is first choice
(Score $\geq$ cutoff) $\times$ (Interaction)	0.027***	-0.005	0.019***	0.008*	-0.006	0.013
	(0.0047)	(0.0046)	(0.0044)	(0.0046)	(0.0042)	(0.0074)
Observations	373653	394345	411692	395119	450972	433342
Adjusted R^2	0.151	0.153	0.159	0.152	0.154	0.162
Mean of DV 1 pt below cutoff	0.144	0.145	0.144	0.145	0.142	0.141
Bandwidth	10.0	10.0	10.0	10.0	10.0	10.0
Panel C. Quadratic, BW=10	(1)	(2)	(3)	(4)	(5)	(6)
Interaction variable	Older sibling is firstborn	Parent has high school education or above	Younger sibling has middle school GPA > median	Younger sibling has hours studied > median	Younger sib MS average COMIPEMS score > median	Distance from home to cutoff school < 2km
Dependent variable	Cutoff school is first choice	Cutoff school is first choice	Cutoff school is first choice	Cutoff school is first choice	Cutoff school is first choice	Cutoff school is first choice
(Score $\geq$ cutoff) $\times$ (Interaction)	0.027***	-0.005	0.019***	0.008*	-0.006	0.013
	(0.0047)	(0.0046)	(0.0044)	(0.0046)	(0.0042)	(0.0074)
Observations	373653	394345	411692	395119	450972	433342
Adjusted R^2	0.151	0.153	0.159	0.152	0.154	0.162
Mean of DV 1 pt below cutoff	0.144	0.145	0.144	0.145	0.142	0.141
Bandwidth	10.0	10.0	10.0	10.0	10.0	10.0

Table D.8: Differential effect of older sibling admission on school choice by graduation outcome

Panel A. Linear, BW=5	(1)	(2)	(3)	(4)	(5)	(6)
Sample	Full		Older siblings at non-elite cutoff		Older siblings at margin of subsystem admission	
Dependent variable	Cutoff school is first choice	Cutoff school is first choice	Cutoff school is first non-elite choice	Cutoff school is first non-elite choice	Schools chosen in cutoff school subsystem	Schools chosen in cutoff school subsystem
Score $\geq$ cutoff	0.023***	(One per	0.054***	(One per	0.101	(One per
	(0.0071)	cutoff)	(0.0116)	cutoff)	(0.0659)	cutoff)
(Score $\geq$ cutoff) $\times$ (Older sibling graduated)	0.021*	0.020*	0.071***	0.076***	0.160	0.166
	(0.0099)	(0.0102)	(0.0155)	(0.0162)	(0.0893)	(0.0926)
Observations	58497	58473	46284	46260	29099	29058
Adjusted R^2	0.071	0.072	0.096	0.097	0.206	0.204
Mean of DV 1 pt below cutoff	0.085	0.085	0.165	0.165	1.521	1.521
Bandwidth	5.0	5.0	5.0	5.0	5.0	5.0

Panel B. Linear, BW=10	(1)	(2)	(3)	(4)	(5)	(6)
Sample	Full		Older siblings at non-elite cutoff		Older siblings at margin of subsystem admission	
Dependent variable	Cutoff school is first choice	Cutoff school is first choice	Cutoff school is first non-elite choice	Cutoff school is first non-elite choice	Schools chosen in cutoff school subsystem	Schools chosen in cutoff school subsystem
Score $\geq$ cutoff	0.023***	(One per	0.039***	(One per	0.117**	(One per
	(0.0049)	cutoff)	(0.0078)	cutoff)	(0.0449)	cutoff)
(Score $\geq$ cutoff) $\times$ (Older sibling graduated)	0.022***	0.020**	0.071***	0.073***	0.101	0.098
	(0.0068)	(0.0070)	(0.0106)	(0.0110)	(0.0601)	(0.0621)
Observations	110194	110193	87295	87294	55203	55186
Adjusted R^2	0.074	0.076	0.099	0.101	0.209	0.212
Mean of DV 1 pt below cutoff	0.085	0.085	0.165	0.165	1.520	1.521
Bandwidth	10.0	10.0	10.0	10.0	10.0	10.0

Panel C. Quadratic, BW=10	(1)	(2)	(3)	(4)	(5)	(6)
Sample	Full		Older siblings at non-elite cutoff		Older siblings at margin of subsystem admission	
Dependent variable	Cutoff school is first choice	Cutoff school is first choice	Cutoff school is first non-elite choice	Cutoff school is first non-elite choice	Schools chosen in cutoff school subsystem	Schools chosen in cutoff school subsystem
Score $\geq$ cutoff	0.019***	(One per	0.051***	(One per	0.140**	(One per
	(0.0063)	cutoff)	(0.0100)	cutoff)	(0.0567)	cutoff)
(Score $\geq$ cutoff) $\times$ (Older sibling graduated)	0.022***	0.020**	0.071***	0.073***	0.102	0.099
	(0.0068)	(0.0070)	(0.0106)	(0.0110)	(0.0601)	(0.0622)
Observations	110194	110193	87295	87294	55203	55186
Adjusted R^2	0.074	0.076	0.099	0.101	0.209	0.212
Mean of DV 1 pt below cutoff	0.085	0.085	0.165	0.165	1.520	1.521
Bandwidth	10.0	10.0	10.0	10.0	10.0	10.0

Table D.9: Effect of older sibling admission on younger sibling assignment outcomes

Panel A. Linear, BW=5	(1)	(2)	(3)	(4)
Sample	Full		Older siblings at margin of subsystem admission	
Dependent variable	Assigned to cutoff school	Assigned to school below cutoff	Assigned to school in subsystem of cutoff school	Assigned to school in subsystem below cutoff
Score $\geq$ cutoff	0.044***	-0.050***	0.038***	-0.046***
	(0.0024)	(0.0025)	(0.0048)	(0.0049)
Observations	235381	235381	118399	118399
Adjusted R^2	0.023	0.012	0.027	0.011
Mean of DV 1 pt below cutoff	0.071	0.116	0.181	0.224
Bandwidth	5.0	5.0	5.0	5.0

Panel B. Linear, BW=10	(1)	(2)	(3)	(4)
Sample	Full		Older siblings at margin of subsystem admission	
Dependent variable	Assigned to cutoff school	Assigned to school below cutoff	Assigned to school in subsystem of cutoff school	Assigned to school in subsystem below cutoff
Score $\geq$ cutoff	0.046***	-0.048***	0.045***	-0.046***
	(0.0017)	(0.0017)	(0.0033)	(0.0033)
Observations	459529	459529	231571	231571
Adjusted R^2	0.024	0.012	0.028	0.011
Mean of DV 1 pt below cutoff	0.071	0.116	0.181	0.224
Bandwidth	10.0	10.0	10.0	10.0

Panel C. Quadratic, BW=10	(1)	(2)	(3)	(4)
Sample	Full		Older siblings at margin of subsystem admission	
Dependent variable	Assigned to cutoff school	Assigned to school below cutoff	Assigned to school in subsystem of cutoff school	Assigned to school in subsystem below cutoff
Score $\geq$ cutoff	0.045***	-0.053***	0.039***	-0.050***
	(0.0026)	(0.0027)	(0.0051)	(0.0053)
Observations	459529	459529	231571	231571
Adjusted R^2	0.024	0.012	0.028	0.011
Mean of DV 1 pt below cutoff	0.071	0.116	0.181	0.224
Bandwidth	10.0	10.0	10.0	10.0

Table D.10: Effect of older sibling admission to an elite school on younger sibling elite school choice and assignment

Panel A1. Linear, BW=5	(1)	(2)	(3)
Dependent variable	Elite school as first choice		
Baseline predicted probability of selecting elite school as first choice	[0, .6]	(.6, .8]	(.8, 1]
Score $\geq$ cutoff	0.105***	0.067***	0.022**
	(0.0290)	(0.0140)	(0.0085)
Observations	4991	14654	18941
Adjusted R^2	0.034	0.022	0.020
Mean of DV 1 pt below cutoff	0.584	0.763	0.903
Bandwidth	5.0	5.0	5.0

Panel A2. Linear, BW=5	(1)	(2)	(3)
Dependent variable	Number of elite schools chosen		
Baseline predicted probability of selecting elite school as first choice	[0, .6]	(.6, .8]	(.8, 1]
Score $\geq$ cutoff	0.575***	0.323***	0.200*
	(0.1553)	(0.0985)	(0.0920)
Observations	4991	14654	18941
Adjusted R^2	0.048	0.048	0.071
Mean of DV 1 pt below cutoff	2.412	3.840	5.550
Bandwidth	5.0	5.0	5.0

Panel A3. Linear, BW=5	(1)	(2)	(3)
Dependent variable	Elite assignment		
Baseline predicted probability of selecting elite school as first choice	[0, .6]	(.6, .8]	(.8, 1]
Score $\geq$ cutoff	0.032	0.023	−0.009
	(0.0240)	(0.0143)	(0.0139)
Observations	4991	14654	18941
Adjusted R^2	0.025	0.022	0.011
Mean of DV 1 pt below cutoff	0.177	0.223	0.302
Bandwidth	5.0	5.0	5.0

Table D.10 (continued)

Panel B1. Linear, BW=10	(1)	(2)	(3)
Dependent variable	Elite school as first choice		
Baseline predicted probability of selecting elite school as first choice	[0, .6]	(.6, .8]	(.8, 1]
Score ≥ cutoff	0.126***	0.069***	0.030***
	(0.0197)	(0.0095)	(0.0058)
Observations	9720	28919	37213
Adjusted R^2	0.039	0.027	0.018
Mean of DV 1 pt below cutoff	0.582	0.763	0.903
Bandwidth	10.0	10.0	10.0

Panel B2. Linear, BW=10	(1)	(2)	(3)
Dependent variable	Number of elite schools chosen		
Baseline predicted probability of selecting elite school as first choice	[0, .6]	(.6, .8]	(.8, 1]
Score ≥ cutoff	0.518***	0.393***	0.221***
	(0.1055)	(0.0678)	(0.0627)
Observations	9720	28919	37213
Adjusted R^2	0.060	0.054	0.066
Mean of DV 1 pt below cutoff	2.408	3.841	5.550
Bandwidth	10.0	10.0	10.0

Panel B3. Linear, BW=10	(1)	(2)	(3)
Dependent variable	Elite assignment		
Baseline predicted probability of selecting elite school as first choice	[0, .6]	(.6, .8]	(.8, 1]
Score ≥ cutoff	0.034*	0.037***	0.009
	(0.0164)	(0.0099)	(0.0096)
Observations	9720	28919	37213
Adjusted R^2	0.034	0.031	0.019
Mean of DV 1 pt below cutoff	0.176	0.223	0.302
Bandwidth	10.0	10.0	10.0

Table D.10 (continued)

Panel C1. Quadratic, BW=10	(1)	(2)	(3)
Dependent variable	Elite school as first choice		
Baseline predicted probability of selecting elite school as first choice	[0, .6]	(.6, .8]	(.8, 1]
Score ≥ cutoff	0.113***	0.066***	0.023**
	(0.0312)	(0.0152)	(0.0093)
Observations	9720	28919	37213
Adjusted R^2	0.039	0.027	0.018
Mean of DV 1 pt below cutoff	0.582	0.763	0.903
Bandwidth	10.0	10.0	10.0

Panel C2. Quadratic, BW=10	(1)	(2)	(3)
Dependent variable	Number of elite schools chosen		
Baseline predicted probability of selecting elite school as first choice	[0, .6]	(.6, .8]	(.8, 1]
Score ≥ cutoff	0.667***	0.320***	0.227**
	(0.1662)	(0.1059)	(0.0995)
Observations	9720	28919	37213
Adjusted R^2	0.060	0.054	0.067
Mean of DV 1 pt below cutoff	2.408	3.841	5.550
Bandwidth	10.0	10.0	10.0

Panel C3. Quadratic, BW=10	(1)	(2)	(3)
Dependent variable	Elite assignment		
Baseline predicted probability of selecting elite school as first choice	[0, .6]	(.6, .8]	(.8, 1]
Score ≥ cutoff	0.026	0.016	−0.006
	(0.0253)	(0.0153)	(0.0149)
Observations	9720	28919	37213
Adjusted R^2	0.034	0.031	0.019
Mean of DV 1 pt below cutoff	0.176	0.223	0.302
Bandwidth	10.0	10.0	10.0

Table D.11: Test for balance of older sibling covariates

Panel A. Linear, BW=5	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Parental education (years)	Male	Hours studied per week	Middle school GPA	Number of siblings	Birth order (1=oldest)	Mean COMIPEMS score of MS peers	Same-sex sibling pair
Score ≥ cutoff	−0.029	0.000	0.004	−0.005	0.010	0.010	0.002	−0.003
	(0.0298)	(0.0040)	(0.0295)	(0.0060)	(0.0116)	(0.0097)	(0.0544)	(0.0041)
Observations	189166	235381	187241	235381	192270	191734	235360	235381
Adjusted R^2	0.105	0.072	0.071	0.182	0.053	0.030	0.294	0.001
Mean of DV 1 pt below cutoff	10.649	0.431	5.080	8.174	2.322	1.673	63.055	0.519
SD of DV 1 pt below cutoff								
Bandwidth	5.0	5.0	5.0	5.0	5.0	5.0	5.0	5.0
p-value for joint significance of Score ≥ cutoff coefficients								0.69

Panel B. Linear, BW=10	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Parental education (years)	Male	Hours studied per week	Middle school GPA	Number of siblings	Birth order (1=oldest)	Mean COMIPEMS score of MS peers	Same-sex sibling pair
Score ≥ cutoff	0.001	−0.004	0.013	0.004	0.007	0.011	0.037	−0.000
	(0.0205)	(0.0028)	(0.0201)	(0.0041)	(0.0080)	(0.0066)	(0.0375)	(0.0028)
Observations	368919	459529	365134	459529	374962	373935	459485	459529
Adjusted R^2	0.104	0.073	0.072	0.184	0.050	0.027	0.295	0.000
Mean of DV 1 pt below cutoff	10.648	0.432	5.080	8.174	2.322	1.674	63.054	0.519
SD of DV 1 pt below cutoff								
Bandwidth	10.0	10.0	10.0	10.0	10.0	10.0	10.0	10.0
p-value for joint significance of Score ≥ cutoff coefficients								0.95

Panel C. Quadratic, BW=10	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Parental education (years)	Male	Hours studied per week	Middle school GPA	Number of siblings	Birth order (1=oldest)	Mean COMIPEMS score of MS peers	Same-sex sibling pair
Score ≥ cutoff	−0.022	−0.000	0.021	0.002	0.001	−0.000	−0.013	−0.001
	(0.0320)	(0.0044)	(0.0317)	(0.0065)	(0.0126)	(0.0105)	(0.0586)	(0.0045)
Observations	368919	459529	365134	459529	374962	373935	459485	459529
Adjusted R^2	0.104	0.073	0.072	0.184	0.050	0.027	0.295	0.000
Mean of DV 1 pt below cutoff	10.648	0.432	5.080	8.174	2.322	1.674	63.054	0.519
SD of DV 1 pt below cutoff								
Bandwidth	10.0	10.0	10.0	10.0	10.0	10.0	10.0	10.0
p-value for joint significance of Score ≥ cutoff coefficients								0.67