American Economic Review, Volume 114, Number 4, pp. 1049-82

Why do households leave school value added on the table? The roles of information and preferences

By Robert Ainsworth, Rajeev Dehejia, Cristian Pop-Eleches, and Miguel Urquiola*

Romanian households could choose schools with 1 s.d. worth of additional value added. Why do households leave value added "on the table"? We study two possibilities: (i) information and (ii) preferences for other school traits. In an experiment, we inform randomly selected households about schools' value added. These households choose schools with up to 0.2 s.d. of additional value added. We then estimate a discrete choice model and show that households have preferences for a variety of school traits. As a result, fully correcting households' beliefs would eliminate at most a quarter of the value added that households leave unexploited.
JEL: I20, D80, C54
Keywords: school choice, value added, information

Friedman (1955) argued that giving households freedom to choose schools would improve their children's learning. This simple idea underlies numerous programs that expand school choice. Yet research yields surprisingly mixed evidence on the effects of school choice. For example, voucher experiments show that choice can impact students' skills in ways that are highly positive (Bettinger et al. 2017), highly negative (Abdulkadiroglu, Pathak and Walters 2018), or modest (Muralidharan and Sundararaman 2015). Considering evidence on choice among selective schools, Beuermann and Jackson (2020) state: "the lack of robust achievement effects of attending schools that parents prefer is something of a puzzle."

We investigate two possible explanations for this puzzle. First, a lack of *information* may prevent households from choosing schools with high value added. Value added is the change in a student's outcomes due to attending a school. This is considerably more difficult to observe than other school attributes, such as the

* Ainsworth: University of Florida (email: robert.ainsworth@ufl.edu); Dehejia: New York University and NBER (email: rajeev@dehejia.net); Pop-Eleches: Columbia University, BREAD, and NBER (email: cp2124@columbia.edu); Urquiola: Columbia University and NBER (email: miguel.urquiola@columbia.edu). For useful comments, we thank Matt Backus, Peter Bergman, Michael Best, Sarah Cohodes, David Figlio, John Friedman, Dong Woo Hahm, Rema Hanna, Douglas Harris, Caroline Hoxby, Kirabo Jackson, Francisco Martorell, Karthik Muralidharan, Suresh Naidu, Christopher Neilson, Daniele Paserman, Jonah Rockoff, Richard Romano, Bennett Smith-Worthington, Doug Staiger, and seminar participants at APPAM, Berlin, Bocconi, Boston, Bristol, Davis, ERMAS, EUI, Florida, NBER Summer Institute, Nottingham, and Princeton. We especially thank Miikka Rokkanen for early contributions to this project. For financial support, we thank Columbia University, the Center for Development Economics and Policy, the SIPA Faculty Grants Program, New York University's Wagner School of Public Service, and the KDI School of Public Policy and Management. The RCT was registered in the American Economic Association Registry (AEARCTR-0004496). The research was approved by IRBs at NYU, Columbia, and Florida (IRB-FY2018-1896, IRB-AAAU0889, IRB202201553).

quality of a school's facilities or the achievement levels of its students. Thus, it is possible that households do wish to attend high-value added schools, but do not know which those are. Second, households may have *preferences* that lead them to prioritize school attributes other than value added. This possibility arises if school quality is multidimensional (Beuermann et. al 2019, Riehl et al. 2019).

Distinguishing between preferences and information is important. If information is the obstacle, then making it available would improve households' choices and, possibly, spur schools to compete on value added. By contrast, if preferences are the constraint, then policy options to boost value added may be more limited. For instance, school choice may cause schools to invest in other, perhaps less desirable, dimensions of quality (Rothstein 2006).

We explore the distinction between preferences and information by studying high school admissions in Romania. To conduct our analysis, we obtained administrative data on fifteen admissions cohorts. We also implemented surveys and ran an informational experiment.

Two features of the Romanian school system make it an advantageous setting. First, high school is bookended by high-stakes standardized exams. Before entering high school, students take a national admissions test, the "transition exam." Before graduating, they take a national exit test, the "baccalaureate exam." These tests allow us to calculate schools' academic value added. The particular outcome we focus on—performance on the baccalaureate exam—is of central importance to Romanian students.

The second advantageous feature is the student assignment mechanism: a serial dictatorship. Each student receives a score before applying to high school. An algorithm then considers applicants one at a time according to their scores, assigning each to his/her most-preferred school that has not yet reached capacity. A household can rank an unlimited number of options; thus, its dominant strategy is to rank truthfully according to its preferences (Chade and Smith 2006). The serial dictatorship lets us: (i) observe the high schools that a student could attend and (ii) be confident that the one the student enrolls in is her most preferred. Further, the algorithm generates school-specific admissions cutoffs. These provide regression discontinuity (RD) estimates of the effect of access to each school. Following Angrist et al. (2017), we use the RDs to validate our value added estimates, which we find closely match causal effects. In short, administrative data allow us to calculate value added and to see the results of household decision-making.

To probe the mechanics of this decision-making, we visited middle schools and collected a baseline survey. This survey occurred at school-sponsored information sessions that are held to help households apply to high school. In the survey, we interviewed parents to obtain the school preference rankings that they intended to submit. We also asked them to evaluate the high schools in their town along dimensions including location, peer quality, curricular focus, and different types of value added.

We ran our experiment at the end of these sessions. At randomly selected treatment schools, we distributed a ranking of the town's high schools based on academic value added. After the assignment process was complete, we obtained students' official school assignments. We also conducted an endline survey, interviewing parents by phone to gather their submitted school preference rankings and to again elicit their beliefs about schools' value added.

This setup yields four findings, which we use to organize the exposition and frame the paper's contribution. They are as follows.

1. Households leave value added on the table

Schools with higher academic value added face higher demand. The correlation between a school's value added and the selectivity of its admissions cutoff is 0.56.[1] In addition, households choose options that are above average by value added in their feasible choice sets. Nonetheless, they leave considerable value added unexploited. Both low- and high-achieving students could gain, on average, about 1 s.d. worth of additional value added—or a 12 percentage point increase in the probability of passing the baccalaureate exam.[2] By contrast, households come closer to maximizing selectivity. For this school trait, they leave only about 0.3 s.d. unexploited.

These results relate to work asking whether households favor productive schools (Beuermann et al. 2019, Abdulkadiroglu et al. 2020). Our contribution is to exploit a setting in which researchers can: (i) measure all schools' value added, and (ii) precisely observe the set of schools among which each household chooses, as well as the one it most prefers.

2. Households have limited knowledge of value added

When asked to score schools on academic value added, households' scores are off by an average of 1.1 within-town quintiles and explain only 17% of the variation. In contrast, households have more awareness of selectivity. Their scores for this school trait have a mean absolute error of 0.9 within-town quintiles and explain 33% of the variation. Finally, households with high-achieving children have more accurate beliefs than those with low-achieving children.

Our contribution is to provide, to our knowledge, the first comparison between researchers' and households' perceptions of school value added within entire markets.

3. Households respond to information on value added (with heterogeneity)

Our treatment improved the accuracy of households' beliefs and caused them to assign higher preference ranks to high-value added schools. Thus, on average, it induced students to attend schools with 0.05 s.d. worth of additional value added. That said, the treatment had larger effects on beliefs and preference ranks for households with low-achieving students. In addition, it did not alter beliefs or ranks for the two options that a household ranked the highest in the baseline. As a

[1] This is a descriptive result—it could arise because households choose schools based on value added, but it could also arise if households seek a correlate of value added, or if there are positive peer effects.

[2] We classify a student as high- (low-) achieving if her transition score is in the top (bottom) half of her application cohort distribution.

result, its effects on students' school assignments were heterogeneous. Notably, for low-achieving students who were rejected by their two top choices, the treatment resulted in enrollment at schools with 0.2 s.d. worth of additional value added. This implies a 2.5 percentage point (9.8%) increase in their probability of passing the baccalaureate exam. By contrast, for all other students, the treatment had no impact on school assignments.

These results address whether information on school quality affects households' choices. Previous work finds positive effects from information related to schools' *absolute* achievement (Hastings and Weinstein 2008, Andrabi, Das and Khwaja 2017, Ajayi, Friedman and Lucas 2017, Corcoran et al. 2018, Allende, Gallego and Neilson 2019) but limited impacts from information related to value added (Imberman and Lovenheim 2016, Mizala and Urquiola 2013). Our contribution is to conduct, to our knowledge, the first experimental distribution of information on school value added.

4. Preferences for other school traits limit households' demand for value added

We use households' school preference rankings and their elicited beliefs about school traits to study their preferences for these traits. We first estimate preferences using a discrete choice model. We then disentangle the roles that preferences and information play in causing households to leave value added on the table. Specifically, we compare predicted school assignments under accurate beliefs about academic value added with those under baseline beliefs. In different specifications, we predict that correcting households' beliefs would spur low- (high-) achieving students to attend schools with 0.13-0.20 (0.10-0.23) s.d. worth of additional value added. This is 17-25% (11-24%) of the value added that the households would leave unexploited under baseline beliefs. Households would not "max out" on value added under accurate beliefs due mostly to their preferences for curricular focus and peer quality.

These results relate to papers on households' preferences for school traits (Hastings, Kane and Staiger 2005, Burgess et al. 2015, Beuermann et al. 2019, Abdulkadiroglu et al. 2020). Our contribution is to calculate preferences using households' beliefs about these traits, rather than values measured by researchers. More broadly, our paper relates to work assessing the roles of preferences and frictions in driving choices (Bergman et al. 2019, Hastings, Neilson and Zimmerman 2018, Bergman, Chan and Kapor 2020, Bau 2022, Carneiro, Das and Reis 2022).

In the rest of the paper, Section I provides some necessary preliminaries, and Sections II-V report on findings 1-4, respectively. Section VI concludes.

I. Preliminaries

We begin by describing the setting, the administrative data, our value added measures, our surveys, and our experiment.

A. *Institutional setting*

A few features of the Romanian setting are especially relevant to our analysis. First, in Romania, high schools cover grades 9-12 and are divided into *tracks*. These are self-contained units within schools that differ in their "curricular focus," or the set of subjects that they emphasize. Curricular focuses fall into three broad categories: a) humanities, b) math or science, and c) "technical studies" with applied themes such as business or agriculture.

Second, students are assigned to tracks via a centralized process known as a serial dictatorship. This process weights students' track preferences according to their academic performance in middle school (grades 5-8). Specifically, in 8$^{\text{th}}$ grade, each student takes a national high school entrance test. The student's score on this exam is combined with her middle school GPA to generate an admissions score, called the transition score. After finding out its child's transition score, a household submits a ranked list—or *preference ranking*—of its preferred tracks. The government then examines the preference rankings in the order of students' transition scores. It first takes the student with the highest score and assigns her to her most-preferred track. It then proceeds down the score distribution, assigning each student to her most-preferred track that is not yet at capacity.

Third, the track assignment process is incentive-compatible. Households' preference rankings can be of virtually unlimited length (up to 287 choices). As a result, the optimal strategy is to submit a list that truthfully reveals one's preferences.[3]

Fourth, a household's choice set is best thought of as the tracks in its town. Technically, households may rank any track in the country. However, it is uncommon for households to move for educational purposes. In addition, Romanian towns tend to be geographically distinct; thus, few students commute from one town to another. Further, Romanian towns are compact, and high schools are usually located in the town-center; thus, within-town commutes are rarely difficult. As a result of these features, we assume that households consider all the tracks in their town and do not consider options in other towns.[4]

Fifth, at the end of high school, students may elect to take a national standardized test known as the baccalaureate exam. The baccalaureate exam has high stakes: there are benefits both to passing it and to achieving a high score. Students who pass receive a baccalaureate diploma, which is necessary for admission to university—at less selective schools, it is the only requirement. A high score helps students access scholarships and prestigious universities (Borcan, Lin-

[3]Recent work notes that households may reasonably choose not to rank a track if they are certain it is out of reach for their child (Fack, Grenet and He 2019, Artemov, Che and He 2020). Below, we show that our findings are robust to using empirical strategies that account for such "skipping."

[4]In our baseline survey, over 93% of households said that they intended to apply only to tracks within their town. For these households, we find that within-town distances hardly affect track choice (Section V). The one setting where distance may matter is Bucharest, which is by far the largest city. Following the Ministry of Education, we divide Bucharest into six sub-town units. Our results are robust to excluding these units.

dahl and Mitrut 2017). Even for students who do not pursue higher education, performing well on the baccalaureate exam can be a strong labor market signal.

B. Administrative data

We have administrative data on the universe of students admitted to Romanian high schools (Ministry of Education 2014, 2019). We use the data to calculate academic value added for each high school track and to examine whether households choose tracks with high value added. The data cover the 2004-2017 and 2019 cohorts. For all cohorts, they provide information on students' demographics, middle school, middle school GPA, scores on the transition exam, and assigned high school track. For 2004-2014, they also include performance on the baccalaureate exam. On average, a cohort includes about 144,000 students who live in about 400 towns and choose among about 3,800 tracks.[5]

Table 1—Summary statistics for the administrative data

	Mean	**Std. dev.**	**Students**
High school track:			
Number of students	61.8	47.0	2,162,736
Minimum transition score (MTS)	6.94	1.59	2,162,736
Student characteristics:			
Female	0.527	0.499	2,162,736
Transition score	7.70	1.35	2,162,736
Middle school GPA	8.65	0.97	2,162,736
Transition exam score	7.05	1.69	2,162,736
Baccalaureate performance:			
Took the exam	0.686	0.464	1,710,030
Passed the exam	0.533	0.499	1,710,030
Perfect score	0.001	0.025	1,710,030

Notes: The table provides summary statistics for the administrative data. Variables under "High school track" are characteristics of a student's track. Variables under "Baccalaureate performance" are available only for the 2004-2014 cohorts.

Table 1 summarizes these data. One covariate in the table merits a special comment. A track's *minimum transition score* (MTS) is the score of the last student admitted. It is the track's admissions cutoff: students with higher scores are eligible to attend the track, while those with lower scores are not. The MTS is a direct measure of a track's selectivity; in addition, it is a proxy for the demand that a track faces—tracks that are more popular reach capacity earlier in the allocation process and thus have higher cutoffs. When the government announces

[5]Online Appendix Table A1 presents the sample size by year. We impose three restrictions on the sample. First, we exclude 2018 due to a reporting issue. Second, we drop a small number of students who participate in vocational programs that do not offer a path to a baccalaureate diploma. Third, given that we are interested in track choice, we drop students who live in very small towns that offer only a single track.

the set of tracks that will accept students, it provides the tracks' MTS from the previous admissions round. Anecdotal evidence suggests that households pay attention to this information when determining their track preference rankings.

C. Value added

We calculate multiple measures of track academic value added; all relate to a track's effect on a student's performance on the baccalaureate exam. As mentioned, students choose whether to take this exam.[6] This means that value added calculated on students' scores could be biased by sample selection. Our main outcome, therefore, is an indicator for whether a student *passes* the exam; this variable is set to zero if the student fails the exam or does not take it. Unless explicitly noted, when we refer to value added we mean value added on passing the baccalaureate exam.

We also consider two outcomes that directly incorporate information on students' scores. First, the *percentile rank* of a student's baccalaureate performance is the percent of students in the admissions cohort who perform worse than the student, with all students who do not pass being assigned a value of 0.[7] Second, the *imputed exam score* is a version of the score that deals with missing scores using imputations: it fills in a value equal to the 33^{rd} percentile among students who take but fail the exam.

We consider these three measures to be complementary. Value added on passing the exam is free of sample selection. Value added on the other outcomes allows more precise results for selective tracks in which large shares of students pass.

Our main value added measures vary by track and year; for robustness, we also calculate measures that vary by student characteristics. These measures allow a track-year to have different effects depending on whether a student is male or female or whether the student is better at math or language. It turns out that all our measures are highly correlated. For example, our main measure (a track-year effect on passing) has a correlation of over 0.9 with each of the alternative measures (online Appendix Table 2).

We detail our methodology for calculating value added in online Appendices A-C. Three facts about it are worth mentioning. First, we rely on a traditional selection-on-observables model (Rothstein 2010, Angrist et al. 2017). Second, we validate our measures by comparing them with RD causal effects generated by track admissions cutoffs—this involves adapting Angrist et al. (2017) to an RD setting. We show that all of our measures closely match the causal effects; in

[6]In our cohorts, 69% of high school students attempted the exam, 53% passed it, and 0.1% achieved a perfect score (Table 1). Online Appendix Figure A1 shows how these values vary with a student's incoming achievement.

[7]This outcome is motivated by the benefit structure of the baccalaureate exam. First, the benefits from a given score depend to a large extent on the fraction of students who perform worse. As such, households may be interested in the percentile rank of performance associated with the score. Next, there are no benefits to taking the exam but failing it. That is, students who fail the exam gain the same result as those who do not attempt it. Both groups perform better than no other students, and thus we assign them a percentile rank of 0.

addition, the measures that do not allow for heterogeneity by student characteristics perform just as well as those that do. Third, we deal with measurement error by calculating Empirical Bayes (EB) posterior means.

One complication in our setting is that we cannot observe baccalaureate outcomes for post-2014 admissions cohorts; as a result, we cannot directly calculate value added for these cohorts. We handle this by forecasting the missing years using machine learning.[8]

Our analysis uses value added variables which we label V_{jt} (for track-year effects) or V_{jgt} (for effects that vary by student). In this notation, j indexes tracks, t indexes years, and g indexes student types. For the years in which we can estimate value added (2004-2014), these variables equal the Empirical Bayes posteriors; for the years in which we cannot (2015-2017, 2019), they equal the machine learning forecasts.

In terms of magnitudes, we find that Romanian tracks differ significantly in value added. For instance, in 2019, a one standard deviation increase in true value added was equivalent to a 12 percentage point increase in the probability of passing the baccalaureate exam.

D. Baseline survey

To gain insight into households' beliefs and preferences, we conducted a baseline survey. We interviewed parents of 8$^{\text{th}}$ graders to collect: (i) their beliefs about the attributes of tracks in their towns, and (ii) their intended track preference rankings.[9] We did this at information sessions held by middle schools to inform parents about the high school application process. These occur about a month before households submit their final track preference rankings.

To select our sample, we had to choose towns and middle schools within towns. We chose towns using two criteria. First, we considered only moderately-sized towns, defined as those that had between 7 and 28 tracks in 2018. Second, among these towns, we chose those in which value added was easiest to forecast (online Appendix Table A3 provides summary statistics for these towns). To choose middle schools, we randomly selected among those that had at least 15 students and in which it was logistically feasible for our surveyors to visit the information sessions. We wanted to minimize spillovers and general equilibrium effects from our experiment. As a result, we visited only a fraction of middle schools in each town—an average of 11% and never more than a third. Our sample covered 194 middle schools in 48 towns. In 2019—the year in which we conducted the survey—the towns had an average of 13 tracks and 412 students. We interviewed the households of 3,898 students, with an average of 81 students per town.

[8] We obtain the forecasts using a local linear forest (Athey et al. 2019). Our model uses a track's past value added and multiple track and student traits. We assess the model by making out-of-sample predictions in years in which we observe value added. The model predicts almost 80% of the variation in tracks' true value added.

[9] Online Appendix D analyzes these rankings. It shows that households do not omit "out of reach" tracks. It also shows that households rank tracks from multiple curricular focuses.

We asked parents to score the tracks in their town—on a scale of 1 to 5—on a variety of dimensions.[10] Table 2 lists those we covered. We first asked about two school attributes that many studies find households value: location and peer quality. We also asked about our definition of value added ("this track will help my child pass the baccalaureate exam"), as well as alternative types of value added related to college and labor market success. Finally, we asked parents to score tracks on teacher quality, on whether the curricular focus is a good fit for their child, and on whether the track is attractive because it is also used by their child's siblings or friends.

Table 2—Track characteristics covered in the baseline survey

Characteristic	Definition
Location	This track has a convenient physical location (close to my home or preferred transport)
Peer quality	This track attracts academically gifted students
VA: pass the bacc.	This track will help my child pass the baccalaureate exam
VA: college	This track will help my child go to the college that I would like for him or her
VA: wages	This track will raise my child's earnings at age 30
Teacher quality	This track has good teachers
Curricular focus	My child will enjoy this track's curricular focus
Siblings & friends	My child's siblings and friends also attend this track (or this track's school)

Notes: The table displays the definitions of the track characteristics covered in the baseline survey.

A number of checks suggest that these scores are credible. First, the means and standard deviations are similar for the various quality dimensions (online Appendix Table A5). Second, the scores have a reasonable across-dimension correlation matrix (online Appendix Table A6). For instance, the largest correlations are among the three value added dimensions; the lowest are those that include scores for a track's location or for whether the child's siblings/friends attend the track. Third, the value added scores have an intuitive relationship with the other scores: in a multivariate regression, they are explained by scores for teacher quality, curricular focus, and peer quality, but not by scores for location or siblings and friends (online Appendix E).

Online Appendix Table A7 describes other variables in the survey, revealing a few notable facts. First, households did not rank or score all the tracks in their towns. On average, they assigned ranks to 42% and quality scores to 35%.[11]

[10]In Romania, scales of 1-5 are often interpreted in terms of quintiles. However, we were careful not to ask households to group tracks into equal-sized bins. Instead, we requested that they assign each track whatever score they thought was appropriate. Thus, the scores roughly correspond to a household's expectation about a track's quintile, rounded to the nearest integer. This way of scoring tracks can incorporate information about a household's uncertainty. For instance, if a household has imprecise beliefs and is unable to differentiate among tracks, it can assign each a score of 3. By contrast, a confident household can assign a distribution of scores that approximates quintiles. Online Appendix Table A4 summarizes the frequency with which households assign scores of each value. The frequencies are all close to 0.2—although households tend to assign more high than low scores.

[11]Despite this, we find that households considered tracks from across the selectivity distribution (online Appendix D). That is, it is not the case that households with low-achieving children omitted all selective tracks; nor is it the case that households with high-achieving children omitted all non-selective ones.

Second, at the time of the baseline survey, households differed in the degree to which they had settled on their track choices: 39% were "very certain" of their preference rankings, while 46% were "somewhat certain" and 15% were "uncertain". Third, households with low-achieving children tended to be less certain than those with high-achieving children (online Appendix F). Fourth, students in the baseline survey had similar characteristics as those in the administrative data (Table 1).

E. *Experiment and endline survey*

We ran an experiment and an endline survey to explore the impact of providing households with information on value added.

The experiment took place during the middle school information sessions where we conducted the baseline survey. In advance of the sessions, we split the middle schools into treatment and control groups using a clustered randomization process.[12] At the end of the baseline survey, we distributed an informational flyer. In the control middle schools, the flyer provided links to government websites, including one listing the prior-year minimum transition score for each track. In the treatment middle schools, the flyer also explained the concept of value added and included a ranking of the tracks in the town by our value added forecasts. Respondents were allowed to keep the flyers.[13]

After the high school allocation, we obtained students' track assignments from the Ministry of Education. In addition, we phoned households to conduct the endline (or "follow-up") survey. In this survey, we collected the final track preference rankings that households submitted and asked them to again score tracks on a scale of 1 to 5 in terms of value added. The data on track assignments lets us see how the information affected the tracks that students attend. The follow-up survey lets us probe the mechanics by which the information influenced choices.[14]

Online Appendix Table A8 presents summary statistics and balance tests for the experiment. The first row displays the share of students in the experimental sample who were assigned to a high school track. It shows that 85% of students

[12]We matched pairs of middle schools within towns based on school characteristics. We then randomized within these matched pairs. Online Appendix G provides details.

[13]Example flyers are in online Appendix Figures A2-A3. For all households, the flyer included Figure A2; for treated households, it also included Figure A3. Our intervention focused only on value added with respect to passing the baccalaureate exam. We stated that our rankings reveal "which tracks most effectively improve students' chances of passing the baccalaureate exam relative to their 9$^{\text{th}}$ grade starting points." It is possible that this type of value added is not of interest to students with very high or very low chances of passing. However, we believe this is unlikely. Online Appendix E shows that households' beliefs about a track's value added on passing the exam are highly correlated with their beliefs about the track's value added on college quality or on wages. Thus, households may have interpreted our information as a clear signal of value added on these other outcomes.

[14]In the time between the creation of the matched pairs and the baseline survey, some middle schools withdrew their permission for our study. For every school where this occurred, we still conducted the baseline survey in the other school in the pair. However, we removed the pair from the experimental sample. Also, we dropped 226 students who reported at baseline that they did not intend to apply to tracks in their town. Thus, while the baseline sample includes 3,898 students in 194 middle schools in 48 towns, the experimental sample includes only 3,186 students in 170 middle schools in 45 towns.

were assigned, with an insignificant difference of 2.5 percentage points between treatment and control groups.[15] The remaining rows exclude students who were not assigned. Balance tests for this sample suggest that the randomization succeeded. The differences between treatment and control groups are small relative to the variables' standard deviations, and none are statistically significant. Moreover, a test of joint statistical significance returns a p-value of 0.722.

Comparing online Appendix Tables A7 and A8 shows that students in the experiment are mostly representative of those in the baseline survey. However, they ranked and scored a larger share of tracks. Also, they were slightly more likely to be certain of their preference rankings, and they had slightly higher transition scores.[16]

II. Households leave value added on the table

The first question we study is whether households choose tracks with high academic value added. Previous work considers this question in a variety of settings. We consider it further for four reasons. First, it reveals whether Romanian households gain academic benefits from their choices. Second, it shows if there is scope to increase their benefits by providing information. Third, it clarifies the representativeness of our setting: comparing our results with the literature provides a sense of whether our findings in later sections are likely to be externally valid. Fourth, in contrast to other studies, we can (i) estimate value added for all schooling options and (ii) observe each student's feasible choice set—thus, we can quantify exactly how much value added households leave unexploited.

To address our first question, we use the administrative data and we employ two approaches. First, we examine whether a track's value added is correlated with the demand it faces, as measured by the selectivity of its admissions cutoff. This approach is similar to the prior literature and lets us benchmark our results. Second, we characterize households' choices in relation to their available options. This analysis exploits our knowledge of households' choice sets and has not been done before. We note that in both cases the analysis is descriptive. It illuminates whether the tracks that households choose happen to have high value added—not whether households make choices based on value added.

Our first approach is to inspect the relationship between a track's value added and its selectivity. This relationship will be positive if tracks with high value added are popular and reach capacity early in the assignment process.

Figure 1 shows that the relationship is strongly positive for less- and moderately selective tracks, but slightly negative for highly selective tracks. The figure plots

[15] Students were matched with data on track assignments by name and middle school. Students do not appear in these data if they do not submit a track preference ranking. A small number of unassigned students participate in a secondary allocation that occurs at the end of the summer; these students get assigned to tracks that did not reach capacity in the main allocation. The remaining students either drop out or attend vocational schools.

[16] Online Appendix Table A12 compares all the samples used in the paper.

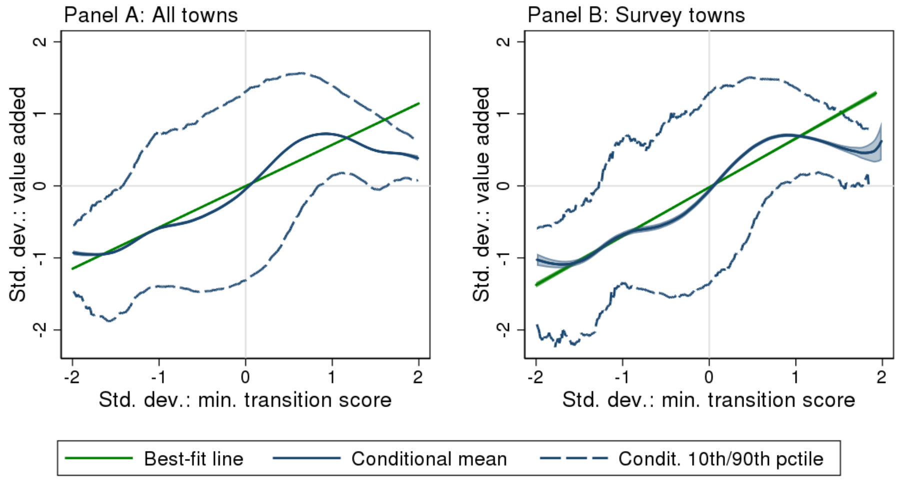

Figure 1. The relationship between value added, V_{jt}, and selectivity, MTS_{jt}

Notes: The figure summarizes the relationship between value added and selectivity. The best-fit line is from a linear regression of standardized values of value added, V_{jt}, on standardized values of minimum transition score, MTS_{jt}. "Conditional mean" plots predictions from a local linear regression, and "conditional 10^{th} and 90^{th} percentiles" are from local quantile regressions. The value added measure is a track-year effect on the probability of passing the baccalaureate exam. Variables are standardized by year.

the conditional mean and the conditional 10^{th} and 90^{th} percentiles of standardized value added, V_{jt}, against standardized values of minimum transition score, MTS_{jt}.[17] It also includes a best-fit line from a linear regression. The described pattern holds for all towns (Panel A), in survey towns (Panel B), and within curricular focus (online Appendix Figure A4).[18]

Table 3 quantifies the results in Figure 1: it presents coefficients from regressions of standardized V_{jt} on standardized MTS_{jt}. The values in the rows labeled "All tracks" match the slopes of the best-fit lines in Figure 1. The remaining rows capture the non-linearity in the figure—they are coefficients from regressions that split the sample by tercile of selectivity. For all towns (Panel A), the overall correlation between value added and selectivity is 0.56. But for the most-selective third of tracks, a one s.d. increase in selectivity corresponds with a 0.24 s.d. decline in value added. For survey towns (Panel B), the values are 0.66 and -0.16.

[17]Here, V_{jt} is our main value added measure: a track-year effect on passing the baccalaureate exam.

[18]The pattern also holds regardless of the value added measure used (online Appendix Figure A5). This mitigates a potential concern regarding Figure 1. In particular, the negative relationship between value added and selectivity for highly selective tracks could be due to a mechanical constraint on value added for these tracks. If students in highly selective tracks are certain to pass the baccalaureate exam regardless of the track they attend, then there will be a cap on these tracks' measured value added. This cap is less likely to be binding for value added on the percentile rank of a student's exam performance or on a student's exam score—for instance, only 0.1% of students achieve a perfect score (Table 1). The non-linearity persists using these alternative measures.

Table 3—Regressions of standardized value added estimates on standardized selectivity

Sample	Coefficient	Std. error	Town-years	Track-years	Students
Panel A: All towns					
All tracks	0.562	0.005	5,969	57,521	2,162,736
By tercile of selectivity:					
Least selective	0.391	0.017	5,710	24,934	723,446
Moderately selective	1.07	0.028	4,325	17,207	723,023
Most selective	-0.243	0.024	2,420	15,380	716,267
Panel B: Survey towns					
All tracks	0.662	0.010	720	11,253	424,508
By tercile of selectivity:					
Least selective	0.408	0.042	717	4,319	135,007
Moderately selective	1.21	0.054	718	3,898	162,887
Most selective	-0.162	0.041	676	3,036	126,614

Notes: The table quantifies the results in Figure 1. It presents coefficients from regressions of standardized value added, V_{jt}, on standardized minimum transition score, MTS_{jt}. The coefficients in the rows labeled "All tracks" match the slopes of the best-fit lines in Figure 1, and can be interpreted as correlation coefficients. "Tercile of selectivity" indicates whether the track is in the lowest, middle, or highest third of MTS_{jt} by year. Regressions are weighted by student; standard errors are clustered by town-year.

The results in Figure 1 and Table 3 are similar to prior work. For instance, in New York City high schools, the overall correlation between value added and peer quality is 0.59 (Abdulkadiroglu et al. 2020). In American higher education, the correlation between a college's selectivity and its earnings value added is 0.63 (Chetty et al. 2020). Also, in various locations, the most selective high schools do not boost achievement relative to students' fallback options (Abdulkadiroglu, Angrist and Pathak 2014, Dobbie and Fryer 2014, Abdulkadroglu et al. 2017). Thus, Romania appears to be representative of other settings.[19]

Our second strategy is to characterize households' choices in relation to their available options. To do this, we exploit our knowledge of a household's feasible choice set—the tracks in the town that the student is eligible to attend, given her transition score and the admissions cutoffs. We conduct the analysis in two ways. First, we compare the value added of the track a household chooses with the value added of its other options. Second, we compute the amount by which a household could increase the value added it receives by switching to its highest-value-added option. We present a parallel analysis for selectivity, asking whether households favor tracks with high-achieving peers.

To elaborate, for each household we calculate two quantities. First, the *per-*

[19]Online Appendix I provides additional replication of existing work. It studies choice behavior using a discrete choice model similar to that in Abdulkadiroglu et al. (2020) and Beuermann et al. (2019). Broadly speaking, we replicate previous findings. As in Abdulkadiroglu et al. (2020), we find that over the full sample, value added does not explain households' utility for tracks after conditioning on selectivity. As in Beuermann et al. (2019), we find that it does—to an extent—for households with high-achieving children.

centile rank of the student's track among feasible tracks is the rank of the student's track (by either value added, V_{jt}, or selectivity, MTS_{jt}) divided by the number of tracks that are available to the student.[20] Second, the *potential increase among feasible tracks* is the difference between the maximum value (of value added or selectivity) within the feasible set and the value for the student's track. It captures how much of an improvement a household could obtain by switching.

Table 4—Summary statistics on households' track choices

	All towns			**Survey towns**		
	All students	Low-achieving	High-achieving	All students	Low-achieving	High-achieving
Panel A: Percent of students with only one feasible track	2.4	4.8	0.0	1.3	2.6	0.0
Panel B: Mean percentile rank of student's track among feasible tracks						
Value added, V_{jt}	67.1	61.0	72.9	67.2	59.9	74.3
Selectivity, MTS_{jt}	81.0	74.9	86.9	79.7	74.6	84.8
Panel C: Mean potential increase among feasible tracks (std. dev.)						
Value added, V_{jt}	1.01	1.01	1.02	0.91	0.93	0.88
Selectivity, MTS_{jt}	0.32	0.34	0.30	0.34	0.34	0.35
Number of students	2,162,736	1,081,075	1,081,661	424,508	211,917	212,591

Notes: The table presents summary statistics on households' track choices. Panel A displays the percent of students who are eligible for only one track. Panels B and C are calculated for students with multiple feasible options; they display means for the "percentile rank of the student's track among feasible tracks" and the "potential increase among feasible tracks." Variables are standardized by year. A student is defined as low- (high-) achieving if his/her transition score is in the bottom (top) half of the within-year distribution.

Table 4 shows that households choose above-average tracks by value added, but leave substantial value added unexploited. Panel A lists the percent of students who have only one track in their feasible set and hence no choice (2% in the full sample). The remaining panels concern the students with choice. For all towns, Panel B reveals that, on average, students attend tracks at the 67$^{\text{th}}$ percentile of value added in their feasible sets. Panel C shows that if students switched to their value added-maximizing options, they would gain an average of 1 s.d. of value added. In 2019, this was equal to a 12 percentage point increase in the probability of passing the baccalaureate exam. The results for survey towns are broadly similar.

Table 4 also shows that households come much closer to "maxing out" on selectivity. Over the full sample, students on average attend tracks with selectivity at the 81$^{\text{st}}$ percentile among their feasible tracks. The average potential increase in selectivity is only 0.32 s.d.

In addition, Table 4 reveals that choice patterns vary little by students' academic achievement. In particular, the results are mostly similar for students

[20] A value of 100 indicates that the household chooses the best option (by value added or selectivity).

with transition scores in the bottom half ("low-achieving") and top half ("high-achieving") of the within-year distribution.

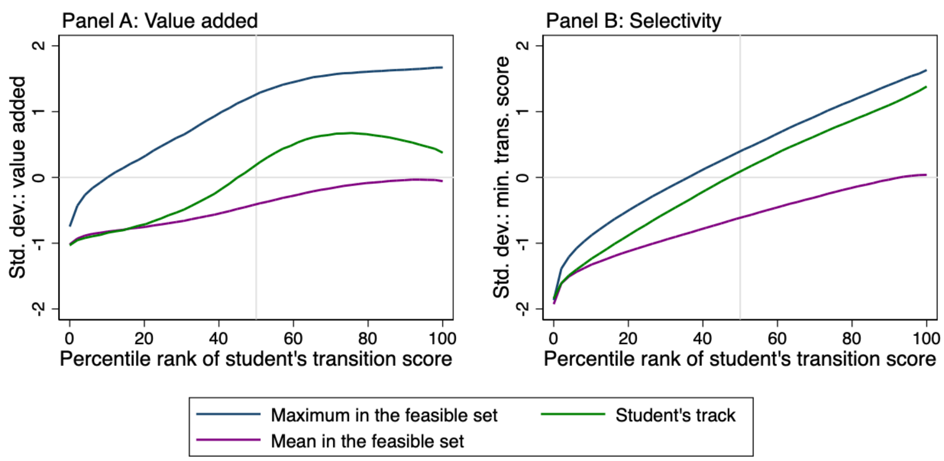

FIGURE 2. CHOICE PATTERNS BY TRANSITION SCORE

Notes: The figure shows how choice patterns vary with a student's transition score. It plots the relationship between the percentile rank of the student's transition score and three variables. The blue line is the maximum value of value added, V_{jt}, or selectivity, MTS_{jt} in the student's feasible set. The purple line is the mean value in the set, and the green line is the value in the track the student attends. The lines are calculated using local linear regressions. The difference between the blue and green lines is the mean potential increase for the given percentile rank. The sample includes students in all towns. See Table 4 for additional details.

The limited heterogeneity is illustrated graphically in Figure 2. It shows how the potential increase in value added or selectivity varies with the percentile rank of a student's transition score. The figure plots the relationship between the student's percentile rank and three variables: (i) the maximum value in the student's feasible set (in standard deviations of value added or selectivity), (ii) the mean value in the set, and (iii) the value for the track the student attends. The difference between the lines for the maximum and for the value of the student's track is equal to the mean potential increase for students with a given transition score. For both value added and selectivity, the potential increases are relatively constant across the transition score distribution. However, they are smaller for the lowest-achieving students, who have limited choice. In addition, for value added, the potential increase is larger for the highest-achieving students.[21]

Finally, we find that households leave significant value added unexploited even within curricular focus. In other words, it does not seem that households sacrifice value added because they willingly exchange it for this other track characteristic.[22]

[21]Figure 2 would be similar using alternative value added measures (online Appendix Figure A6).
[22]Online Appendix Table A14 replicates Table 4 while restricting a household's choice set to the subset

III. Households have limited knowledge of value added

The previous section showed that households leave value added on the table. We now investigate whether this may be because households have inaccurate beliefs about value added. To do so, we make use of the baseline survey. We compare households' elicited beliefs from this survey with the values of track attributes that we observe as researchers (the "measured values"). To our knowledge, this is the first such comparison in the literature.

In the baseline survey, we elicited beliefs by asking households to score tracks on a variety of dimensions on a scale of 1 to 5 (Table 2). We compare these scores with our measured values along two dimensions. First, we compare households' scores for a track's value added on passing the baccalaureate exam with our forecast for this characteristic, V_{jt}. As a benchmark, we also compare households' scores for a track's peer quality with the track's prior-year selectivity, MTS_{jt-1}. The concept of selectivity is well-understood in Romania; in addition, households can view each track's prior-year selectivity on the official admissions website. Thus, this benchmark reflects scores under a scenario of easy access to information.[23]

We characterize households' scores in three ways. First, we quantify the *accuracy* of the scores: we calculate the mean absolute difference between a household's score and the within-town quintile of a track's measured value.[24] This quantity reveals the average amount by which households' scores are incorrect. Second, we quantify the *bias* of the scores: we calculate the mean difference between the scores and the tracks' quintiles. This value shows whether the scores tend to be too high or too low. Finally, we regress the quintiles on the scores. In the regression, the slope coefficient reveals how differences in scores map to differences in quintiles. The R-squared measures how much of the variation in quintiles can be explained by the scores. If households' scores were fully accurate, the slope coefficient and R-squared would both be 1.[25]

The results are in Tables 5 and 6. Table 5 summarizes accuracy and bias for two sets of tracks: (i) all those that a household scored and (ii) the two that the household reported as being most preferred. We include this latter set because

of feasible tracks whose curricula fall into the same focus as that of its child's track. It shows that the average student attends a track with value added (selectivity) at the 64$^{\text{th}}$ (80$^{\text{th}}$) percentile among this restricted choice set. On average, these students could gain increases in value added (selectivity) of 0.55 (0.26) standard deviations.

[23] Arguably, households' peer quality scores should reflect a track's *current-year* rather than prior-year selectivity. A rational household might combine the data on prior-year selectivity with its knowledge of recent changes in tracks' traits. In our main analysis, we refrain from using current-year selectivity because this variable may be impacted by our experiment. Nonetheless, we find that results would be similar if we were to use it.

[24] We compare scores with within-town quintiles because the scores are on a scale of 1 to 5. As discussed, a quality score can be roughly interpreted as a household's rounded expectation about a track's quintile. It is possible that households assigned scores based on the national—rather than within-town—distribution of tracks. However, we believe this is not the case. In results not shown, we replicated the analysis using national quintiles. We find that doing so causes the scores to have less predictive power.

[25] These quantities are not mechanically related. To see this, suppose a household assigns all tracks but one a score of 3. Suppose it then assigns the correct score to the remaining track. In this example, the slope coefficient would be 1 while the R-squared would be small.

Table 5—The accuracy and bias of households' quality scores

	All tracks			Two most-preferred tracks		
	All students	Low-achieving	High-achieving	All students	Low-achieving	High-achieving
Panel A: Accuracy (mean abs. dif.)						
Value added: s_{ij}^{V} v. quint(V_{jt})	1.13	1.19	1.09	0.99	1.06	0.95
Selectivity: s_{ij}^{PQ} v. quint(MTS_{jt-1})	0.90	1.01	0.84	0.78	1.06	0.62
Panel B: Bias (mean dif.)						
Value added: s_{ij}^{V} v. quint(V_{jt})	0.35	0.45	0.29	0.63	0.64	0.63
Selectivity: s_{ij}^{PQ} v. quint(MTS_{jt-1})	0.17	0.36	0.06	0.35	0.65	0.17
Students	2,370	883	1,487	2,283	837	1,446
Student-tracks	17,460	6,433	11,027	3,900	1,420	2,480

Notes: The table summarizes the accuracy and bias of households' scores. Panel A describes accuracy: it displays the mean absolute difference between a household's score for value added on passing the baccalaureate exam, s_{ij}^{V} (peer quality, s_{ij}^{PQ}) and the within-town quintile of measured value added, quint(V_{jt}) (prior-year selectivity, quint(MTS_{jt-1})). Panel B describes bias: it exhibits mean differences between the quantities. "Two most-preferred tracks" are the two that the household ranked highest at baseline. The sample drops: (i) student-track observations where the respondent did not score the track on both value added and peer quality and (ii) 152 students with missing transition scores.

households may know more about their favored options. Table 6 provides the regression results using all scored tracks.

Table 6—Regressing track attributes on households' quality scores

	All students		Low-achieving		High-achieving	
	quint(V_{jt})	quint(MTS_{jt-1})	quint(V_{jt})	quint(MTS_{jt-1})	quint(V_{jt})	quint(MTS_{jt-1})
Score: VA-pass, s_{ij}^{V}	0.416***		0.380***		0.435***	
	(0.019)		(0.032)		(0.018)	
Score: Peers, s_{ij}^{PQ}		0.572***		0.507***		0.611***
		(0.016)		(0.032)		(0.012)
R-sq.	0.17	0.33	0.12	0.23	0.20	0.39
Clusters	188	188	171	171	177	177
Students	2,370	2,370	883	883	1,487	1,487
Student-tracks	17,460	17,460	6,433	6,433	11,027	11,027

Notes: The table presents regressions of within-town quintiles of measured values of track characteristics on households' scores. The notes to Table 5 describe the sample. Standard errors are clustered by middle school. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

The tables reveal a few notable points. First, households have relatively limited knowledge of tracks' value added on passing the baccalaureate exam. On average, households' scores for this characteristic are off by 1.1 quintiles (Table 5). A 1 point increase in a household's score is associated with only a 0.42 quintile increase in measured value added, and households' scores explain only 17% of the variation in quintiles (Table 6).

Second, households are more—if still imperfectly—aware of track selectivity. On average, households' peer quality scores are off by 0.90 quintiles in predicting prior-year selectivity (Table 5). A 1 point increase in a peer quality score is associated with a 0.57 quintile increase in the true value, and households' scores explain 33% of the variation (Table 6).

Third, households with high-achieving children have more accurate scores than those with low-achieving children. For the former, value added (peer quality) scores are off by an average of 1.09 (0.84) quintiles and explain 20% (39%) of the variation. For the latter, values are 1.19 (1.01) quintiles and 12% (23%) of the variation.

Fourth, households' scores concerning their two most-preferred tracks are only slightly more accurate than their scores concerning all tracks (Table 5). Further, these scores are biased; households tend to think their favored tracks are better than they actually are. For instance, for value added, households with low-achieving (high-achieving) children on average over-estimate the quality of their preferred tracks by 0.64 (0.63) quintiles. For prior-year selectivity, the bias is an over-estimate of 0.65 (0.17) quintiles.[26]

IV. Households respond to information on value added

The previous sections showed that households leave value added on the table and have only partially accurate beliefs regarding this attribute. We now test whether informing households about value added can influence their track choices. This could occur if information causes households to update their beliefs or if it alters their preferences over track characteristics (e.g., by making value added more salient).

A. *Effects on students' assigned tracks*

Our main outcome is the academic value added of the track that a student attends. In order to calculate the treatment effect on this outcome, we estimate:

$$\mathrm{sd}(\mathrm{V}_i) = \eta_0 + \eta_1 \cdot T_i + {\eta_X}' \cdot X_i + \eta_i. \tag{1}$$

Here, $\mathrm{sd}(\mathrm{V}_i)$ is the value added of the track of student i in standard deviation units, T_i is an indicator for whether i is in the treatment group, and X_i is a vector of i's covariates.[27] The coefficient of interest is η_1; it captures the average treatment effect of providing information.

Using this specification, Table 7 shows that the intervention had a substantial effect, but only among households with low-achieving children. Over the full sample ("All students"), providing information caused students to attend tracks with value added that was higher by 0.05 s.d. (significant at a 10% level). This amounts to an increase in the probability of passing the baccalaureate exam of 0.58 percentage points, which is small relative to the 63% predicted pass rate. For

[26]Online Appendix J shows that these results are highly robust.

[27]In our primary specification, X_i includes (i) an indicator for whether the student ranked a feasible track in the baseline survey and (ii) the value added of the track to which the student would have been assigned based on the baseline preference ranking. This latter covariate is calculated as the value added of the feasible track that the student ranked highest in the baseline survey. It is set to zero if the student did not rank any feasible tracks.

Table 7—Average treatment effects on the value added of students' tracks

	All students	Low-achieving	High-achieving
Treated	0.048*	0.121**	-0.002
	(0.025)	(0.049)	(0.023)
Effect in percentage points	0.58	1.45	-0.02
Predicted pass rate	62.9	29.2	83.2
Clusters	78	78	77
Students	2,692	1,012	1,680

Notes: The table presents results from regression (1). Low- (high-) achieving students are those with transition scores in the bottom (top) half of the national distribution. "Effect in percentage points" is the effect on the probability that a student passes the baccalaureate exam. We calculate this by multiplying the effect in standard deviation units by the 2019 standard deviation of true value added. "Predicted pass rate" is the share of students in the regression sample who are predicted to pass. We calculate this in two steps. First, we predict the probability of passing for each student by calculating the share of students with the same transition score percentile rank who passed in the 2004-2014 admission cohorts. Second, we average these values over the students in the regression sample. Standard errors are clustered by the middle school treatment-control pairs within which we conducted the randomization. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

low-achieving students, the treatment effect is 0.12 s.d. (significant at a 5% level). This is a 1.45 percentage point increase in the probability of passing, as compared to a 29% predicted pass rate. For high-achieving students, the treatment effect is virtually zero and statistically insignificant. These results are robust to a variety of alternative specifications of regression (1).[28] They are also not confounded by informational spillovers.[29]

We next investigate whether treatment effects vary based on whether a student was eligible for the tracks she preferred in the baseline. Over 95% of households expected their child to be admitted to at least one of the two tracks that they ranked highest in this survey (online Appendix H). Thus, it is possible that households were more willing to change their choices over the other tracks, given that they did not expect those choices to be relevant for their children's assignments. In this scenario, treatment effects would be larger for students who did not end up being eligible for their two top baseline choices and smaller for those who did. Importantly, almost a quarter of students fall into the former group.

The results (Table 8) are consistent with the above story. The treatment had little impact for students who were admitted to one of their top baseline choices, but it had a large effect for students who were not. In Table 8, the columns indicate which baseline choice a student was eligible for.[30] The effects are statis-

[28]For example, they are robust to controlling for different covariates and using a difference-in-difference design; see online Appendix Table A15.

[29]If treated households shared information with others in their towns—including individuals in the control group—then the effects would be biased toward zero. Online Appendix K tests for spillovers by examining whether effects are smaller in towns where we visited a larger fraction of middle schools. We find no evidence for this.

[30]The first column is for students who scored above the cutoff for their most-preferred baseline choice. The second column is for students who scored above the cutoff for their second-most-preferred baseline

Table 8—Effects on value added by eligibility for the tracks preferred in the baseline

	Eligible for x^{th} most-preferred track in the baseline					
	Most-preferred	2nd-most-preferred	$\geq$ 3rd-most-preferred	$\geq$ 4th-most-preferred	$\geq$ 5th-most-preferred	$\geq$ 6th-most-preferred
Treated	0.019	-0.072	0.184***	0.173**	0.171**	0.190**
	(0.018)	(0.102)	(0.065)	(0.066)	(0.075)	(0.084)
Effect in percentage points	0.23	-0.86	2.21	2.08	2.05	2.28
Predicted pass rate	75.6	51.7	32.8	30.5	29.0	28.3
Clusters	77	72	76	75	73	71
Students	1,766	288	638	507	427	375

Notes: The table presents results from regression (1) for subsets of students by eligibility for the tracks ranked highly in the baseline survey. "Most-preferred" is the set of students who were eligible for their most-preferred baseline track. "2nd-most- preferred" is the set who were eligible for the track that they ranked second highest in the baseline, but not for the track they ranked highest. "$\geq$ 3rd-most-preferred" is the set who were not eligible for either of their two most-preferred tracks, and so on.

tically insignificant and close to zero for students who were eligible for either their most- or second-most-preferred baseline choice. For the remaining students, the effects are always significant and range from 0.17 to 0.19 s.d. These magnitudes translate into increases in the probability of passing the baccalaureate exam of over 2 percentage points—substantial relative to predicted pass rates of 28-33%.

We next examine how the heterogeneity in Table 8 interacts with student characteristics. We estimate regression (1) for different types of students, always distinguishing between those who were eligible for their two top baseline choices and those who were not. Table 9 contains the results. Panel A provides the effects for the students who were eligible. The impacts for these students remain small and statistically insignificant regardless of achievement, gender, or mother's schooling. Panel B is for the ineligible students. The effects for these students vary little by gender or mother's schooling, but they do vary by achievement. For low-achieving students, the effect is 0.20 s.d. (significant at a 1% level). This is a 2.45 percentage point increase in the probability of passing, over a 25% predicted pass rate. Meanwhile, for high-achieving students, the effect is statistically insignificant and close to 0.

We do not find evidence that households made large tradeoffs in order to attend higher-value added tracks. Table 10 re-estimates (1) using additional track characteristics as outcomes. As before, we provide results separately for students who were and were not eligible for their most-preferred baseline choices. The results indicate that the treatment had little impact on track characteristics other than value added. This is the case even for students who were rejected by their top choices. Notably, these students attended tracks with 0.18 s.d. worth of additional value added; yet they did not experience any difference in selectivity, peer

choice, but not for their top choice. The remaining columns are for students who were eligible for only their third-most-preferred track or worse, fourth-most-preferred or worse, etc.

Table 9—Effects on value added by additional student characteristics

	All	Achievement		Gender		Mother's schooling	
		Low	High	Female	Male	$\leq$ 12 years	$>$ 12 years
Panel A: Eligible for at least one of two top baseline choices							
Treated	0.007	0.035	0.000	0.001	0.013	0.021	-0.004
	(0.024)	(0.058)	(0.022)	(0.024)	(0.038)	(0.032)	(0.028)
Effect in percentage points	0.08	0.42	0.00	0.01	0.16	0.25	-0.05
Predicted pass rate	72.3	33.7	84.0	75.3	68.6	63.4	80.3
Clusters	78	72	77	78	77	78	77
Students	2,054	479	1,575	1,120	934	981	1,073
Panel B: Ineligible for two top baseline choices							
Treated	0.184***	0.204***	-0.023	0.193**	0.180*	0.154*	0.221*
	(0.065)	(0.069)	(0.123)	(0.096)	(0.105)	(0.082)	(0.127)
Effect in percentage points	2.21	2.45	-0.28	2.32	2.16	1.85	2.65
Predicted pass rate	32.8	25.1	72.1	32.7	32.9	28.0	42.7
Clusters	76	76	28	71	67	75	64
Students	638	533	105	306	332	430	208

Notes: The table presents results from regression (1) for subsets of students. The subsets represent the interaction between student characteristics (achievement, gender, or mother's schooling) and whether the student was eligible for at least one of the two tracks s/he listed as most preferred in the baseline survey. See the notes to Table 7 for additional details on the regressions.

socio-economic status, or location quality. The only tradeoff that these students were induced to make relates to curricular focus: they were 7 percentage points less likely to attend a technical track. The reason for this is that, conditional on selectivity, technical tracks tend to have lower academic value added than humanities or math and science tracks (online Appendix Figure A4).[31]

In short, the treatment had heterogeneous effects on the value added of students' tracks and little effect on other track characteristics. For value added, the treatment had no impact for high-achieving students or for low-achieving students who were admitted to one of their preferred baseline choices. By contrast, it had large impacts for low-achieving students who were rejected by these choices. We next try to understand this heterogeneity by using the follow-up survey to investigate effects on beliefs and track preference rankings.

[31]The fact that the treatment induced students to switch out of technical tracks is a potential concern. The impacted students were largely low-achieving (Table 9), and it is possible that technical tracks—being more career-oriented—are a good option for students unlikely to succeed in college. In this story, these tracks' low test-related value added would mask higher wage value added. We are unable to measure wage value added. Nonetheless, we explore this story using our data on households' beliefs. For each value added dimension we asked about at baseline, we calculate the mean quality score (across households) for a given track. These variables exploit the "wisdom of the crowd" to measure value added. We use the variables as outcomes in regressions akin to those in Table 10. If the tracks that students switched out of were believed to have high wage value added, then treatment effects on mean quality scores for this outcome would be smaller than on scores for academic value added. The results do not support this (online Appendix Table A16). They show positive and similarly sized treatment effects on scores for all value added dimensions. That is, the treatment induced students to switch to tracks believed to be better for both academics and wages.

Table 10—Effects on other characteristics of students' tracks

	Value added	Selectivity	Peer SES	Location quality	Curricular focus		
					Humanities	Math & science	Technical
Panel A: Eligible for at least one of two top baseline choices							
Treated	0.007	0.009	0.017	0.015	-0.013	0.012	0.001
	(0.024)	(0.017)	(0.013)	(0.020)	(0.017)	(0.016)	(0.010)
Clusters	78	78	78	78	78	78	78
Students	2,054	2,054	2,054	1,978	2,054	2,054	2,054
Panel B: Ineligible for two top baseline choices							
Treated	0.184***	0.006	-0.009	0.073	0.039	0.030	-0.069**
	(0.065)	(0.050)	(0.052)	(0.086)	(0.030)	(0.027)	(0.033)
Clusters	76	76	76	76	76	76	76
Students	638	638	638	492	638	638	638

Notes: The table estimates regression (1) for different outcomes. Columns 1 and 2 refer to value added and the minimum transition score. Column 3 refers to the average transition score in the middle schools of a track's students. This is a measure of track peer SES because Romania has neighborhood-based middle schools. The outcomes in Columns 1-3 are all in standard deviation units. The outcome in Column 4 is a household's baseline score for a track's location quality. Columns 5-7 refer to a track's curricular focus. Regressions control for values of the outcome variable for the feasible track that the household's ranked highest at baseline. This is the track to which the household would have been assigned based on its baseline ranking. The regressions also include indicators for students who did not rank any feasible tracks. Standard errors are clustered by middle school treatment-control pairs.

B. *Effects on beliefs regarding value added*

This subsection presents treatment effects on beliefs—it explores whether providing information increased the accuracy of households' scores for academic value added. Before turning to results, we note that there are two ways in which this analysis may understate effects on beliefs. First, it is possible for information to influence the precision of households' beliefs without changing their scores.[32] Second, the scores may contain measurement error: the follow-up survey took place a few weeks after households submitted their preference rankings; by this time, households may have forgotten some of what they knew when they were deciding their track preferences. Despite these caveats, our sense is that treatment effects on quality scores are a useful proxy—and possibly a lower bound—for those on beliefs.

To conduct the analysis, we estimate:

$$|\text{quint}(\text{V}_{jt}) - s^{V}_{ij,\text{fs}}| = \eta_0 + \eta_1 \cdot T_i + {\eta_X}' \cdot X_{ij} + \eta_{ij}. \tag{2}$$

Here, $|\text{quint}(\text{V}_{jt}) - s^{V}_{ij,\text{fs}}|$ is the absolute difference between: (i) the value added of track j in units of within-town quintiles, $\text{quint}(\text{V}_{jt})$, and (ii) household i's

[32]Suppose that a track's true score is 4, and that during the baseline a household believed the track had an equal chance of being a 3, 4, or 5. In this case, the household would assign a score of 4 but would be uncertain about its decision. Our treatment would remove the uncertainty, although the score would stay the same.

quality score for the track's value added from the follow-up survey, $s^V_{ij,\text{fs}}$. As in regression (1), the coefficient of interest is η_1. It represents the average treatment effect on the absolute error of households' quality scores. If the treatment caused households' scores to become more accurate, then η_1 will be negative.

Table 11—Effects on the accuracy of households' value added scores

	x^{th} **most-preferred track in the baseline**						
	All tracks	Most-preferred	2nd-most-preferred	$\geq$ 3rd-most-preferred	$\geq$ 4th-most-preferred	$\geq$ 5th-most-preferred	$\geq$ 6th-most-preferred
Treated	-0.055	0.032	-0.033	-0.101**	-0.124**	-0.156**	-0.181***
	(0.034)	(0.041)	(0.053)	(0.045)	(0.053)	(0.060)	(0.063)
Mean abs. difference: baseline	1.02	0.93	1.07	1.06	1.11	1.16	1.18
Mean abs. difference: follow-up	1.00	0.86	1.03	1.06	1.12	1.14	1.15
Clusters	76	76	75	76	76	76	76
Students	1,525	1,263	962	1,352	1,134	967	868
Student-tracks	4,970	1,263	962	2,745	2,100	1,727	1,487

Notes: The table presents results from regression (2). The values in the row labeled "Treated" are the estimates for η_1. The columns provide results for different sets of tracks. "Most-preferred" refers to the track that a household ranked highest in the baseline survey. "2nd-most-preferred" is the track ranked second highest. "$\geq$ 3rd-most-preferred" are all tracks other than the two most preferred. The remaining columns are defined analogously. The regressions include indicators for the value of the absolute difference between: (i) the within-town quintile of a track's value added, quint(V_{jt}), and (ii) the household's baseline score for the track on this dimension, s^V_{ij}. "Mean abs. difference: baseline" is the mean absolute difference between quint(V_{jt}) and s^V_{ij} for the sample. Similarly, "Mean abs. difference: follow-up" is the mean absolute difference between quint(V_{jt}) and $s^V_{ij,\text{fs}}$. Standard errors are clustered by the middle school treatment-control pairs within which we conducted the randomization.

The results indicate that the treatment increased the accuracy of households' value added scores, but only for their less-preferred tracks. In Table 11, the first column is for all tracks, while the others distinguish by a track's position in a household's baseline preference ranking. For the full set of tracks, the treatment led to a statistically insignificant improvement in accuracy of 0.06 quintiles—small relative to the mean inaccuracy of about one quintile. For the tracks that households' ranked highest and second-highest in the baseline survey, effects are close to zero. For the remaining tracks, improvements are sizable and are all significant at either the 1% or 5% confidence level. The results also reveal that the changes in accuracy grow larger for tracks that are farther down a household's baseline preference ranking. For instance, among tracks other than a household's two top baseline choices, the improvement is 0.10 quintiles; among tracks other than a household's five top baseline choices, it rises to 0.18 quintiles.

The above pattern holds regardless of whether households have low- or high-achieving children. In each case, providing information increased the accuracy of scores, but only for tracks that households did not initially prefer (online Appendix Tables A17 and A18). That said, magnitudes are twice as large for households with low-achieving children as for those with high-achieving ones.

C. Effects on track preference rankings

We next analyze whether information caused households to assign higher preference ranks to tracks with higher value added. To do this, we calculate treatment effects on the association between the within-town percentile rank of a track's value added and the percentile rank of the track in a household's preference ranking. We estimate:

$$\text{(3)} \qquad \text{ppr}_{ij,\text{fs}} = (\delta_1 + \delta_2 \cdot T_i) \cdot \text{pr}(\text{V}_{jt}) + (\delta_{X,1} + \delta_{X,2} \cdot T_i)' \cdot X_{ij} + \delta_{ij}.$$

Here, $\text{ppr}_{ij,\text{fs}}$ is household i's percentile preference rank for track j, as reported in the follow-up survey.[33] It is calculated by dividing a track's rank in the household's preference ranking by the number of tracks in the town. The variable is ordered such that a value of 1 indicates a household's most-preferred track.[34] Next, $\text{pr}(\text{V}_{jt})$ is track j's within-town percentile rank of value added. It is calculated by dividing the track's within-town value added rank by the number of tracks in the town. To be consistent with $\text{ppr}_{ij,\text{fs}}$, it is ordered such that a value of 1 indicates the town's best track by value added. X_{ij} is a set of indicators for track j's position in household i's baseline preference ranking. The coefficient of interest is δ_2; it measures the effect of the treatment on the association between value added and preference ranks.[35]

The results exhibit a pattern similar to that for effects on the accuracy of households' value added scores (Section IV.B): providing information increased the association between preference ranks and value added, but only among tracks that households did not initially prefer. Table 12 shows that among all tracks, the treatment caused the association to be higher by 0.05 percentiles (significant at the 10% level). For the two tracks that a household ranked highest in the baseline survey, the effect is insignificant and of the wrong sign. After excluding these tracks, the effect rises to 0.06 percentiles and is significant at the 1% confidence level. Moreover, the effect continues to grow for tracks that are farther back in the baseline preference ranking.

This pattern persists for households with low- and high-achieving children. For both, effects exist only among tracks other than the two top baseline choices. That said, for households with high-achieving children, the effects are small and mostly insignificant; for those with low-achieving children, they are sizable and

[33] At the time of the baseline survey, we told households that we would contact them after the allocation, for a follow-up survey. We requested that they save a copy of their official track preference ranking. During the follow-up, we asked households to find their copy and read off the ranking. We therefore believe that the rankings reported in the follow-up closely approximate those submitted. For instance, 99% of respondents report that their child attends the track that we observe them attending in the administrative data.

[34] We set $\text{ppr}_{ij,\text{fs}}$ equal to 0 for tracks that households do not rank, since students cannot be assigned to these. In particular, if a household ranks all J_i tracks in its town, its least-preferred track has $\text{ppr}_{ij,\text{fs}} = 1/J_i$. Unranked tracks should thus have a value of $\text{ppr}_{ij,\text{fs}}$ that is less than $1/J_i$; we choose 0.

[35] To see this, note that δ_1 is the average slope of conditional-on-X_{ij} best-fit lines between $\text{ppr}_{ij,\text{fs}}$ and V_{jt} for households in the control group; $\delta_1 + \delta_2$ is the average slope of these lines for treated households.

Table 12—Effects on the association between value added and preference rankings

	x^{th} most-preferred track in the baseline					
	All tracks	Two most-preferred	$\geq$ 3rd-most-preferred	$\geq$ 4th-most-preferred	$\geq$ 5th-most-preferred	$\geq$ 6th-most-preferred
Value added: treated	0.049*	-0.072	0.062***	0.064***	0.068***	0.069***
	(0.026)	(0.103)	(0.023)	(0.023)	(0.023)	(0.023)
Association: baseline	0.434	0.018	0.269	0.179	0.102	0.055
Association: follow-up	0.345	0.067	0.213	0.168	0.149	0.141
Clusters	76	76	76	76	76	76
Students	1,533	1,523	1,533	1,533	1,533	1,514
Student-tracks	20,029	2,937	17,092	15,849	14,779	13,938

Notes: The table presents results from regression (3). The values in the row labeled "Value added: treated" are the estimates for δ_2. The columns provide results for different sets of tracks. "Two most-preferred" refers to the two tracks that households ranked highest at baseline. "$\geq$ 3rd-most-preferred" are all tracks other than the two most preferred. The remaining columns are defined analogously. The regressions include indicators for the interaction between a track's position in a household's baseline ranking and whether the household is in the treatment group. "Association: baseline" is the slope coefficient from a regression of the percentile preference rank from the baseline survey, ppr_{ij}, on $\text{pr}(\text{V}_{jt})$. "Association: follow-up" is the slope coefficient from a regression of $\text{ppr}_{ij,\text{fs}}$ on $\text{pr}(\text{V}_{jt})$. Standard errors are clustered by the middle school treatment-control pairs within which we conducted the randomization.

significant at a 1% level (online Appendix Tables A19 and A20).

D. Discussion

The results for beliefs and preference rankings help to explain the heterogeneity in impacts on the value added of students' tracks. For high-achieving students, providing information had modest effects on beliefs and little effect on preference rankings. As a result, it did not cause these students to attend tracks with higher value added. For low-achieving students, the information did affect beliefs and preference rankings, but only for tracks that were initially less preferred. Thus, for this group, impacts on track assignments differ depending on whether a student was eligible for her top baseline choices. The treatment had no influence on assignments for students who were eligible for these choices, but it had significant impacts for students who were not.

As noted, the fact that households were more receptive to information for tracks other than their top baseline choices is likely a consequence of the approach that they used to rank tracks. The tracks that households ranked highest were ones that they thought would be feasible and that they thus expected their child to attend.[36] It may be that households were less attached to their beliefs and preference rankings for the other tracks because they did not think those tracks would be relevant. This behavior is consistent with evidence that searching for information on school quality is costly (Arteaga et al. 2021).

A separate question is why responses were larger for households with low-

[36] Recall that more than 95% of households expected their child to be admitted to at least one of their two top baseline choices (online Appendix H).

achieving children than for those with high-achieving children. One explanation is that households with high-achieving children may have been more certain of their beliefs and rankings at the time of the baseline. If so, our intervention may have come too late in their decision-making process. We assess this by comparing the two groups' self-reported certainty about their baseline preference rankings. Households with high-achieving children were indeed more certain; 45% reported being very certain, 43% were somewhat certain, and 12% were uncertain. For households with low-achieving children, the corresponding percentages are 33%, 50%, and 17% (online Appendix F).

A second explanation is that households with low-achieving children may be more trusting of information provided by outside authority figures. If so, treatment effects on beliefs and preference rankings would be larger for low-achieving students even after conditioning on certainty. Two results emerge in this regard (online Appendix Table A21). First, for both low- and high-achieving students, effects are larger for households who reported being uncertain or somewhat certain than for those who were very certain. Second, within the two certainty groups, effects are larger for low-achieving students.

In short, the evidence suggests that both stories play a role in explaining why households with low-achieving children were more receptive to the information. These households were less likely to have settled on their beliefs and preference rankings when the intervention occurred. In addition, they exhibited larger responses conditional on their degree of certainty.

V. Preferences for other traits limit demand for value added

We next explore how demand for academic value added is constrained by households' preferences for other track characteristics. We first estimate preferences—in a discrete choice model, we explain households' track preference rankings using their quality scores. We then predict how track choices would change if households' scores for value added were made to be fully accurate. This exercise allows us to decompose the value added that households leave unexploited into two components: one that is due to preferences and another that is due to inaccurate beliefs.

We caution that the results do not provide a firm upper bound on the impact of providing information. This is because households' choices may depend on the precision of their beliefs in ways not captured by the quality scores; in addition, providing information may influence preferences, such as by signaling the importance of value added. With these caveats in mind, we conclude by comparing the magnitude of our experimental treatment effects with those predicted under accurate quality scores.

A. Households' preferences for track characteristics

We estimate preferences for track characteristics by relating households' track preference rankings to their beliefs about the attributes of the tracks. For simplicity, we use baseline values of track preference rankings and beliefs.[37]

Specifically, we assume that households rank tracks according to expected utility.[38] We then write a household's baseline expected utility from a track as a linearly separable function of its baseline survey scores for the track (on a scale from 1-5; Table 2) on various quality dimensions. This is:

$$U_{ij} = \sum_q \beta_q \cdot s^q_{ij} + \epsilon_{ij}, \tag{4}$$

where U_{ij} is household i's baseline expected utility from track j, s^q_{ij} is the household's baseline score for the track on quality dimension q, and ϵ_{ij} is an error term. The β_q coefficients reflect households' preferences for track attributes; they represent the change in expected utility associated with a one-unit increase in a given quality score. To estimate (4), we assume that the error term, ϵ_{ij}, is independent and follows a Type-1 Extreme Value distribution. We then fit the model to households' baseline preference rankings using a rank-ordered logit.[39]

Table 13 presents the estimates for the β_q coefficients, which indicate that households care about a variety of track characteristics. Column 1 presents our benchmark model. It shows that households have similar preferences for a track's location (coef. estimate of 0.28), for whether a child's siblings and friends use the track (0.34), for the track's peer quality (0.34), and for the track's value added on passing the baccalaureate exam (0.34). By contrast, households have consid-

[37] In online Appendix L, we instead use endline values, and we exploit experimental variation in beliefs. We show that results are very similar.

[38] This is weakly dominant since the track assignment mechanism is incentive compatible. The dominance is strict for tracks that a household believes its child has a chance of attending. In our main analysis, we assume that households consider all the tracks in their towns. Nonetheless, the results are robust to excluding tracks that households may have considered "out of reach."

[39] A rank-ordered logit is a series of multinomial logits corresponding to each choice in a preference ranking. In practice, we do not use all the constituent multinomial logits. We use just those for a household's top choices. In our main results, we consider a households' *two* top choices. That is, we maximize the probability that a household prefers its highest-ranked track, r_{i1}, to all other tracks in the town times the probability that the household prefers its second-highest-ranked track, r_{i2}, to all other tracks except r_{i1}. Letting $\mathcal{J}_i$ be the set of tracks in i's town, this likelihood is:

$$\Pr[r_{i1}, r_{i2} | \mathcal{J}_i, \{s^q_{ij}\}_{q,j}] = \prod_{l=1}^{2} \frac{\exp[\sum_q \beta_q \cdot s^q_{ir_{il}}]}{\sum_{k \in \mathcal{J}_i \setminus \{r_{im}: m<l\}} \exp[\sum_q \beta_q \cdot s^q_{ik}]}. \tag{5}$$

We focus on the first two choices because most households appear to have settled on these by the time of the baseline survey. However, in Section V.B, we also run specifications with different numbers of choices. In addition, we sometimes restrict attention to tracks that are plausibly feasible. When we do this, we re-define a household's top choices as the most preferred among this narrower set. There are three ways in which our approach may fail to recover preferences. First, we may be omitting a quality dimension that is correlated with both utility and the s^q_{ij} covariates. This concern is mitigated because we have scores on a large number of quality dimensions. Second, the quality scores may contain measurement error—we consider this issue in the main text. Third, it is possible that households are risk averse with respect to track characteristics. In this case, utility would not be linear in the characteristics—as in (4)—but rather strictly concave. Thus, expected utility would depend on the precision of households' beliefs in a manner not captured by quality scores. For instance, a household would gain more expected utility when it knows a track is a 4 than when it thinks the track has an even chance of being a 3, 4, or 5. We ignore this issue because accounting for it would require data on the full density of beliefs.

Table 13—Households' preferences for track attributes

	(1)	(2)	(3)	(4)	(5)
Location	0.276***	0.292***	0.277***	0.281***	0.289***
	(0.069)	(0.069)	(0.069)	(0.066)	(0.069)
Siblings and friends	0.336***	0.326***	0.319***	0.344***	0.311***
	(0.048)	(0.048)	(0.048)	(0.048)	(0.048)
Peer quality	0.344***	0.317***	0.318***	0.380***	0.298***
	(0.069)	(0.066)	(0.069)	(0.067)	(0.074)
Curricular focus	0.931***	0.789***	0.877***	0.986***	0.763***
	(0.071)	(0.069)	(0.068)	(0.073)	(0.068)
VA: pass the bacc.	0.337***				0.013
	(0.082)				(0.083)
VA: college		0.519***			0.347***
		(0.073)			(0.082)
VA: wages			0.485***		0.320***
			(0.064)		(0.069)
Teacher quality				0.180**	-0.026
				(0.088)	(0.080)
R-sq.	0.33	0.33	0.33	0.32	0.34
Clusters	150	150	150	150	150
Students	1,170	1,157	1,151	1,168	1,137
Student-tracks	11,575	11,395	11,382	11,573	11,220

Notes: The table presents results from equation (4). The model is estimated by maximizing the log-likelihood corresponding to equation (5). The sample is limited to students in experimental middle schools. Standard errors are clustered by middle school.

erably stronger preferences over a track's curricular focus (0.93). All estimates are significant at a 1% confidence level.

Columns 2-5 explore preferences for alternative dimensions of value added: value added on college quality, value added on wages, and a track's teacher quality. When included one at a time, all value added dimensions are statistically significant (Columns 2-4). However, the results suggest that households care most about value added on college and wages. For instance, in a horse race (Column 5), the β_q estimates are large and significant for these dimensions, but small and insignificant for passing the baccalaureate exam and for teacher quality. One interpretation is that households may see the two latter dimensions as inputs into the former two.

We find some heterogeneity in preferences between households with low- and high-achieving children (online Appendix Table A22). First, the two groups differ in their preference for peer quality: coefficient estimates are large (0.43-0.56) and statistically significant for high-achievers, but small (0.06-0.13) and insignificant for low-achievers. Second, the groups' quality scores have differing degrees of explanatory power. Depending on the specification, these explain 40-41% of the variation in track choices for high-achievers and only 20-22% for low-achievers.

Our preference estimates are robust to a variety of potential issues. A first concern is that few households provide scores for all the tracks in their towns.

Missing scores could introduce bias if a household's propensity to score a track depends on its preference for the track. We gauge the impact of missing scores using two approaches. First, we limit the sample to households without missing scores. Second, we impute the missing scores using a random forest.[40] In both cases, results are similar to those for our main specification (online Appendix Table A23).

A second concern stems from the "skipping" issue highlighted by Fack, Grenet and He (2019), Artemov, Che and He (2020). If households refrain from ranking tracks that they believe will not admit their child, then their rankings will not reflect their true preferences. To assess this issue, we run two specifications that exclude tracks that households may have considered "out of reach". Again, results are unchanged (online Appendix Table A24). Importantly, the preference estimate for peer quality among low-achieving students remains small. Thus, this value appears to be a reflection of preferences rather than an artifact of skipping.

A final concern is that the quality scores may contain measurement error—that is, they may be a noisy proxy for the expectation of a household's beliefs. Measurement error would cause the preference estimates to be attenuated. To explore this concern, we use a horse race proposed by Kapor, Neilson and Zimmerman (2020). The approach involves re-estimating the preference model while adding controls for measured values of track characteristics. Specifically, for the attributes of curricular focus, peer quality, and academic value added, we have both quality scores and measured values. Thus, we can test the quality scores by including the measured values. If the quality scores are noisy, then the measured values may contain additional information about households' beliefs, in which case they would provide additional explanatory power for expected utility.[41] The results suggest that measurement error is a modest issue (online Appendix Table A25). The measured values are often statistically significant; however, they generate only small increases in R-squared, and they exert limited impact on the coefficients for the quality scores.[42]

B. *Track choices under accurate beliefs*

Next, we simulate how track choices would change if households had accurate beliefs about academic value added.[43] Using the preference model, we predict choice outcomes under two sets of quality scores. The first set, *Inaccurate scores*,

[40]For each quality dimension, we predict a household's score for a track using covariates including: (i) characteristics of the track, (ii) characteristics of the student, and (iii) quality scores for the track from other households in the same town or middle school as the student. We replace missing scores with these predictions.

[41]The measured values may be significant even absent measurement error. For example, they may be correlated with omitted quality dimensions. In this way, the test can provide evidence for measurement error but not proof.

[42]For scores for location, siblings and friends, and curricular focus, the coefficients are not changed at all. For scores for value added, coefficients fall by 31% (51%) for low- (high-) achieving students. For scores for peer quality, coefficients fall for high-achieving students and rise slightly for low-achieving ones.

[43]The analysis in this section holds constant households' feasible choice sets. Thus, it lends insight into choice behavior, but it does not reveal the impacts of a large-scale policy of information provision.

uses households' baseline scores for value added. The second set, *Accurate scores*, replaces these with within-town quintiles of measured value added.[44]

For each track in a household's feasible choice set, we predict the probability that the household would prefer the track, given the scores. We then use the probabilities to predict the value added of the track the student would attend. This latter prediction is a weighted average of the value added of each feasible track, with weights that are equal to the predicted preference probabilities. For *Inaccurate scores*, it is:

$$\mathrm{V}_{i,\mathrm{IS}} \equiv \sum_{j \in \mathcal{J}_i^e} \mathrm{sd}(\mathrm{V}_{jt}) \cdot \frac{\exp[\sum_q \hat{\beta}_q \cdot \tilde{s}_{ij}^q]}{\sum_{k \in \mathcal{J}_i^e} \exp[\sum_q \hat{\beta}_q \cdot \tilde{s}_{ik}^q]}.$$

Here $\hat{\beta}_q$ is a coefficient estimate from the preference model (4), $\mathcal{J}_i^e$ is household i's feasible choice set, and $\tilde{s}_{ij}^q$ is a score in *Inaccurate scores*. For the prediction for *Accurate scores*, $\mathrm{V}_{i,\mathrm{AS}}$, the formula is analogous but with correct scores for value added.

Under each set of scores, we produce four versions of our predictions. These reflect different assumptions about the preference model and about how households update their beliefs in response to information. Our first specification is titled "Just quality scores". It uses a preference model akin to that in Column 1 of Table 13, controlling for scores for location, siblings and friends, peer quality, curricular focus, and value added on passing the baccalaureate exam.[45] The second specification—"With measured attributes"—adds measured values of track characteristics. The third specification—"Update on all VA dimensions"—supposes that households may update their beliefs on all the value added dimensions that we asked about in the survey, not just on value added with respect to passing the baccalaureate exam. The preference model for this specification is similar to that in "With measured attributes"; however, it includes quality scores for each of the four value added dimensions. Further, when calculating $\mathrm{V}_{i,\mathrm{AS}}$ for this specification, we "correct" the scores for all of these dimensions.[46] Finally, the fourth specification—"Adjust for measurement error"—accounts for the fact that the preference estimate for value added may be attenuated by measurement error.

In particular, if all households were somehow made to have correct beliefs, then the choice setting would change due to dynamic effects on track selectivity, teacher sorting, value added, etc. We leave these effects for future work.

[44]For the other quality dimensions, we always use households' baseline scores. The only exception is peer quality, where we substitute the within-town quintile of a track's selectivity. We make this substitution because households have access to information on selectivity and may absorb this information before making their final choices. We impute missing baseline scores using a random forest.

[45]In practice, the model differs in two ways from that in Table 13. First, we allow coefficients to vary based on whether a student is low- or high-achieving (as in online Appendix Table A22). Second, we estimate the model after imputing missing quality scores with the predictions from the random forest (as in online Appendix Table A23). Thus, the coefficients for this preference model are those in the third and sixth columns of online Appendix Table A23. We present coefficients for all the preference models that we use in this section in online Appendix Table A26.

[46]We replace them with the within-town quintile of value added on passing the baccalaureate exam.

It uses the same preference model as "With measured attributes" but inflates the value added coefficient by a factor of 1.5.

Before turning to results, we validate our approach by comparing our predictions with households' observed choices. In particular, *Inaccurate scores* is meant to represent households' beliefs in the absence of the experiment. As such, for households in the control group, $V_{i,\mathrm{IS}}$ should match the value added of students' actual tracks.[47] We find that $V_{i,\mathrm{IS}}$ has strong predictive power.[48]

Table 14—The effect of accurate beliefs on the value added of students' tracks

	Potential increase in VA		**Change**	**Share of**
	$V_{i,\mathrm{IS}}$	$V_{i,\mathrm{AS}}$	**in VA**	**pot. incr.**
Panel A: Just quality scores				
All students	0.894	0.679	0.215	0.240
Low-achieving	0.847	0.649	0.198	0.234
High-achieving	0.922	0.697	0.225	0.244
Panel B: With measured attributes				
All students	0.853	0.743	0.110	0.129
Low-achieving	0.733	0.607	0.126	0.172
High-achieving	0.925	0.824	0.101	0.109
Panel C: Update on all VA dimensions				
All students	0.860	0.691	0.169	0.196
Low-achieving	0.736	0.569	0.167	0.227
High-achieving	0.934	0.764	0.170	0.182
Panel D: Adjust for measurement error				
All students	0.848	0.690	0.158	0.187
Low-achieving	0.723	0.542	0.181	0.251
High-achieving	0.923	0.778	0.145	0.157

Notes: The table describes our predictions for how the value added of students' tracks would change under accurate beliefs. Results are reported in standard deviations of value added. Results in levels of value added can be obtained by multiplying by 12 percentage points (the 2019 standard deviation), and are presented in online Appendix Table A27. "Potential increase in VA" is the mean difference between (i) the maximum value added in students' feasible sets and (ii) the listed variables. "Change in VA" is the mean difference between $V_{i,\mathrm{AS}}$ and $V_{i,\mathrm{IS}}$. "Share of potential increase" is the ratio of "Change in VA" to the potential increase under inaccurate scores. See Section V.B for the definitions behind Panels A-D. The sample includes 997 low-achieving and 1,680 high-achieving students. It is similar to that for the experimental treatment effects from Section IV. However, it excludes 15 students who did not score above the cutoff for any tracks that existed in both 2018 and 2019. (These students were assigned to tracks that were newly created in 2019 and for which we did not elicit beliefs.)

Table 14 presents the results of the simulation, indicating that inaccurate beliefs play a limited role in explaining why households leave value added unexploited. This finding holds for each of the four specifications that we use to make predictions. The column labeled "Potential increase in VA: $V_{i,\mathrm{IS}}$" reveals how much

[47]The analogous exercise for households in the treatment group would compare $V_{i,\mathrm{AS}}$ with the value added of students' tracks. However, this would be inappropriate because the treatment did not fully influence households' beliefs about value added. For instance, it had no effect on beliefs regarding the two tracks that households ranked the highest in the baseline survey.

[48]Online Appendix Figure A7 plots the value added of control students' tracks against $V_{i,\mathrm{IS}}$. It also includes best fit lines from linear regressions. For the "Just quality scores" specification, the R-squared of the best-fit line is 0.61; for the other specifications, it is 0.67. The slope coefficients are all close to 1.

value added households would leave unexploited, on average, under *Inaccurate scores*. Across specifications, values range from 0.72 to 0.85 s.d. for low-achieving students and 0.92 to 0.93 s.d. for high-achieving ones. "Potential increase in VA: $V_{i,AS}$" provides corresponding results for *Accurate scores*. These range from 0.54 to 0.65 s.d. for low-achieving students and 0.70 to 0.82 s.d. for high-achieving ones. "Change in VA" displays the mean difference between $V_{i,AS}$ and $V_{i,IS}$—or the predicted treatment effect of correcting households' scores. This column shows that under correct scores, low- (high-) achieving students would, on average, attend tracks with 0.13 to 0.20 (0.10 to 0.23) s.d. of additional value added. Finally, "Share of potential increase" divides the predicted treatment effects by the potential increase in value added under *Inaccurate scores*. It thus reveals the percent of the unexploited value added that is due to inaccurate beliefs. For low- (high-) achieving students, these percentages vary from 17-25% (11-24%).[49]

The results are depicted graphically in Figure 3, which is similar to Figure 2 in Section II. The figure shows how our predictions for the value added of students' tracks compare to students' feasible options; in addition, it reveals how these patterns vary based on a student's achievement. As in Table 14, we can see that correcting households' value added scores would cause students to attend tracks with higher value added; however, the effects would represent only a fraction of the potential increase. This holds throughout the achievement distribution.

The fact that households would continue leaving value added unexploited under accurate beliefs reflects that they are constrained by their preferences for other track characteristics—mostly by those for curricular focus and peer quality. Table 15 shows the mean potential increase in value added (under accurate beliefs) if we were to set certain preference coefficients, $\hat{\beta}_q$, to 0 when calculating $V_{i,AS}$. If households cared only about value added and curricular focus, they would leave an average of 0.44 s.d. of value added on the table (Column 1). If they cared only about value added, curricular focus, and peer quality, they would leave 0.57 s.d. unexploited (Column 3). Meanwhile, in our main simulation—in which households care about a variety of track characteristics—they are predicted to leave 0.69 s.d. unexploited (Column 5). Thus, 83% of the value added left on the table is due to preferences for curricular focus and peer quality.[50]

We conclude by comparing the results from our simulation with the experimental treatment effects from Section IV. We note that the simulation results may seem small given the size of the observed treatment effects. For instance, for the students for whom the treatment had an impact—low-achieving students who were not admitted to their two top baseline choices—the effect was 0.20 s.d.

[49]Table 14 estimates the preference model using only the first two choices in a household's preference ranking. Further, it defines a choice set as the full set of tracks in the household's town. We explored alternative ways of estimating the model (online Appendix Table A28). In some specifications we use different numbers of choices and in others we limit the choice set to tracks that households could have expected to be feasible. These changes do not alter our findings. For instance, across all specifications, the largest predicted treatment effect is 0.27 s.d.

[50]There is some heterogeneity by achievement. For low-achieving students, choices are constrained largely by preferences for curricular focus (Column 1) and unexplained factors (Column 5).

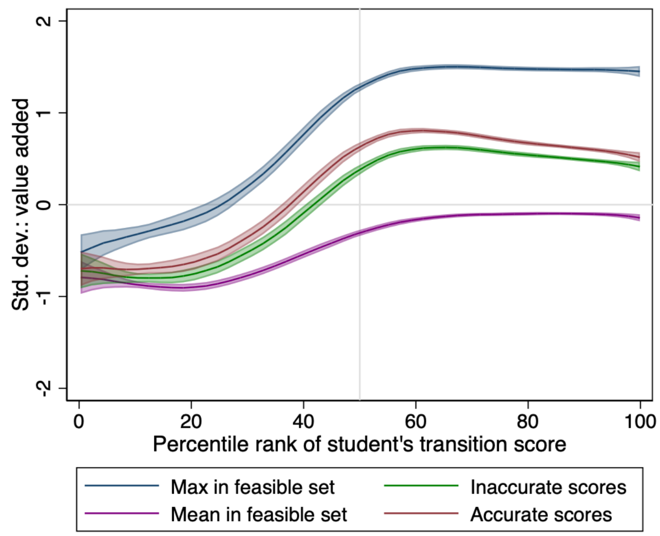

FIGURE 3. THE VALUE ADDED OF STUDENTS' TRACKS UNDER ACCURATE BELIEFS

Notes: The figure shows how the value added of students' tracks would change under accurate beliefs. It plots the relationship between the percentile rank of the student's transition score and multiple value added variables. The blue (purple) line is the maximum (mean) value added in the student's feasible set. The green and brown lines are the value added predictions, $\mathrm{V}_{i,\mathrm{IS}}$ and $\mathrm{V}_{i,\mathrm{AS}}$, respectively. All lines are calculated using local linear regressions. The predictions, $\mathrm{V}_{i,\mathrm{IS}}$ and $\mathrm{V}_{i,\mathrm{AS}}$, are from the "Adjust for measurement error" specification. The sample is as in Table 14.

This value is as large as the largest predicted treatment effect for low-achieving students in Table 14.[51]

A possible explanation is that the treatment may have affected choices via channels other than the accuracy of households' beliefs. As discussed, expected utility may depend on the precision of beliefs in ways not captured by the quality scores. In addition, providing information may cause households to care more about value added. If so, then the predicted treatment effects in Table 14 do not represent an upper bound on the impact of providing information. Nonetheless, these other channels would have to be sizable in order to change our main finding. In other words, it appears that households are likely to leave substantial academic value added on the table, even under correct beliefs. This is due to their preferences for other track attributes.

[51] In fact, for low-achieving who were not admitted to their two top baseline choices, the largest predicted treatment effect is only 0.19 s.d. Moreover, it seems unlikely that the treatment caused these households to have fully accurate beliefs about value added.

Table 15—The effect of preferences on the amount of unexploited value added

	Potential increase in VA: $\mathrm{V}_{i,\mathrm{AS}}$				
	(1)	(2)	(3)	(4)	(5)
All students	0.44	0.24	0.57	0.60	0.69
Low-achieving	0.22	0.03	0.23	0.28	0.54
High-achieving	0.57	0.36	0.78	0.78	0.78

Notes: The table shows how preferences for track attributes constrain households' choices with respect to value added. It presents the mean potential increase in value added, under *Accurate scores*, for different versions of the simulation. The versions differ in that they set $\hat{\beta}_q = 0$ for different quality dimensions in calculating $\mathrm{V}_{i,\mathrm{AS}}$. All versions use the "Adjust for measurement error" specification. Column (1) is if households care only about value added and curricular focus; (2) is if they care only about value added and peer quality; (3) is for value added, curricular focus, and peer quality; (4) is for value added, curricular focus, peer quality, siblings and friends, and location; (5) is for all the attributes in "Adjust for measurement error", including unexplained idiosyncratic factors. (5) is the same as in Panel D of the second column of Table 14. Results are reported in standard deviations of value added; results in levels can be obtained by multiplying by 12 percentage points (the 2019 standard deviation). See Table 14 for details on the sample.

VI. Conclusion

Recent research studies how to allocate students to schools—for example, how to implement mechanisms that free households from strategizing as they apply to schools. A separate question concerns what incentives the resulting demand patterns generate for schools. If demand reflects households' desire for value added, then schools may feel pressure to raise their value added. By contrast, if demand reflects households' desire for peer quality, then schools might focus on becoming more selective, and so forth (Rothstein 2006, Abdulkadiroglu et al. 2020).

Why might households not always demand the schools that researchers deem most productive? What constraints or factors might lead them to leave value added on the table? We have considered two possibilities. First, households may lack *information.* Value added is difficult to observe, even for researchers with access to ample data. Thus, it is possible that households do wish to attend high-value added schools, but do not know which those are. On the other hand, it may be that households' *preferences* lead them to prioritize other school traits. For example, a given school might not provide the largest gains in skill, but it may offer a short commute or desirable peers. In this case, households may willingly give up value added in exchange for other dimensions of quality (MacLeod and Urquiola 2019).

Our results suggest that both candidate explanations are relevant. We find that essentially all types of households make school choices that leave value added on the table. In addition, our experiment shows that distributing information can affect households' school rankings, placements, and value added—particularly for households with low-achieving students. That said, our simulations suggest that even correcting all informational shortfalls would leave households far from maximizing school value added.

We note that the effects of information could be larger or smaller in other settings. On the one hand, Romanian public schools are relatively homogenous in terms of resources. This may make it difficult for households to observe value added. On the other hand, the towns we studied contain fairly standardized markets, with a clear value added measure and few other constraints on choice, such as cost or distance. This suggests that the market mechanism may work even less well in other settings.

Finally, we note issues for further research. First, one of our robust findings is that households attach great weight to their top school choices—it is difficult to influence their decisions on these. This might generalize to other settings where, at least anecdotally, households tend to prioritize "favorite" schools (e.g., those used by previous generations of the family). Such behavior could reflect aspects related to costly attention (Arteaga et al. 2021). Second, the effects of information on value added might be larger and of a general equilibrium nature if information can be delivered in greater doses and in a more sustained fashion than we did. Third, our results leave open questions on whether information interventions change only students' information sets, as opposed to affecting their preferences; these might have different implications in terms of wellbeing and schooling outcomes.

REFERENCES

Abdulkadiroglu, Atila, Joshua Angrist, and Parag Pathak. 2014. "The Elite Illusion: Achievement Effects at Boston and New York Exam Schools." *Econometrica*, 82(1): 137–196.

Abdulkadiroglu, Atila, Parag A. Pathak, Jonathan Schellenberg, and Christopher R. Walters. 2020. "Do Parents Value School Effectiveness?" *American Economic Review*, 110(5): 1502–39.

Abdulkadiroglu, Attila, Parag Pathak, and Christopher R. Walters. 2018. "Free to choose: Can school choice reduce student achievement?" *American Economic Journal: Applied Economics*, 10(1): 175–206.

Abdulkadroglu, Atila, Joshua D. Angrist, Yusuke Narita, Parag A. Pathak, and Roman A. Zarate. 2017. "Regression Discontinuity in Serial Dictatorship: Achievement Effects at Chicago's Exam Schools." *American Economic Review*, 107(5): 240–45.

Ainsworth, Robert, Rajeev Dehejia, Cristian Pop-Eleches, and Miguel Urquiola. 2019. "How parents choose schools: Evidence from the high-school admissions process in Romania." AEA RCT Registry. https://doi.org/10.1257/rct.4496.

Ainsworth, Robert, Rajeev Dehejia, Cristian Pop-Eleches, and Miguel Urquiola. 2022. "Replication data for: Why do households leave school value added on the table? The roles of information and preferences." *American Economic Association.* Inter-university Consortium for Political and Social Research. https://doi.org/10.3886/E181263V1.

Ajayi, Kehinde F., Wila H. Friedman, and Adrienne M. Lucas. 2017. "The Importance of Information Targeting for School Choice." *American Economic Review: Papers and Proceedings*, 107(5): 638–643.

Allende, Claudia, Franciso Gallego, and Christopher Neilson. 2019. "Approximating the Equilibrium Effects of Information Interventions." Princeton University Mimeo.

Andrabi, Tahir, Jishnu Das, and Asim Khwaja. 2017. "Report Cards: The Impact of Providing School and Child Test Scores on Educational Markets." *American Economic Review*, 107(6): 1535–15635.

Angrist, Joshua D., Peter D. Hull, Parag A. Pathak, and Christopher R. Walters. 2017. "Leveraging Lotteries for School Value-Added: Testing and Estimation." *The Quarterly Journal of Economics*, 132(2): 871–919.

Arteaga, Felipe, Adam J. Kapor, Christopher A. Neilson, and Seth D. Zimmerman. 2021. "Smart Matching Platforms and Heterogeneous Beliefs in Centralized School Choice." Princeton University Working Paper.

Artemov, Georgy, Yeon-Koo Che, and Yinghua He. 2020. "Strategic 'Mistakes': Implications for Market Design Research." Working Paper.

Athey, Susan, Rina Friedberg, Julie Tibshirani, and Stefan Wager. 2019. "Local linear forests." Stanford University.

Bau, Natalie. 2022. "Estimating An Equilibrium Model of Horizontal Competition in Education." *Journal of Political Economy*, 130(7): 1717–1764.

Bergman, Peter, Eric Chan, and Adam Kapor. 2020. "Housing Search Frictions: Evidence from Detailed Search Data and a Field Experiment." Columbia University Working Paper.

Bergman, Peter, Raj Chetty, Stefanie DeLuca, Nathaniel Hendren, Lawrence F Katz, and Christopher Palmer. 2019. "Creating Moves to Opportunity: Experimental Evidence on Barriers to Neighborhood Choice." National Bureau of Economic Research Working Paper 26164.

Bettinger, Eric, Michael Kremer, Maurice Kugler, Carlos Medina, Christian Posso, and Juan Esteban Saavedra. 2017. "Can educational voucher programs pay for themselves?" Harvard University Mimeo.

Beuermann, Diether, C. Kirabo Jackson, Laia Navarro-Sola, and Francisco Pardo. 2019. "What is a Good School, and Can Parents Tell? Evidence on the Multidimensionality of School Output." Northwestern University.

Beuermann, Diether W., and C. Kirabo Jackson. 2020. "The Short and Long-Run Effects of Attending The Schools that Parents Prefer." *Journal of Human Resources.*

Borcan, Oana, Mikael Lindahl, and Andreea Mitrut. 2017. "Fighting Corruption in Education: What Works and Who Benefits?" *American Economic Journal: Economic Policy*, 9(1): 180–209.

Burgess, S., E. Greaves, A. Vignoles, and D. Wilson. 2015. "What Parents Want: School Preferences and School Choice." *Economic Journal*, 125(587): 1262–1289.

Carneiro, Pedro, Jishnu Das, and Hugo Reis. 2022. "The Value of Private Schools: Evidence from Pakistan." RISE Working Paper No. 22/091 Report.

Chade, Hector, and Lones Smith. 2006. "Simultaneous Search." *Econometrica*, 74(5): 1293–1307.

Chetty, Raj, John N Friedman, Emmanuel Saez, Nicholas Turner, and Danny Yagan. 2020. "Income Segregation and Intergenerational Mobility Across Colleges in the United States*." *The Quarterly Journal of Economics*, 135(3): 1567–1633.

Corcoran, Sean P., Jennifer L. Jennings, Sarah R. Cohodes, and Carolyn Sattin-Bajaj. 2018. "Leveling the Playing Field for High School Choice: Results from a Field Experiment of Informational Interventions." National Bureau of Economic Research Working Paper No. 24471.

Dobbie, Will, and Jr. Fryer, Roland G. 2014. "The Impact of Attending a School with High-Achieving Peers: Evidence from the New York City Exam Schools." *American Economic Journal: Applied Economics*, 6(3): 58–75.

Fack, Gabrielle, Julien Grenet, and Yinghua He. 2019. "Beyond Truth-Telling: Preference Estimation with Centralized School Choice and College Admissions." *American Economic Review*, 109(4): 1486–1529.

Friedman, Milton. 1955. "The Role of Government in Education." In *Economics and the Public Interest.* , ed. Robert Solow. NJ:Trustees of Rutgers College.

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, Christopher A. Neilson, and Seth D. Zimmerman. 2018. "The Effects of Earnings Disclosure on College Enrollment Decisions." National Bureau of Economic Research Working Paper No. 21300 Mimeo.

Hastings, Justine, Thomas Kane, and Douglas Staiger. 2005. "Parental Preferences and School Competition: Evidence from a Public School Choice Program." National Bureau of Economic Research Working Paper No. 11805 Mimeo.

Imberman, Scott, and Michael Lovenheim. 2016. "Does the market value value-added? Evidence from housing prices after public release of school and teacher value added." *Journal of Urban Economics*, 91: 104–121.

Kapor, Adam J., Christopher A. Neilson, and Seth D. Zimmerman. 2020. "Heterogeneous Beliefs and School Choice Mechanisms." *American Economic Review*, 110(5): 1274–1315.

MacLeod, W. Bentley, and Miguel Urquiola. 2019. "Is education consumption or investment? Implications for the effects of school competition." *Annual Review of Economics*, 11: 563–589.

Ministry of Education. 2014. "Data on the Romanian baccalaureate exam." Government of Romania. http://static.bacalaureat.edu.ro/.

Ministry of Education. 2019. "Data on the Romanian high school allocation." Government of Romania. http://static.admitere.edu.ro/.

Mizala, Alejandra, and Miguel Urquiola. 2013. "Parental choice and school markets: The impact of information on school effectiveness." *Journal of Development Economics*, 103: 313–335.

Muralidharan, Karthik, and Venkatesh Sundararaman. 2015. "The Aggregate Effect of School Choice: Evidence from a Two-stage Experiment in India." *The Quarterly Journal of Economics*, 130(3): 1011–1066.

Riehl, Evan, Juan Saavedra, and Miguel Urquiola. 2019. "Learning and earning: An approximation to college value added in two dimensions." In *Productivity in Higher Education.* , ed. Caroline M. Hoxby and Kevin Stange. Chicago, IL:The University of Chicago Press.

Rothstein, Jesse. 2006. "Good Principals or Good Peers: Parental Valuation of School Characteristics, Tiebout Equilibrium, and the Incentive Effects of Competition Among Jurisdictions." *American Economic Review*, 96(4): 1333–1350.

Rothstein, Jesse. 2010. "Teacher Quality in Educational Production: Tracking, Decay, and Student Achievement." *The Quarterly Journal of Economics*, 125(1): 175–214.