Housing Search Frictions: Evidence from Detailed Search Data and a Field Experiment

Peter Bergman Eric Chan Adam Kapor*

December 12, 2019

Abstract

This paper shows that imperfect information about school quality causes low-income families to live in neighborhoods with lower-performing, more segregated schools. We randomized the addition of school quality information onto a nationwide provider of online housing listings for families with housing vouchers. We find that this information causes families to move to areas with higher-performing, more racially disparate schools. To understand the value of this information and its implications for models of neighborhood choice, we develop a dynamic model of households' search and location choices that incorporates subjective beliefs and imperfect information about school quality. We use the model to estimate how much families value school quality and how much they would appear to value it if we ignored information frictions. If we had ignored information frictions, control group families would appear to value school quality relative to their commute downtown by less than half as much as treatment group families. Moreover, we show that control-group households have biased beliefs that underestimate school quality and mispredict it as a function of other neighborhood characteristics.

*Bergman: Teachers College, Columbia University, bergman@tc.columbia.edu; Chan: Babson College, echan1@babson.edu; Kapor: Princeton University and NBER, akapor@princeton.edu. The authors would like to thank Dan Ackerberg, Magdalena Bennett, Eric Chyn, Tom Vogl, and seminar and conference participants at HUD, UT Austin, ASU, UVA, NWU, Princeton, Harvard (Opportunity Insights), NBER Education, Brookings, and AEFP for useful feedback. We are especially grateful to Greg Sullivan and Mirav Rathod from GoSection8, Lynn Rodgers and Veronica Helms from HUD, and Sam Olivieri and Weezie Hough from GreatSchools for data provision and assistance. This project was funded by the Walton Family Foundation, the Laura and John Arnold Foundation and J-PAL. The experimental portion of this study was pre-registered on OSF, which can be accessed here. All errors are our own.

1 Introduction

School quality varies substantially both within and across neighborhoods. This geographic variation matters because high-quality schools improve children's long-run outcomes, and, for most of the United States, residential choice is school choice: where you live determines where your child goes to school (Reardon, 2018; Deming et al., 2014; Angrist et al., 2013; Chetty et al., 2011, 2014; Dobbie and Fryer, 2015). Nevertheless, low-income families are disproportionately segregated into neighborhoods zoned to low-performing schools. This pattern of segregation is particularly true for the recipients of Housing Choice Vouchers (Section 8), a 20 billion dollar federal program that helps low-income families pay for their rent (Ellen et al., 2016; HUD, 2011). While low-income families may prefer neighborhoods for reasons other than schools, such as proximity to their family or employment, there may be frictions that impede their choices as well.

This paper tests one such friction: whether families lack information about school quality at the time of their housing search. We show how providing school-quality information to families affects their search for homes and where they choose to live. We then show how this friction matters for estimating valuations of neighborhood and school characteristics, and what our results imply for families' beliefs about school quality.

We conducted a randomized controlled trial adding school quality information onto the largest provider of online housing listings for families with housing vouchers. Using data and variation from this experiment, we estimate a dynamic model of households' search and location choices that incorporates subjective beliefs and imperfect information about school quality. We use the model to estimate how much families value school quality, and how much they would appear to value it if we ignored imperfect information. We then estimate what beliefs about the distribution of school quality, conditional on neighborhood characteristics, can rationalize the choices of households who do not see the additional school quality information.

To implement our research design, we assembled a unique dataset on low-income households' search behavior. We partnered with GoSection8, which provides online housing listings for housing voucher recipients across the country, to add school-quality information onto their listings for a random sample of users and then tracked detailed information on users' search behavior. We obtained nationwide data on school proficiency ratings from a non-profit organization. We also used a nationwide dataset on school attendance zones to assign units to their neighborhood schools. To estimate the impact of this information on users' search behaviors and residential choice, we then merged individual-level data from the listings provider with the universe of residential data on voucher recipients from the Department of Housing and Urban Development (HUD).

We find that this information problem causes low-income families to live in neighborhoods with lower-performing, more segregated schools. Providing information about zoned schools and their performance affects households' search behavior and causes them to move to locations with higher-performing zoned schools. The additional information increases the number of inquiries families made to landlords of units assigned to areas with higher-performing schools. Using the post-intervention location data from HUD, we find that families in the treatment group live in areas assigned to schools that have 0.10 standard deviation higher ratings—equivalent to roughly 1.5 percentage point higher share proficient on state exams—than schools assigned to the locations of control group families.

The potential effects of information depend on households' informed preferences and beliefs as well as the supply of desirable apartment units. The model allows us to measure demand and information while holding supply fixed, and provides context for the magnitude of the information effects. We estimate that families in the treatment group would trade an additional 52 minutes of commute to downtown for a 10 percentile point increase in school quality, holding other neighborhood characteristics constant. If we naively ignored these information frictions, we would understate this valuation by more than 50%; families would trade 27 minutes of commute time for a 10 percentile point increase in school quality.

The experimental variation also allows us to interpret families' prior beliefs about the distribution of quality and neighborhood amenities. Intuitively, a control-group household may appear to value an amenity such as (low) poverty rates both because of a direct preference for this amenity and because they believe that it predicts a high level of school quality. If the household's belief is correct, then our treatment will inform them only about the residual component of school quality that is orthogonal to the observables used to form the prediction. In contrast, if a control-group household overestimates the extent to which low poverty predicts good schools, then our treatment will cause them to be systematically negatively "surprised" about units in low-poverty areas; that is, our treatment will weaken their apparent preference for low-poverty neighborhoods.

We find that families do not use observable neighborhood characteristics well to predict school quality. We test the hypothesis that households' beliefs coincide with "rational expectations" beliefs, which we estimate using the housing data and OLS regression. We conduct a counterfactual in which households who are not treated with information continue to observe school quality with noise, but form posterior beliefs according to Bayes' rule and the true prior distribution, rather than according to a misspecified prior. We find that, if households were misinformed only in the sense of having unbiased noise around the true measure of quality, our information intervention would not cause families to live in neighborhoods with higher-performing schools. This result suggests that allowing for departures from rational expectations based on neighborhood characteristics may be important in feature to incorporate into models neighborhood choice for low-income families.

Previous research has shown that housing vouchers can induce moves to lower-poverty neighborhoods, which improves the mental and physical health of adults (Katz et al., 2001; Kling et al., 2007; Clampet-Lundquist and Massey, 2008; Ludwig et al., 2013). These moves to low-poverty neighborhoods also improve the long-run outcomes of children who move at young ages (Chetty et al., 2016; Chyn, 2018). However, voucher-induced moves result in substantially smaller changes in school characteristics than in neighborhood characteristics (Jacob, 2004; Sanbonmatsu et al., 2011; Gennetian et al., 2012; Jacob et al., 2014). More recently, Bergman et al. (2019) show how barriers impede moves to neighborhoods that promote upward mobility. Our research explicitly incorporates imperfect information into a model of residential choice to examine the extent to which this friction can explain why families move to neighborhoods with low-performing schools.

Several studies use data from centralized allocation mechanisms to estimate families' demand for schools as well as housing characteristics. The former typically find that distance to the school and its racial composition are strong determinants of families' choices (Hastings et al., 2009; Glazerman and Dotter, 2017). Families also adjust their listed preferences for schools when information is shown to them about test scores (Weinstein and Hastings, 2008; Corcoran et al., 2018; Allende et al., 2018), though families either cannot infer or do not respond to school value added (Abdulkadiroglu et al., 2017). Waldinger (2018) and van Dijk (2019) use observational data from centralized assignment mechanisms to estimate the demand for housing characteristics in low-income populations. We contribute to this research by testing how low-income families respond to information in settings where schooling decisions and residential decisions are simultaneously determined, and incorporate imperfect information into a model of residential choice.

Lastly, several papers show that test scores are capitalized into housing prices (Black, 1999; Figlio and Lucas, 2004; Bayer et al., 2007). While in recent years the test scores of schools are shown prominently next to housing listings on all of the major real-estate websites, this information is generally *not* provided to low-income families and voucher holders focused on the rental markets. Prior to our study, the two largest providers of housing listings did not provide any information on school quality.[1] As a result of our work, school-quality information is now provided to low-income families nationwide across the GoSection8 platform.

The rest of the paper is organized as follows. Section 2 describes the setting and relevant background on housing choice vouchers, study partners, the intervention, and subject recruitment. Section 3 reviews the data, and Section 4 presents the results of the experiment. Section 5 and Section 6 describes the model and estimation, respectively. The model results are in Section 7. We conclude in Section 8.

[1]These two providers are GoSection8 and Socialserve.

2 Background and Experimental Design

2.1 Housing Choice Voucher Program

The Housing Choice Voucher (HCV) program, managed by the Department of Housing and Urban Development (HUD) program, supports 2.2 million families each year by providing a voucher to help families pay for their rent in the private market. Vouchers are administered locally by Public Housing Authorities (PHAs).

Eligibility for the program is determined by total annual gross income, family size, and citizenship.[2] PHAs are required to provide 75 percent of vouchers to very low income families, defined as families whose income does not exceed 30 percent of the area's median income. In practice, almost all families fall below this cutoff. The program is not an entitlement, so even eligible applicants are often placed on a waiting list. The length of time spent on the wait lists ranges widely between PHAs.

Voucher holders are generally required to pay 30% of their adjusted gross income each month for rent and utilities.[3] The housing subsidy then covers the remainder of the rent up to a maximum amount which is a function of the Voucher Payment Standard.[4] An important implication of this rule is that the typical voucher holder faces little price variation within the set of units that will accept a voucher.

Qualified families who receive a voucher are then free to search for and choose any housing that meets the requirements of the program as long as the voucher is accepted by the property owner as rent. Each month, the housing subsidy portion is paid to the landlord by the PHA on behalf of the family. The family is responsible for the different between the rent charged by the landlord and the amount subsidized by the program.

[2] Some categories of non-citizens with legal immigration status may also be eligible.

[3] It is possible to lease a unit over the payment standard, but the family is required to pay more than 30%. The maximum a family can pay is up to 40% of their adjusted gross income, which is usually only allowed in the first year of lease up. We find that leases in our data rarely exceed 30%.

[4] This payment standard is a function of the Fair Market Rent, which is measure of area rental prices constructed by HUD.

2.2 Study Partners

The implementation of the intervention relied on two study partners: GreatSchools and GoSection8. GreatSchools is a non-profit organization that provides free and accessible school quality ratings to families via a web-based platform on www.greatschools.org. The site provides school information, test score information, and school-level ratings for over 200,000 public, charter, and private PK-12 schools nationwide. The school-level rating, which we use as our measure of "school quality" for the treatment, aims to help families compare schools within states. Ratings are whole numbers from one to ten and, at the time of implementation, were constructed by average the proficiency rates across subjects and grade levels within a schools, and then converting this index into deciles. GreatSchools labeled these deciles as below average [1-3], average [4-7], and above average [8-10] and color-coded these labels as green for above average scores, orange-yellow for average scores, and red for below average scores. This measure is used by most large-market share real-estate websites, such as *Zillow*, *Trulia*, and *Redfin*.

GoSection8.com (GS8) is the largest online platform for rental listings in the Housing Choice Voucher housing market in the United States. Their website provides a database for low-income families, both with and without vouchers, to locate and compare rental listings. There is no charge to families for this web-based service. GS8 partners with housing authorities directly to host listings for their areas, and both landlords and housing authorities can post listings. In this manner, the GS8 database has become a primary source of rental listings for many local housing authorities. At the time of the study in, 2015, the site received approximately 400,000 unique visitors per month and registers 11,000-13,000 prospective tenants per month

GS8's interface is similar to that of Zillow or other housing search platforms. Upon arriving to the site, a potential tenant logs in and types in a search query, which can be a county, city, or zip code. This query results in a set of listing titles shown to the user. The listing results displays the address, rent and the number of bedrooms and baths for each of a group of listings. Both treatment and control group users would observe the same information at this point in their search. Clicking on a listing brings up a detailed listing page, which displays a photo (if available) and characteristics of the unit and neighborhood. These characteristics include location, security deposit, and whether a voucher is required. In addition, the detailed listing page contains contact information of the landlord, as well as a text box in which to submit a "direct inquiry" to the landlord via the website. This implies that users can contact landlords by calling them directly or by using the GS8 direct inquiry button on the unit's listing page. As we describe below, this detailed listings page is what differs between the treatment and control group. We present screen shots of the website through each stage of the subject recruitment process as well.

2.3 Intervention

A registered user on GS8 was randomized into either the treatment group or the control group. Those randomized into treatment saw the GreatSchools Rating module within each property's detailed listing profile (Figures 1 and 2). The module observed only by the treatment group consists of a school quality ratings legend, a map showing the ratings of the assigned schools to the property being viewed, and a listing of the schools shown on the map.[5] The school listing below the map presents the rating, the school name, school type (all are traditional public), grades served, distance from the property, and a check mark if the school is the assigned school for that particular property. Ratings are displayed within a colored circle, as described before.[6] The treatment is only preceded by the general property details and description first seen when the user clicks on a listing resulting from their search query (Figure 1).[7]

Both treatment and control group users can view property information other than the GreatSchools Ratings module (which is the same as what a user who refuses participation in the study views). This information consists of a property photograph, property details, a narrative property description, a Walk-

[5] If an assigned school is relevant to a unit's location because it is in an area with centralized school choice, users are shown the closest elementary, middle, or high school.

[6] Clicking on a school 'details' link takes the user to a web page on GreatSchools.org, where they can view additional information for the specific school.

[7] Since our study, GoSection8 has since scaled the intervention across their site and updated the intervention to make it easier to find listings in areas with better schools, including a text message alert system, a link to find more listings in a given school zone, and the ability to sort listings by school ratings.

Score section, utilities payment information, additional amenities and appliance details, and handicap accessibility features of the property. The property details also include rental price, number of beds and baths, and an occasional short description written by a landlord.

2.4 Subject Recruitment

Implementation of the intervention began in late May 2015 and ran through February 2017. After visiting the GS8 landing page (Figure 3) and entering a location into the search box, users saw a pop-up asking whether they would like to sign up for the study (Figure 4). If a user selected "no," they were directed to the website as a non-participant of the study and would not see school-quality information. If a user instead selected "yes," they were taken to a page with the consent agreement and were asked to fill out an initial survey (Figure 5). Upon completion of the survey, the user is registered both as a tenant registrant on the website and as a participant of our study. Randomization into treatment and control was then automated by the website at the moment of survey completion, and any subsequent browsing, searching, and page-viewing were tracked via IP addresses and cookies. Treatment assignment was stratified by voucher status.

We recruited 5,743 participants. Of these, 3,012 self-reported that they hold a voucher. Of the voucher holding participants, we were able to match 1,969 (66%) heads of households with stated vouchers to HUD data. There is no significant difference in the treatment status between matched and unmatched users.[8] There are various reasons why not all families who reported having a voucher were matched to HUD. First, HUD only has data on households that lease up. National lease-up rates conditional on voucher issuance are approximately 69%, comparable to our match rate (Finkel and Buron, 2001). Second, there may also be self-reporting error regarding first and last names or date of birth. We cannot, unfortunately, discern our consent rate because GS8 did not track a denominator to compute this number. We can, however, compare our study sample to a representative sample from the universe of voucher holders.

Table 1 shows this comparison. Study participants are much less likely to be classified as disabled

[8] See Table 3, which we discuss further below.

and have more total household members, but total income differences are small and not insignificant. Demographic differences, while statistically significant given the large sample from HUD, are all less than 5 percentage points. Table 2 shows how the school ratings associated with the study participants' neighborhoods compares to that of the overall HUD population, GS8 website listings, and US elementary schools. Study participants' school ratings are low, and minority and free-reduced-priced lunch shares are high, compared to the overall HUD population. Ratings for GS8 listings are comparable to overall elementary schools, though these listings tend to be much higher minority share and higher share free-reduced-priced lunch than the average elementary school.

3 Data

In this section, we describe the baseline and outcome data, which are primarily from two sources: GoSection8 and HUD. A third data source, GreatSchools, provided some of the baseline and school characteristics and ratings used for the intervention.

3.1 GoSection8 Data

From GoSection8, we obtained data on users who consented to be in the study and property information. Tenant registration collected information on treatment status, initial sign up date, and results from the baseline survey. The baseline survey, which can be seen in Figure 5, collected the registrant's name, contact information, head of household's name and date of birth, whether the family has children and their age ranges, intended move date, primary reason for moving, and whether they have a voucher and the associated housing agency. We used name, date of birth, and public housing authority (PHA) in the baseline survey to link to HUD administrative data, which is described further below.

In addition to tenant registration and survey data, we also received a list of properties viewed and direct inquiries (i.e. contacts made via the website to a landlord) by each study participant. With these data, we observe the set of properties a participant viewed, along with property descriptions, rental prices, location, and other characteristics. Direct inquiries are online inquiries sent from prospective tenants to a landlord via a "contact the landlord" link on shown on the right hand side of every GS8 rental listing. It is one method for tenants to show interest in a property, however, it is not the only method. Prospective tenants can also directly call or email the landlord, provided that either or both is available. Inquiry data is useful because tenants cannot view school quality data prior to viewing the detailed property listing.

3.2 GreatSchools Data

GreatSchools provided school-quality ratings, geocoded school locations, and school attendance zones.[9] As of writing, ratings are based on multiple measures of academic quality (student achievement as measured by performance on standardized tests, student progress as measured by year-over-year test score gains, college readiness, etc.), however, at the time of the intervention, ratings were based on each state's test-score proficiency rates, as described above.

The distribution of the universe of school ratings is shown in Figure 6 for elementary, Figure 7 for middle, and Figure 8 for high schools. Ratings tend to peak at six or seven and are approximately normally distributed around a mean of 5.5 and standard deviation of 2.5. However, the highest ratings on average are from elementary schools, and the lowest are from high schools, which also have the lowest share of schools rated above average.

Using clear bars with red outlines, Figures 6-8 also show the distribution of ratings for properties listed on GoSection8. These properties, which reflect rental supply available to voucher holders, have significantly lower ratings, on average, compared to the general population of schools. Further, choice is restricted by the low variation in ratings within metropolitan areas. For example, 88% of the top 100 cities by number of GS8 rental property listings has a standard deviation in elementary-school ratings of less than 2.5, which is the standard deviation of the universe of schools. The standard deviation of GS8 listings is, on average, less than 2.0.[10].

[9] School attendance zones were purchased from Maponics.

[10] As an example, we show elementary school rating standard deviations by county, city, district, and zip code in figure

3.3 HUD Data

HUD provided quarterly data on the universe of voucher recipients, which includes household member's demographics and income as well as information about where they live: the bedrooms, rent, and geocoded locations.[11] We linked these data to GoSection8 tenant participant data based on first name, middle name, and the date of birth of the head of household.[12]

As part of the study on GoSection8, we recruited 2,968 eligible participants who reported that their household had a voucher. Of these participants, we were able to match 1,965 heads of households (66%) to HUD data based on name and date of birth using data from 2013 through 2018. Our HUD data span roughly seven quarters after the end of the experiment.

We show baseline sample characteristics and treatment-control group balance in Table 3 by regressing an indicator for treatment assignment on baseline covariates. The regression to check for balance is as follows:

$$\mathrm{X}_i = \gamma_0 + \gamma_1 \mathrm{Treatment}_i + \eta_i,$$

where X is the baseline variable of interest for family i and γ is the coefficient of interest for the treatment-control differences. Here, we find that treatment and control groups are generally similar in terms of baseline characteristics. There are some significant differences for the head of household being Hispanic and the probability of being issued a voucher from a housing authority with Moving to Work (MTW) designation (this designation means that housing authority has greater autonomy over how it allocates its resources). However, an omnibus test for treatment-control differences is not significant (p-value=0.19).

As part of our baseline survey, we also asked participants the primary reason for their move. Thirty-two

?? and **??** The first figure includes duplicated schools, which shows that when schools aligned to all properties are included, there is often very little variation in school quality within a geographic area. The second figure shows unique school instances, though this is less representative of the school quality as faced by GoSection8 users.

[11] The data provided to us are from users who formally filed a Form 50058 with their PHA, which is required annually and when any transaction, such as a new move, occurs.

[12] We merged data from GoSection8 to HUD based on fuzzy merging Stata package, dtalink.

percent of registrants noted that the primary reason for moving is to "get better schools for my children." This share is high relative to prior results from the Moving to Opportunity study (Kling et al., 2007), but this result is in line with more recent evidence from Bergman et al. (2019), which also finds that one third of families' primary reason for moving is for better schools. Also worth noting, more than 75% of the sample intends to move within three months. This intention is in line with the time families have to use a voucher before it expires.

Empirical Strategy

To measure the effect of the treatment on various outcomes, we estimate the following:

$$Y_i = \beta_0 + \beta_1 * Treatment_i + \varepsilon_i$$

where Y is the outcome of interest for family i, and β is the coefficient of interest, on the treatment variable. Our primary outcome is the average rating of the elementary, middle and high school of the assigned schools to a given property. We also look at the associated school demographic characteristics, and the neighborhood and property characteristics as well, such as tract poverty rates, demographics, commute to downtown, and walk score. The walk score is a proprietary measure from Walkscore.com that assesses the "walkability" of an area based on a neighborhood's access to public transportation, friendliness to biking, and local amenities. In the following section, we provide the estimated impacts of the intervention.

4 Results

4.1 Impacts on Search Behaviors

The upper panel of Figure 9 shows the average number of properties families view each week relative to their intended move date, which is normalized to 0 on the X-axis. Weeks prior to the intended move date are therefore shown as negative and weeks after are positive on the X-axis. For both the treatment and control group, there is a rise in the number of properties viewed as the intended move date gets closer, which peaks around an average of two views in one week. These averages include users who did not view any properties in a given week. After the intended move date, the rate of additional properties viewed rapidly declines. There is no significant difference in the patterns across group, including in the overall number of views (we show this formally in Table 4 discussed below). These results provides evidence that the intervention did not change how often families used the website, either in terms of whether a family made any views or how many views.

At the outset, it is less likely we should observe differences in the quality of views because families do not observe school quality before clicking to view a property. However, it is possible they learn and infer quality from neighborhood characteristics. Nonetheless, we do not observe significant differences in the school quality of listings viewed by families. Panel (a) of Figure 10 shows the average number of views a treatment group user makes broken out by the average school quality (averaged across elementary, middle, and high school). Most views have average school quality in the 1 to 4 range, and there are very few views made to listings associated with schools rated 6 and above, which is likely because the supply of such listings is low.

Table 4 summarizes these and other search-related regression results. Families make, on average, 34 views. Roughly 80% of families make at least one view and view properties in 14 different census tracts. There are no significant effects on these outcomes. We also assess the effects on the number of views for schools rated at some school quality rating or above. Here, we find almost no effects with the exception of number of views rated 10. Given the small number of schools rated 10—the average control user only saw 0.07 listings with a 10-rated school—this finding just be due to chance.

After a family views a unit they may choose to make an inquiry to the landlord about that unit either by phone or via a link on the website. GoSection8 recorded whether a family sends an inquiry to the landlord through their platform, though we cannot observe whether families called landlords. The lower panel of Figure 9 plots the average number of inquiries made each week relative to families' intended move date.

Again, these averages include families who made no inquiries in a given week. The pattern is similar to the pattern of views over time: inquiries increase over time with respect to the move date. In contrast to the views, however, the treatment group has a larger increase in inquiries in the week prior to when they would like to move, and slightly higher inquiries before then. As we test below, there is evidence that the treatment group makes more total inquiries than the control group.

Once a family views a unit, they also observe the quality of schools assigned to that unit. Therefore we might see a change in the school quality of the units that families inquire about. Figure 10, panel (b), shows that there is an increase in the school quality associated with the units families inquire about. Most of the increase inquiries is concentrated among units with average school quality ratings between 3 and 5.

Table 5 summarizes the inquiry-related regression results. The intervention increased the number of inquiries for properties with schools rated 2 or above through 4 or above. We also find that the treatment increased total inquiries by 0.80 over the control mean of 2.3, which is significant at the 10% level. The intervention did not, however, significantly increase the probability that a user inquired about any property nor the number of tracts inquired about.

4.2 Endline School Quality

The additional information causes families to move to higher-rated schools. Using address information from HUD, we assigned the zoned elementary, middle and high school, and their respective Greatschools ratings, to each of our study participants' endline location. Table 7 shows the results across several measures of quality. School quality as measured by the average elementary, middle and high school rating improved by 0.26 points, which is statistically significant. We also show the maximum or minimum school quality across the schools assigned to an address; the impact for the maximum is 0.32 and the minimum is 0.17, and both remain statistically significant. Figure 11 graphs the density of school quality associated with where families live for the treatment and control groups. There is a shift away from ratings in the 1 to 3 range and toward ratings in 5 and 6 range, similar to where the impacts fall on the quality of inquiries. The effects are slightly larger for the middle school rating than for the elementary or high school rating, though this difference is not significant. To add context to these effects, the average treatment effect is approximately 0.10 standard deviations relative to the control group. Measured in terms of the share of students marked "proficient" or above across a schools' state exams, the treatment effect is 1.5 percentage points.

The interventions also causes families to move less segregated schools. Treatment group families live in neighborhoods assigned to schools with 3 percentage points fewer black or Hispanic students and 2 percentage points fewer students receiving free or reduced-priced lunch, which is an indicator of family income. These effects likely arise because the school ratings are based on test score levels, which strongly correlate with the demographics of the students attending the school.

Lastly, we explore whether the effects are concentrated in schools that parents may expect to send their children in the future. We define a "next school" measure based on the age of the youngest child. If the child is 0-4, the "next school" is the zoned elementary school.[13] If the child is 5-10, the next school is the middle school, etc.. Table 8 shows that families with very young to elementary-school-aged children target the next school level, but this pattern disappears for the parents of children in middle or high school. The bottom of the table shows the overall effect on "next-level" school rating, which is 0.31 points. This is slightly larger than the effect on average ratings, and not nearly statistically different.

4.3 Endline Neighborhood Quality

We also examine whether the addition school quality information affects the characteristics of the neighborhoods in which families. Table 9 shows these results. In terms of neighborhood demographics, this is generally not the case. There are small differences in the racial composition of neighborhoods, none of which is statistically significant. There is a 1 percentage point decrease in the neighborhood poverty rate, which may be due to the correlation between school ratings and family income. Figure 12 shows that there

[13] We break down the child's age in terms of ranges because when asking the participant the age of their children on the baseline survey, one requirement from the website was to make it as easy as possible. We decided to construct ranges associated with the age range of school levels. For example, children tend to be in elementary school when they are 5-10 years of age, and in middle school from 11-13, and so on.

is a slight but broad-based shift leftward in the distribution of neighborhood poverty rates for treatment families. For comparison, the MTO experiment induced moves to neighborhoods with 15 percentage point lower poverty rates for families with young children (Kling et al., 2007; Chetty et al., 2016).[14] We also find no impact on the median rent in families' neighborhoods.

There are larger, statistically significant effects on the commute to downtown and the walk score of the neighborhood. The latter reflects what local amenities and public transportation are nearby. Commute to downtown increases by 2 minutes (the control-group mean is 15 minutes), which implies families are moving slightly farther out from the city center to live in areas with better schools. The walk score decreases by 3.9 points (out of 100) compared to a control-group mean of 50. Families appear to move slightly out of downtown in favor of better schools, though we cannot discern from the reduced-form results to what extent families value commute to downtown *vis-à-vis* school quality.

HUD records limited data on the characteristics of the units families live in. Table 10 shows the treatment effects on unit rent, the number of bedrooms, and beds per household member. However, the design of the voucher program implies that there is unlikely to be an affect on these variables. Most families maximize the amount of voucher support they receive, and housing authorities place a maximum on the percent of income families can pay toward rent. Nearly the entirety of our sample pays precisely 30% of their income toward rent. Moreover, families are eligible for a number of bedrooms and subsidy amount based on their household composition, and they cannot choose fewer bedrooms and put the additional funds toward a nicer unit. Thus, it is not surprising there is no impact on the rent paid and the number of bedrooms. Instead, as discussed above, families appear to trade off unit location—distance from downtown—for additional school quality.

Finally, we explore whether families appear to respond to the portion of school quality that is not correlated with the neighborhood's other characteristics. We regress the school-quality measure on neighborhood racial composition, poverty rate, median gross rent, walk score, and distance to downtown, and then regress

[14] 15 percentage points is the Intent-to-Treat Effect.

this residualized measure on the treatment indicator.

The effect on residual school quality, shown in the bottom row of table 9, is smaller but still large and statistically significant, suggesting that households are not just using school quality as a signal of the other amenities that we observe. We also show the other component of this regression of school quality on neighborhood characteristics: predicted school quality. If households had formed prior beliefs whose means coincided with OLS estimates of predicted quality, then our treatment would have informed them only about the residual component, which is mean-independent of the predicted value. Hence, if households were using these variables to formulate a "correct" prior, we should see a zero treatment effect on predicted quality. Our findings indicate that households may not be using this information to predict quality well. In what follows, we formalize this logic in our model.

5 Model

We now specify and estimate a model of households' search behavior and moving decisions, recovering households' preferences, their costs of using the platform and of inquiring about and visiting housing units, and their prior beliefs about the joint distribution of school quality and other characteristics. The model allows us to assess the importance of imperfect information for a model of neighborhood choice, and, in particular, how imperfect information affects estimated valuations of school quality *vis-à-vis* other neighborhood characteristics.

We do this as follows. First, we obtain estimates of search costs, treated households' preferences for neighborhood charactristics and school quality, and the "as-if-full-information" preferences of control households that an econometrician would estimate when not incorporating imperfect information. A direct comparison of the treated households' preference parameters to the "as-if" preferences of the control group reveals the bias in estimated preferences for neighborhood characteristics and school quality when imperfect information is ignored. Second, we use the estimated parameters to solve for families' prior beliefs about school quality under the hypothesis that they are Bayesian updaters under our treatment, and then we test whether symmetric measurement error around rational expectations of school quality based on observable neighborhood characteristics could be a viable approach to modeling imperfect information in this setting.

The model is designed to capture the following features of the market. First, if a household fails to use its voucher within a given period, the household loses the voucher. In contrast, once a household signs a lease, it commits to pay rent at that location for a year. Therefore, households in our model face a stopping problem with a finite deadline. Second, because demand exceeds supply, vacancies are short-lived. At the same time, it may take landlords a few days to respond to inquiries. Our model therefore treats search as simultaneous within a period (in our empirical specification: one week), but we assume no recall across periods.[15] Third, not all search is conducted online. Households may discover other apartments through social networks or by chance. Our model therefore allows for off-platform search as well as on-platform search.

5.1 A model of search

Time is discrete: $t = 1, \ldots, T < \infty$. Household $i \in I$ in market $m \in M$ becomes active in period 1 (i.e. the head of household receives a new voucher, or decides to move), and has until time T to find an apartment. The set of available apartments in market m is denoted $\mathcal{J}_m$. If i moves to an apartment $j \in \mathcal{J}_m$ before time T, it receives a payoff of

$$u_{ij} = x_j \beta^x + q_j \beta^q + \epsilon^0_{ij} + \epsilon^1_{ij}$$

where x_j is a vector of neighborhood characteristics of unit j, q_j is school quality, and ϵ^0_{ij} and ϵ^1_{ij} are idiosyncratic preference shocks. There is no discounting.[16] However, if the household fails to accept an apartment by time T, it receives an outside-option payoff which is normalized to zero.

Within each period, each active household makes three choices: whether to use the online platform;

[15] A potential tenant may submit multiple inquiries in one session, then wait a few days for a response, but must then decide quickly. We conducted an audit study of landlords' responses to inquiries; we found that 41 out of 100 landlords responded to online inquiries, typically with a lag of a day or two.

[16] An interpretation is that the household is searching for a lease which begins at time T.

which subset of units to inquire about and/or visit; and whether to accept a unit. We now discuss these in detail.

At the beginning of period t, household i observes a cost draw

$$c_{it}^{search} \sim F_{c^{search}}(\cdot),$$

i.i.d. across periods. i then chooses whether to pay this cost in order to use the search platform. If so, it discovers a set of on- and off-platform units. If not, it discovers only off-platform units. On-platform units $J_{it}^{on} \subset \mathcal{J}_m$ are drawn at random from a distribution over sets of apartments,

$$F_{go8}(\{x_{ij}, q_{ij}, \epsilon_{ij}^0, \epsilon_{ij}^1\}_{j \in J_{it}^{on}}; m),$$

conditional on searching. If i does not use the platform in period t, then $J_{it}^{on} = \emptyset$. Off-platform units J_{it}^{off} are drawn from an analogous distribution,

$$F_{off}(\{x_{ij}, q_{ij}, \epsilon_{ij}^0, \epsilon_{ij}^1\}_{j \in J_{it}^{off}}).$$

Let $J_{it} = J_{it}^{on} \cup J_{it}^{off}$ denote the set of units that i discovers in period t.

Upon discovering a unit, a household observes characteristics x_j, preference shock component ϵ_{ij}^0, and a signal of school quality, $\hat{q}_{ij}$. In order to observe observe the second preference ϵ_{ij}^1, the household must pay an additional cost $c_{it}^{inquire}$ *per unit* that it wishes to learn about. We model this second decision as a simultaneous search problem. After observing the set of units, the household learns its inquiry cost draw,

$$c_{it}^{inquire} \sim F_{c^{inquire}}(\cdot),$$

drawn i.i.d. across periods, and chooses a set of units to submit inquiries to,

$$J_{it}^{inquiry} \subseteq J_{it}$$

at cost $|J_{it}^{inquiry}|c_{it}^{inquire}$.

Finally, the household observes ϵ_{ij}^1 for each unit that it inquired about, then chooses whether to accept a unit. If the household accepts unit j it receives u_{ij}. Otherwise, it continues to the next period. Households may not accept units that they have not inquired about.

If household i is not part of the treatment group, it observes a noisy signal of quality, $\hat{q}_{ij} \sim G(\hat{q}_{ij}|q_{ij})$. If i is treated with information, then $\hat{q}_{ij} = q_{ij}$ almost surely.[17] Households are risk-neutral, and form expectations $E(u_{ij}|\hat{q}_{ij}; x_{ij}, G_{q|x}(\cdot))$ where $G_{q|x}(q|x)$ is a belief over the conditional distribution of q_{ij} given characteristics x_{ij}.

$G_{q|x}(q|x)$ arises from an information structure in which households observe the sum of the true quality q and additive measurement error η. Households assume that true quality and measurement error are jointly normal conditional on x. In particular, for prior-mean parameters γ and variance parameters σ_η^2 and σ_q^2, for untreated households, the signal distribution is given by

$$\begin{pmatrix} q \\ q+\eta \end{pmatrix} |x \sim N\left(\begin{pmatrix} x'\gamma \\ x'\gamma \end{pmatrix}, \begin{pmatrix} \sigma_q^2 & \\ \sigma_q^2 & \sigma_q^2+\sigma_\eta^2 \end{pmatrix}\right).$$

Importantly, the prior mean belief $x'\gamma$ need not coincide with the "rational expectations" prior that the econometrician would estimate from data on all listed units. Households who are Bayesian updaters will shrink the signal that they observe toward this (potentially misspecified) subjective mean.

In our empirical specification, preference terms x_j include a constant, ACS demographics (median gross rent, poverty rate, percent Black households), walk score, distance to downtown, and school quality.

[17] We have estimated specifications in which treated households may have imperfect signals of quality for off-platform listings, $\hat{q}_{ij} \sim G_{\text{off}}(\hat{q}_{ij}|q_{ij})$. We find that their information is close to perfect for these listings.

Households may also use nonschool characteristics to form beliefs about school quality. Here, median rent proxies for the cost and amenities of the neighborhood.[18] Controlling for median rent absorbs the variation in school quality that is capitalized into the price of a typical housing unit in the neighborhood. The term ϵ^0_{ij} is intended to capture match-specific characteristics, such as the proximity of the unit to family or childcare providers, which households can infer from the listing, while ϵ^1_{ij} represents characteristics such as the quality and condition of the housing unit which cannot be learned without contacting the landlord or visiting the unit.

We complete the model with the following distributional assumptions on preferences, costs and listings:

$$\begin{aligned}
c^{\text{search}}_{it} &\sim LogNormal(\mu_{\text{search}}, \sigma^2_{\text{search}}) \\
c^{\text{move}}_{it} &\sim LogNormal(\mu_{\text{inquire}}, \sigma^2_{\text{inquire}}) \\
\epsilon^0_{ij} &\sim N(0, \sigma^2_\epsilon) \text{ i.i.d.} \\
\epsilon^1_{ij} &\sim \mathit{Gumbel}() \text{ i.i.d.}
\end{aligned}$$

These assumptions guarantee that costs are positive. Independence of costs across periods is not essential, and can, in principle, be relaxed.

On-platform listings are sampled from the empirical distribution of $\{J_{it}\}_{i \in I_m,\ t=1,\dots,T}$ by market. We assume off-platform units are drawn from the same distribution as on-platform units, and that the number of off-platform units drawn in a period is distributed $Poisson(\lambda)$.

5.2 Optimal play and qualitative predictions

Let V_{it} denote the expected payoff of household i if it is active at the beginning of period t. Fix $V_{i,T+1} = 0$. Let

$$\hat{u}_{ij} \equiv E(u_{ij} | \hat{q}_{ij}, x_{ij}, \epsilon^0_{ij}, \epsilon^1_{ij}; G(\cdot))$$

[18] ACS median gross rent is not the rent that the tenant would pay. Recall that tenants typically pay 30% of their income, regardless of the unit chosen.

denote i's expected utility from property j when i observes signal $\hat{q}_{ij}$ and has prior belief $G(\cdot)$. i accepts a unit with a successful inquiry if and only if $\hat{u}_{ij} > V_{i,t+1}$.

Let

$$\hat{v}_{ij} = E(x_j\beta^x + q_j\beta^q + \epsilon^0_{ij}|\hat{q}_j)$$

denote the expected value of unit j, not including the final shock ϵ^1_{ij}, when the household observes signal $\hat{q}_{ij}$. If the household is going to inquire about k units in period t, the optimal units are those with the k highest values of $\hat{v}_{ij}$. The value of this set of inquiries is given by

$$U^k_{m,t}(\theta) = V_{t+1}(\theta) + \log\left(1 + \sum_{j=1,\ldots,k} \exp(v_{ij}(\theta) - V_{t+1}(\theta))\right).$$

Because the cost is linear in k and the marginal return to an additional inquiry is decreasing, the optimal strategy is to choose the smallest k such that the marginal gain from the $k+1$th unit is smaller than $c^{inquire}_{it}$.

The payoff from on-platform search is given by

$$V^1_{it} = E\left(\max\left\{V_{i,t+1}, \max\{\hat{u}_{ij}|j \in J^{inquiry}_{it}\}\right\} - c^{inquire}_{it}|J^{inquiry}_{it}|\right),$$

where the expectation is over the sets $J_{it} = J^{off}_{it} \cup J^{on}_{it}$ and the inquiry cost $c^{inquiry}_{it}$.

Analogously, let V^0_{it} denote the expected payoff from no on-platform search in period t, i.e. when $J^{on}_{it} = \emptyset$ with probability 1.

i chooses to search in period t if

$$V^1_{it} - V^0_{it} > c^{search}_{it},$$

The value V_{it} is then given by the value of the option to search, i.e.

$$V_{it} = V^0_{it} + \left(V^1_{it} - V^0_{it} - E(c^{search}_{it}|c^{search}_{it} < V^1_{it} - V^0_{it})\right) F_{c^{search}}(V^1_{it} - V^0_{it}).$$

Intuitively, the search cost threshold is falling in time. With a long horizon, there is a good chance of low cost draws in future periods, and moreover it is more likely that a good unit will arrive via off-platform search in the future. With a short horizon, search is more valuable. Thus, conditional on being active the probability of search and number of inquiries is higher as the deadline draws closer. Close to the deadline, however, many households may no longer be active, so that peak period of search activity may be interior.

6 Identification and Estimation

We first provide an intuitive discussion of identification and sources of variation. We then discuss parametric identification of our model's belief parameters and present our estimation procedure. Our parametric assumptions are convenient, but we do not believe that normality of signals or linearity of the utility index is essential.

6.1 Identification and testable restrictions

Observe that one may estimate a full-information model, obtaining estimates of the distribution of preference and cost parameters, using only treated households. Identification of treated households' demand is standard. In particular one may restrict attention to the final period, reducing the problem to a static problem.[19] The key novelty here is the estimation of subjective prior beliefs and signal parameters, which relies on the variation induced by the randomized experiment.

Under Bayesian updating with a correctly specified prior, beliefs are Martingales. Let $v = v(x, q, \epsilon^0)$ be any random variable with finite first moment, and suppose utility is given by $u_{ij} = v_{ij} + \epsilon^1_{ij}$ with $\epsilon^1_{ij} \perp v_{ij}$. When households have a correctly specified prior belief $H(u|x, \epsilon^0)$, we have

$$E(E(v|x, \epsilon^0, q)|x, \epsilon^0) = E(E(v|x, \epsilon^0, \hat{q})|x, \epsilon^0) = E(v|x, \epsilon^0).$$

[19] Conditional on inquiry decisions, the problem is standard. There is selection into the decision to inquire about a property, but the characteristics of other units j' discovered in period t shift the threshold cost at which a household inquires about unit j.

Because the average of the agent's posterior mean utility conditional on our treatment is identical to the average of the posterior mean utility conditional on a control-group agent's information for all x, treatment effects on mean utility cannot systematically vary with x. That is, while the provision of information about q may increase demand for units with higher q conditional on x, it cannot systematically raise or lower the agent's expected utility of a unit given its information *as a function of* x.

In contrast, a nonzero treatment effect on the expected utility of units with characteristics (x', q, ϵ_0) relative to those with (x, q, ϵ_0) indicates a violation of the hypothesis of Bayesian learning with a correctly specified prior.[20] For instance, households may believe that high poverty rates in a neighborhood generally predict low school quality to a greater extent than warranted by the data. Intuitively, our treatment will then increase the expected utility that households believe they receive from high-poverty neighborhoods, and therefore the probability of choosing such neighborhoods.

We have stated this intuition in terms of expected utilities. However, mean utilities $E(v|x, q)$ are an invertible function of choice probabilities under fairly general conditions (Berry, Gandhi, Haile 2015). Thus, estimating prior beliefs about the quality of units with characteristics x relative to those with characteristics x' intuitively amounts to estimating treatment effects on the demand for such units. Under Bayesian learning with a correctly specified prior, this treatment effect should be zero.

Intuitively, we may also learn about the informativeness of control households' signals from their demand for units as a function of q. If demand is completely unresponsive to quality q conditional on other characteristics x, we may conclude that control-group households observe no additional information about q. A higher probability of choosing high-q units would indicate that control households observe a signal that conveys some information.

[20] We may integrate out ϵ_0.

6.2 Beliefs and expected utilities

We now make this argument precise in the context of our model. We begin by deriving the distribution of indirect utilities according to the econometrician. The agent's expectation is given by

$$E(q|q+\eta) = s \cdot (q+\eta) + (1-s) \cdot \mu_q,$$

where $s \equiv \frac{\sigma_q^2}{\sigma_q^2+\sigma_\eta^2}$. Dropping subscripts, expected utility given i's information is

$$\begin{aligned}
\hat{u} &= x'\beta + \left(s\hat{q} + (1-s)x'\gamma\right)\beta_q + \epsilon \\
&= x'\left((1-s)\beta_q\gamma + \beta_x\right) + (q+\eta)s\beta_q + \epsilon \\
&= x'\left((1-s)\beta_q\gamma + \beta_x\right) + qs\beta_q + (s\beta_q\eta + \epsilon) \\
&= x'\tilde{\beta}_x + q\tilde{\beta}_q + \tilde{\epsilon},
\end{aligned}$$

where

$$\begin{aligned}
\tilde{\beta}_x &= (1-s)\beta_q\gamma + \beta_x \\
\tilde{\beta}_q &= s\beta_q \\
\tilde{\epsilon} &\sim N(0, s^2\beta_q^2\sigma_\eta^2 + \sigma_\epsilon^2).
\end{aligned}$$

It follows that, for each $(\beta_x, \beta_q, \gamma, \sigma_\eta^2, s, \sigma_\epsilon^2)$, there is an observationally equivalent "full information" model, in which q is observed without error. In this model, the household acts as if it has preference parameters $\tilde{\beta}_x, \tilde{\beta}_q$, and $\tilde{\sigma}_\epsilon^2$. The cost parameters are unchanged.

$\tilde{\beta}_x$ and $\tilde{\beta}_q$ are the parameters would we would estimate if we ignored imperfect information. The model specifies the resulting bias: β_q will be shrunk by the signal to noise ratio, s, so it will appear families value school quality less than they actually do. The bias in $\tilde{\beta}_x$ arises as families use other neighborhood characteristics to predict school quality via γ. For example, say families dislike high neighborhood poverty rates, which implies that $\beta_{poverty}$ is negative. If families further believe that the poverty rate negatively correlates with school quality, the γ associated with this characteristic is negative as well. Then, ignoring imperfect information, the estimated distaste for the neighborhood poverty rate would be larger than that estimated under perfect information.

Our approach is to estimate the "full-information" preference parameters separately by treatment status, recovering the parameters $\beta_x, \beta_q, \tilde{\beta}_x, \tilde{\beta}_q, \sigma^2_\epsilon, \tilde{\sigma}^2_\epsilon, \mu_{c_{search}}, \sigma^2_{c_{search}}, \mu_{c_{inquire}}, \sigma^2_{c_{inquire}}$, and λ. These parameters are sufficient for obtaining "as if" willingness-to-travel measures $\frac{\tilde{\beta}_q}{\tilde{\beta}_{commute}}$ relating taste for quality to distaste for distance from city center.

We then solve for the values of γ, s, σ_η and subjective prior variance σ_q that uniquely rationalize the estimated parameters. In particular, given estimates $\hat{\beta}_x, \hat{\tilde{\beta}}_x, \hat{\beta}_q, \hat{\tilde{\beta}}_q, \hat{\sigma}^2_\epsilon, \hat{\tilde{\sigma}}^2_\epsilon$, we compute

$$\hat{s} = \frac{\hat{\tilde{\beta}}_q}{\hat{\beta}_q}$$

$$\hat{\gamma} = \frac{\hat{\tilde{\beta}}_x - \hat{\beta}_x}{\hat{\beta}_q - \hat{\tilde{\beta}}_q}$$

$$\hat{\sigma}^2_\eta = \frac{\hat{\tilde{\sigma}}\epsilon^2 - \hat{\sigma}^2_\epsilon}{\hat{\tilde{\beta}}^2_q}$$

$$\hat{\sigma}^2_q = \frac{\hat{\sigma}^2_\eta \hat{s}^2}{1 - \hat{s}^2}.$$

Our model has a unique γ that exactly rationalizes the estimates. When the "objective" γ can be estimated, we can test the hypothesis of Bayesian updating under maintained assumption that treatment operates only via the information channel. In particular, we test whether $\gamma = \gamma_{OLS}$, where γ_{OLS} are the coefficients that the econometrician would obtain from a regression of q on neighborhood characteristics x.

6.3 Estimation procedure

We allow for "inquiries" to take place via channels which we do not observe, such as phone calls. We estimate a parameter, $Pr(\text{observe inq.})$, which denotes the probability that an inquiry occurs via the "direct inquiry" text box. We assume that this probability is invariant to our treatment, and that inquiries by text box and by phone or other means are otherwise identical. To better fit the data, we also allow for the presence of a "passive" type, which completes our survey but never views a listing or moves. We estimate the probability, $Pr(\text{passive})$, that households are of this type.

We estimate the following parameters via the method of simulated moments: $\beta^x(\text{treat})$, $\beta^q(\text{treat})$, $\beta^x(\text{control})$, $\beta^q(\text{control})$, μ_c^{search}, σ_c^{search}, μ_c^{inquire}, $\sigma_c^{\text{inquire}}$, $\sigma_\varepsilon(\text{treat})$, $\sigma_\varepsilon(\text{control})$, $Pr(\text{observe inq.})$, λ, $Pr(\text{passive})$.

In order to restrict λ and cost parameters to be identical across treatment status, we estimate the model jointly on treatment and control groups. Doing so is not essential; an alternative procedure would involve estimating the model separately by group.

Our moments are of the form

$$g = \frac{1}{n}\sum_{i=1}^{n}\left(E(y_i^{model}(\theta)) - y_i^{data}\right) w_i$$

where w_i are indicators for treatment groups, and y are the following objects:

1. $\sum_{t=1}^{T} Pr(\text{search in period } t) * b_k(t)$, for basis vectors $b_k(\cdot)$;
2. $\sum_{t=1}^{T} Pr(\text{number of inquiries in period } t) * b_k(t)$;
3. $\sum_{t=1}^{T} Pr(\text{number of inquiries in period } t)^2$;
4. $1(\text{lease up at new unit})$;
5. mean inquiry characteristics (x, q);
6. mean leased unit characteristics (x, q).

We choose the following basis vectors in order to reduce the total number of moments needed:

$$b_k(t) = \cos(\pi k t/(T-1)) \text{ for } k \in 0, 1$$

. Thus, our moments relating to the timing of views and inquiries consist of the total number of views and inquiries in the $k = 0$ case, respectively, and decreasing functions that capture time trends for $k = 1$.

The variance moment is crucial for estimating the distribution of inquiry costs. Intuitively, when the variance of costs is high, so is the variance of the number of inquiries, as households will either inquire at no units or all the units in a week with high probability.

To compute the moments, we solve the model separately for each treatment group at each trial parameter vector. We use the standard two-step approximation to the optimal GMM weighting matrix. We report asymptotic standard errors, and obtain standard errors of estimates of derived parameters such as $\hat{\gamma}$ via the delta method. We provide computational details in the appendix.

In order to obtain "rational expectations" estimates $\hat{\gamma}_{OLS}$ and $\sigma^2_{q,OLS}$, we estimate an OLS regression of q on x in the sample of housing units ever shown to a user on GoSection8. We then conduct a Wald test of the hypothesis $H_0 : \gamma_{OLS} = \gamma$.

7 Results

In Figures 13 and 14 we show the fit of targeted moments. In general our model fits the data well, especially for key moments such as school quality at endline and mean quality of units at which households inquire. Figures 16 and 17 show model-predicted and observed patterns of views and inquiries by period. The model-predicted series are too variable, but tend to match the level and slope of the observed series.

We present parameter estimates in Table 11. Control households appear to dislike high-walkscore neighborhoods, perhaps because they view them as signals of poor schools. Households observe roughly one off-platform unit per week. We find that both treatment and control-group households have apparent tastes for higher-rated schools and lower commute time to city center. A treated household is willing to trade off commute distance and mean school quality at a rate $\beta_q/\beta_{\text{commute}} = -0.858$, or 52 minutes per GreatSchools rating point.[21] In contrast, the apparent preferences of control-group households imply a ratio of $\tilde{\beta}_q/\tilde{\beta}_{\text{commute}} = -0.442$, or 27 minutes per point, implying that, had we estimated the model ignoring information frictions, we would have mistakenly found roughly half the true willingness to travel for school quality.

We observe that the implicit counterfactual underlying this willingness-to-travel measure involves moving people from neighborhoods with characteristics (x,q) to those with characteristics (x',q'). This is the standard counterfactual in the housing literature, e.g. see Black (1999), Bayer, Ferreira and McMillan (2007).[22] Because this counterfactual does not change the joint distribution of (x, q, ϵ), we do not need to distinguish between a taste for q directly and a belief that high test scores predict other measures of school or neighborhood quality.

In Table 12 we compare households' belief parameters (γ, σ_η) to those from an OLS regression of mean quality on characteristics in our sample of housing units. In addition, we report confidence intervals on the prior mean parameters, γ. Estimates of individual elements of γ are noisy, but point estimates indicate that households are pessimistic about school quality. A Wald test of equality of $\gamma = \gamma^{\text{OLS}}$ overwhelmingly rejects this hypothesis ($p < 0.0001$). Families appear to slightly underestimate the variance of q conditional on x. Moreover, control-group families have a signal with $(1 - s) \approx 0.3$, implying that the signals are informative about school quality but not perfectly so.

We plot the implied distributions of residual quality $q - E(q|x)$ in Figure 15. This residual is the object that our information treatment provides to households when their signal η is completely uninformative. The OLS residuals are centered around zero by construction while the subjective beliefs are shifted substantially to the right. This implies that households generally believe school quality is lower than would be predicted

[21] Not all households work in the city center, so distaste for commute to work location would likely be larger than β_{commute}, resulting in a smaller ratio.

[22] A different counterfactual, which we do not consider in this paper, would involve improvements in the technology of schools so as to raise proficiency rates, "all else equal".

by the econometrician; households are pessimistic about locations at which they do not currently live. Our treatment causes households to update positively about the presence of higher-quality schools in their choice set. Overall, our findings imply that we cannot model households' beliefs as measurement error around rational expectations.

8 Conclusions

In this paper, we ask whether low-income families lack information at the time of their search for a home, and whether providing this information affects how they search and where they live. We show that providing this information causes families to live in neighborhoods with higher-performing, less segregated schools. We estimate a model of residential choice that incorporates imperfect information about school quality. Our estimates show that untreated households act as if they are mistaken about the distribution of school quality and other housing characteristics. As a result, there may be returns to providing information that can help households make informed choices. In this paper, however, we do not consider the general equilibrium impacts of information provision that might arise if scaled.

Low-income households encounter many barriers to moving. Households can find it difficult to pay security deposits and application fees, and landlords can discriminate against their race or source of income. Moreover, a lack of supply may incentivize households to accept units even where school quality is low. Policies to ameliorate these frictions, such as subsidies or loans to help families pay for moving costs, landlord incentives to participate in low-income housing programs, construction incentives, or increases in rent formulas can reduce the barriers that impede neighborhood choice. Our results suggest that, if the goal is to help households move to neighborhoods with better schools or other amenities, information provision may be a valuable complement to these more-expensive policies.

Lastly, we demonstrated that families respond to one particular type of school-quality information. However, future work could study whether families respond differently to information about other school and neighborhood characteristics.

References

Abdulkadiroglu, Atila, Parag A Pathak, Jonathan Schellenberg, and Christopher R Walters, "Do Parents Value School Effectiveness?," Working Paper 23912, National Bureau of Economic Research October 2017.

Ackerberg, Daniel A., "A new use of importance sampling to reduce computational burden in simulation estimation," *QME*, Dec 2009, *7* (4), 343–376.

Allende, Claudia, Francisco Gallego, and Christopher Neilson, "The Equilibrium Effects of Informed School Choice," Working Paper 2018.

Angrist, Joshua D., Parag A. Pathak, and Christopher R. Walters, "Explaining Charter School Effectiveness," *American Economic Journal: Applied Economics*, October 2013, *5* (4), 1–27.

Bayer, Patrick, Fernando Ferreira, and Robert McMillan, "A unified framework for measuring preferences for schools and neighborhoods," *Journal of political economy*, 2007, *115* (4), 588–638.

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

Black, Sandra E., "Do Better Schools Matter? Parental Valuation of Elementary Education," *The Quarterly Journal of Economics*, 1999, *114* (2), 577–599.

Chetty, Raj, John N. Friedman, and Jonah E. Rockoff, "Measuring the Impacts of Teachers II: Teacher Value-Added and Student Outcomes in Adulthood," *American Economic Review*, 2014, *104* (9), 2633–79.

_ , _ , Nathaniel Hilger, Emmanuel Saez, Diane Whitmore Schanzenbach, and Danny Yagan, "How Does Your Kindergarten Classroom Affect Your Earnings? Evidence from Project Star," *The Quarterly Journal of Economics*, 2011, *126* (4), 1593–1660.

__ , **Nathaniel Hendren, and Lawrence F. Katz**, "The Effects of Exposure to Better Neighborhoods on Children: New Evidence from the Moving to Opportunity Experiment," *American Economic Review*, 2016, *106* (4), 855–902.

Chyn, Eric, "Moved to Opportunity: The Long-Run Effect of Public Housing Demolition on Children," *American Economic Review*, October 2018.

Clampet-Lundquist, Susan and Douglas S. Massey, "Neighborhood Effects on Economic Self Sufficiency: A Reconsideration of the Moving to Opportunity Experiment," *American Journal of Sociology 114.1: 107-143*, 2008.

Corcoran, Sean P, Jennifer L Jennings, Sarah R Cohodes, and Carolyn Sattin-Bajaj, "Leveling the Playing Field for High School Choice: Results from a Field Experiment of Informational Interventions," Working Paper 24471, National Bureau of Economic Research March 2018.

Deming, David J., Justine S. Hastings, Thomas J. Kane, and Douglas O. Staiger, "School Choice, School Quality, and Postsecondary Attainment," *The American Economic Review*, 2014, *104* (3), 991–1013.

Dobbie, Will and Roland G. Fryer, "The Medium-Term Impacts of High-Achieving Charter Schools," *Journal of Political Economy*, 2015, *123* (5), 985–1037.

Ellen, Ingrid Gould, Keren Mertens Horn, and Amy Ellen Schwartz, "Why don't housing choice voucher recipients live near better schools? Insights from Big Data," *Journal of Policy Analysis and Management*, 2016, *35* (4), 884–905.

Figlio, David N and Maurice E Lucas, "What's in a grade? School report cards and the housing market," *American Economic Review*, 2004, *94* (3), 591–604.

Finkel, Meryl and Larry Buron, "Study on Section 8 Voucher Success Rates: Volume I Quantitative Study of Success Rates in Metropolitan Areas. Report prepared for the US Department of Housing and

Urban Development, Office of Policy Development and Research," 2001.

Gennetian, Lisa, Matthew Sciandra, Lisa Sanbonmatsu, Jens Ludwig, Lawrence Katz, Greg Duncan, Jeffrey Kling, and Ronald Kessler, "The Long-Term Effects of Moving to Opportunity on Youth Outcomes," *Cityscape*, 2012, *14* (2).

Glazerman, Steven and Dallas Dotter, "Market Signals: Evidence on the Determinants and Consequences of School Choice From a Citywide Lottery," *Educational Evaluation and Policy Analysis*, 2017, *39* (4), 593–619.

Hastings, Justine, Thomas Kane, and Douglas Staiger, "Heterogeneous Preferences and the Efficacy of Public School Choice," Working Paper June 2009.

HUD, "HUD, Greatschools Team up to give public housing, voucher families tools to make informed school choices," 2011.

Jacob, Brian A., "Public Housing, Housing Vouchers, and Student Achievement: Evidence from Public Housing Demolitions in Chicago," *American Economic Review, 94(1): 233-258*, 2004, *94* (1), 233–258.

_ , Jens Ludwig, and Max Kapustin, " The Impact of Housing Assistance on Child Outcomes: Evidence from a Randomized Housing Lottery *," *The Quarterly Journal of Economics*, 11 2014, *130* (1), 465–506.

Katz, Lawrence F., Jeffrey B. Liebman, and Jeffrey R. Kling, "Moving to Opportunity in Boston: Early Results of a Randomized Mobility Experiment," *Quarterly Journal of Economics*, 2001, *116* (2), 607–54.

Kling, Jeffrey R., Jeffrey B. Liebman, and Lawrence F. Katz, "Experimental Analysis of Neighborhood Effects," *Econometrica 75 (1): 83-119*, 2007, *75* (1), 83–119.

Ludwig, Jens, Greg J. Duncan, Lisa A. Gennetian, Lawrence F. Katz, Ronald C. Kessler, Jeffrey R. Kling, and Lisa Sanbonmatsu, "Long-Term Neighborhood Effects on Low-Income Families: Evidence from Moving to Opportunity," *American Economic Review Papers and Proceedings 103(3):*

226-31, 2013.

Reardon, Sean F, "Educational Opportunity in Early and Middle Childhood: Variation by Place and Age," CEPA Working Paper 17-12, Center for Education Policy Analysis, Stanford University March 2018.

Sanbonmatsu, Lisa, Jens Ludwig, Lawrence F. Katz, Lisa A. Gennetian, Greg J. Duncan, Ronald C. Kessler, Emma Adam, Thomas McDade, and Stacy Tessler Lindau, *Moving to Opportunity for Fair Housing Demonstration Program: Final Impacts Evaluation*, Washington, DC: U.S. Department of Housing and Urban Development, Office of Policy Development and Research, 2011.

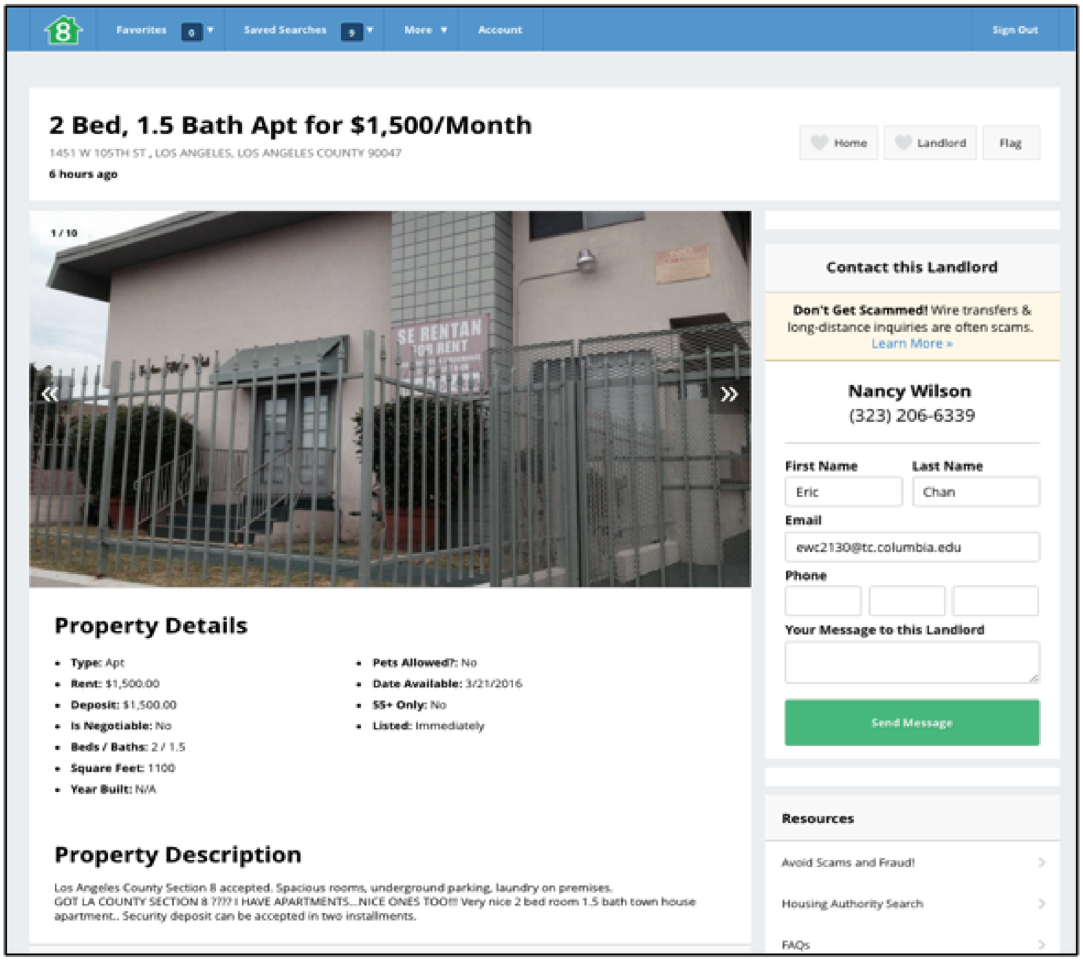
van Dijk, Winnie, "The Socio-Economic Consequences of Housing Assistance," Working Paper 2019.

Waldinger, Daniel, "Targeting In-Kind Transfers Through Market Design: A Revealed Preference Analysis of Public Housing Allocation," Working Paper 2018.

Weinstein, Jeffrey M. and Justine S. Hastings, "Information, School Choice, and Academic Achievement: Evidence from Two Experiments*," *The Quarterly Journal of Economics*, 11 2008, *123* (4), 1373–1414.

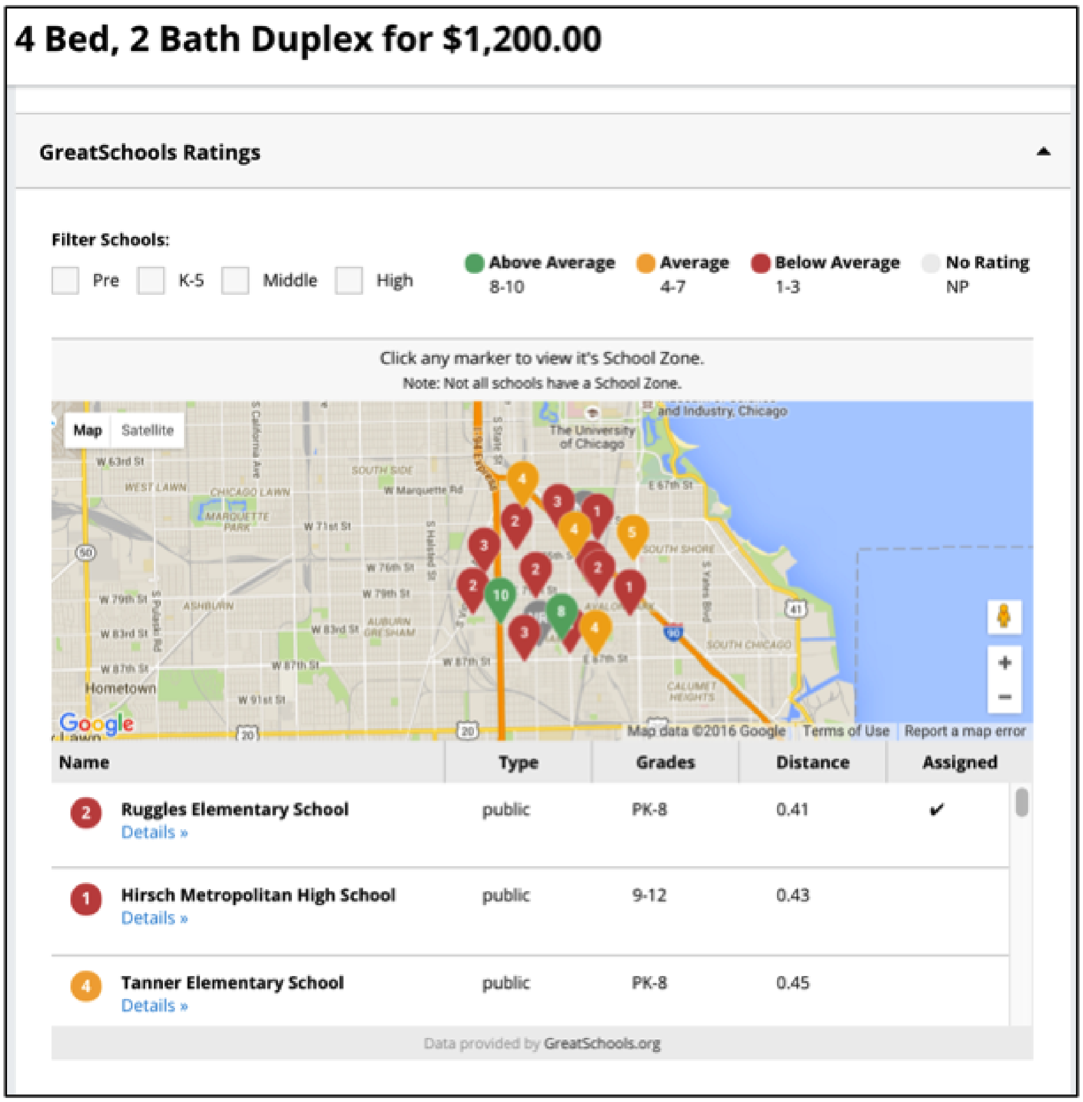
Figures

Figure 1: Property Listing



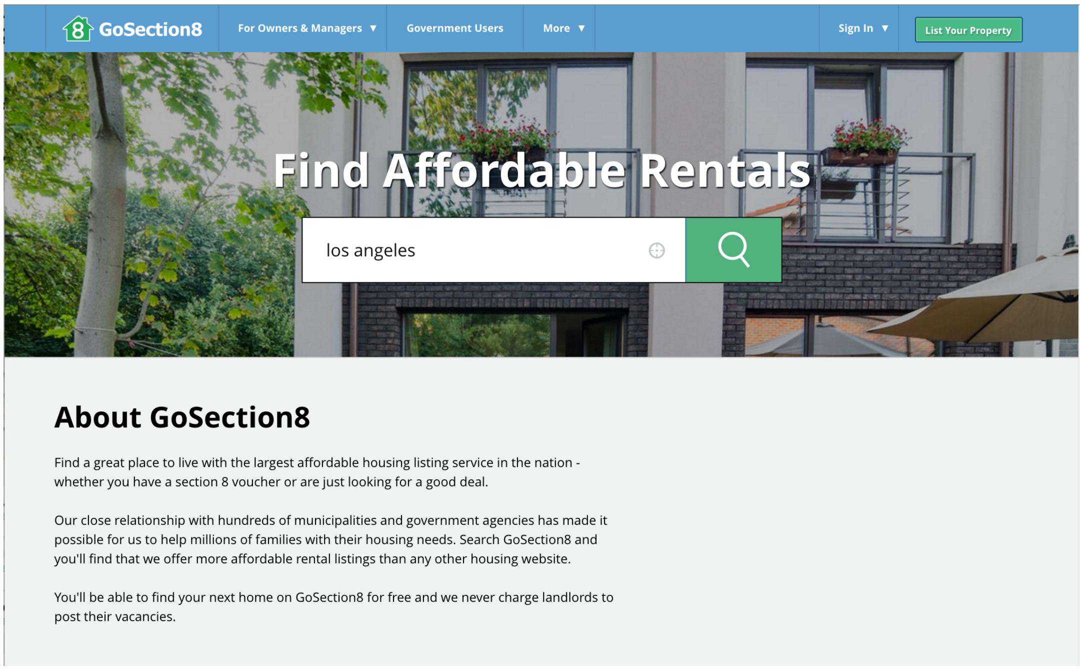
This figure shows the property listing page on GoSection8.com. The intervention module is directly underneath the shown area. See next figure.

Figure 2: Greatschools module



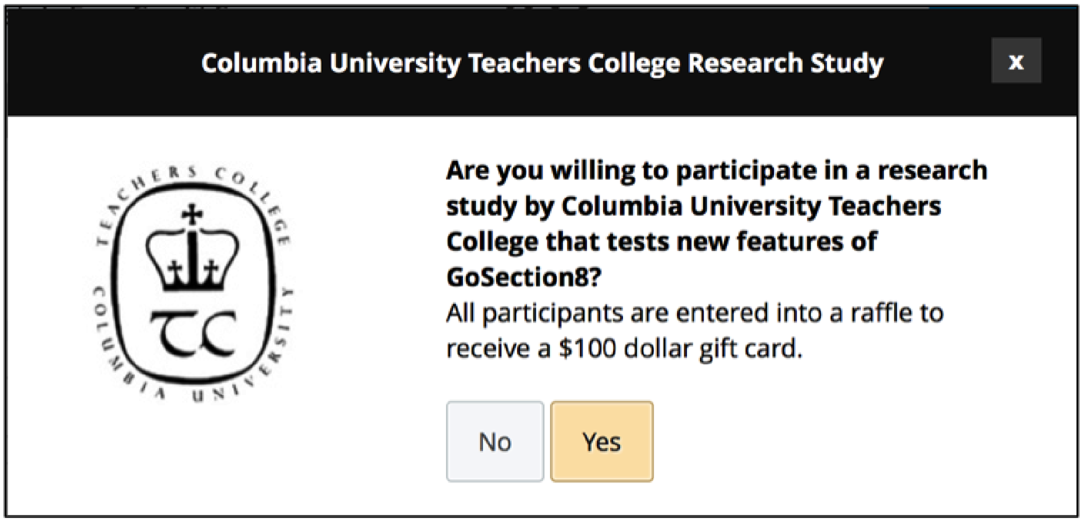
This figure shows the intervention module with school quality information on each of GoSection8's property listing page. It includes a map, the assigned (or closest school if no assigned schools) schools at each school level, and school quality information.

Figure 3: Landing page



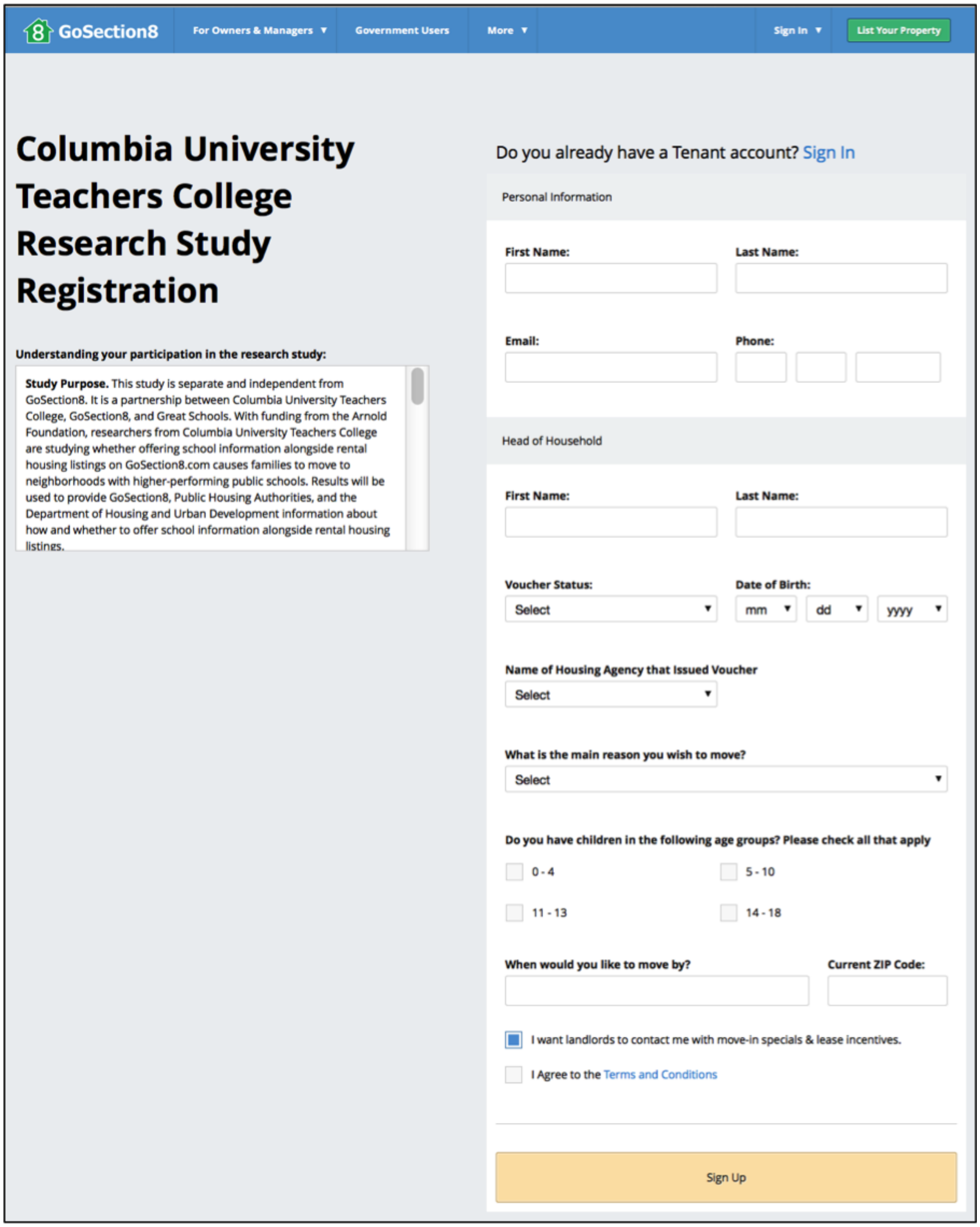
This figure shows the landing page on GoSection8.com, where users can search for rental properties by zip code, city, or county.

Figure 4: Consent pop up



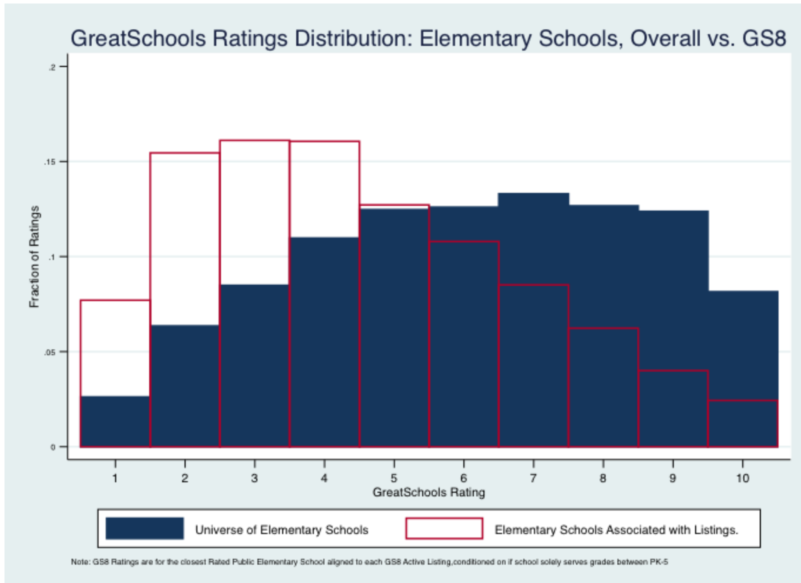
This figure shows the consent pop up when users first log into GoSection8.com.

Figure 5: Intake survey



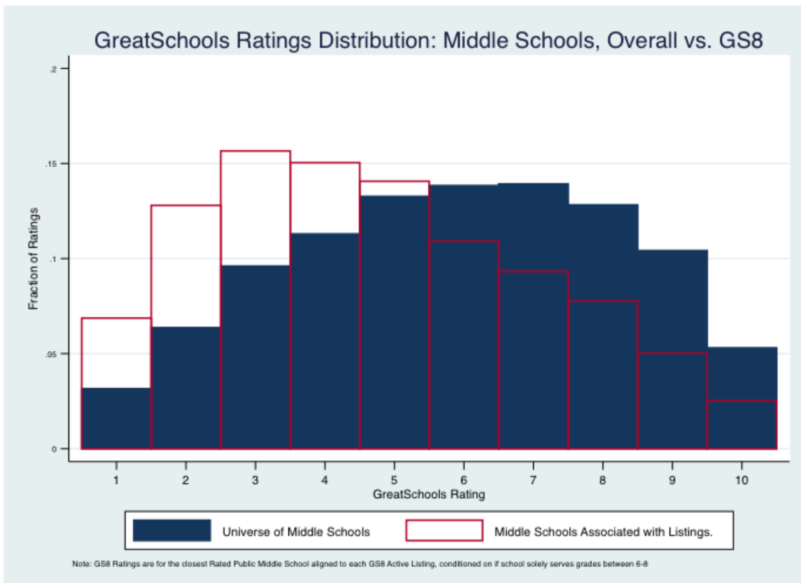
This figure shows the survey after users click 'yes' to the consent pop up on GoSection8.com.

Figure 6: Distribution of GreatSchools Public School Ratings Nationally vs. GoSection8 Properties



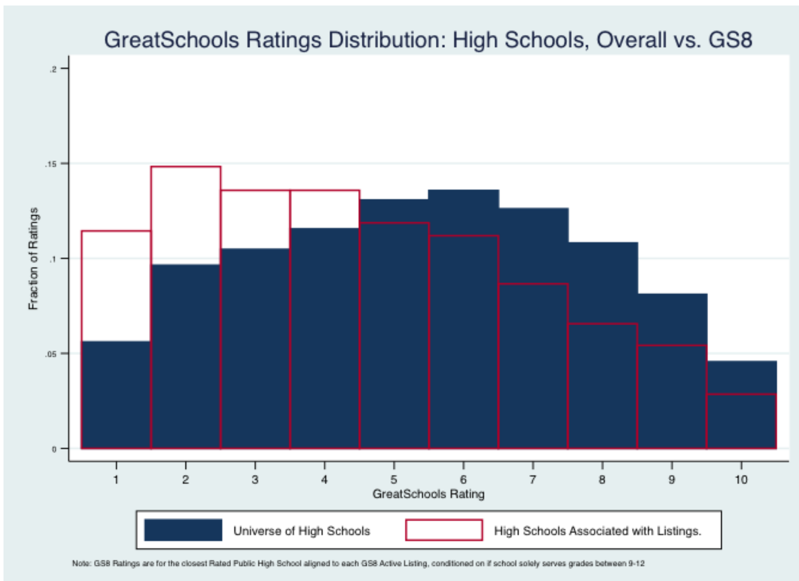
This figure compares the GreatSchools Public School Ratings for the universe of elementary schools vs. the elementary schools associated with the GoSection8 properties listed on its website. Data from GreatSchools and GoSection8.

Figure 7: Distribution of GreatSchools Public School Ratings Nationally vs. GoSection8 Properties



This figure compares the GreatSchools Public School Ratings for the universe of middle schools vs. the middle schools associated with the GoSection8 properties listed on its website. Data from GreatSchools and GoSection8.

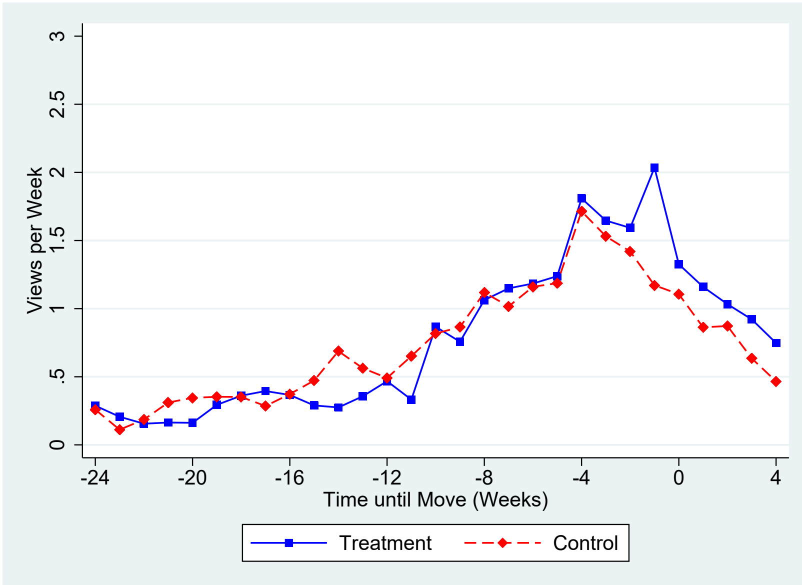
Figure 8: Distribution of GreatSchools Public School Ratings Nationally vs. GoSection8 Properties



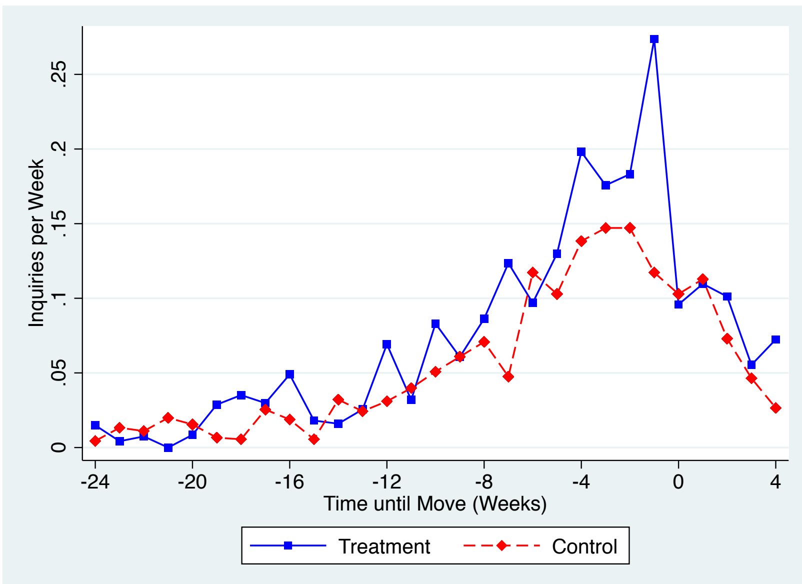
This figure compares the GreatSchools Public School Ratings for the universe of middle schools vs. the middle schools associated with the GoSection8 properties listed on its website. Data from GreatSchools and GoSection8.

Figure 9: Number of Views and Inquiries Made Relative to Move Date

(a) Views by week



(b) Inquiries by week

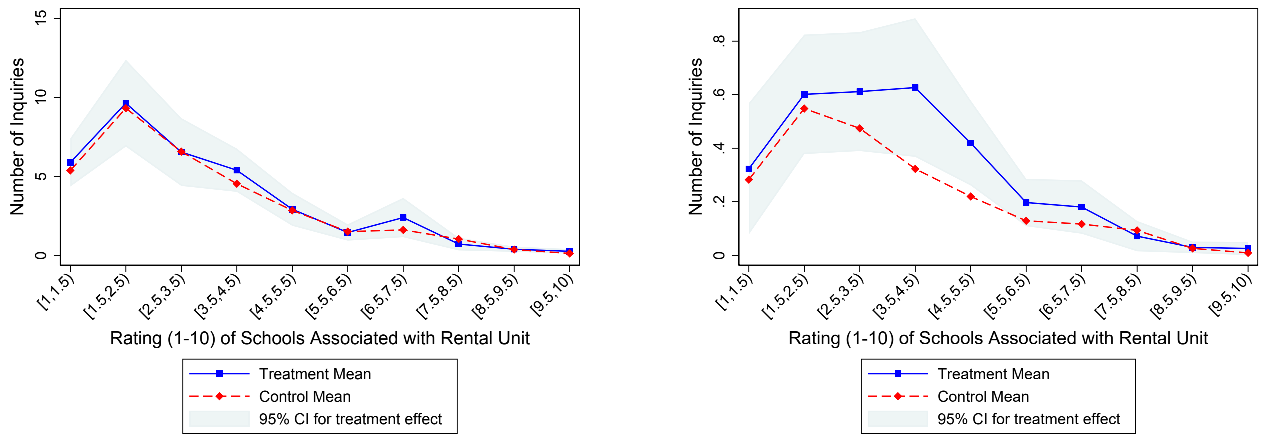


These figures compare the treatment and control group views and inquiries made relative to time until move (in weeks), where 0 is the week when user reported that they desire to move. Data from GreatSchools and GoSection8.

Figure 10: Searches and Inquiries by School Quality

(a) Number of Views by Mean School Quality

(b) Number of Inquiries by Mean School Quality



These figures show the number of views and inquiries by mean school quality.

Figure 11: Density of Average Ratings

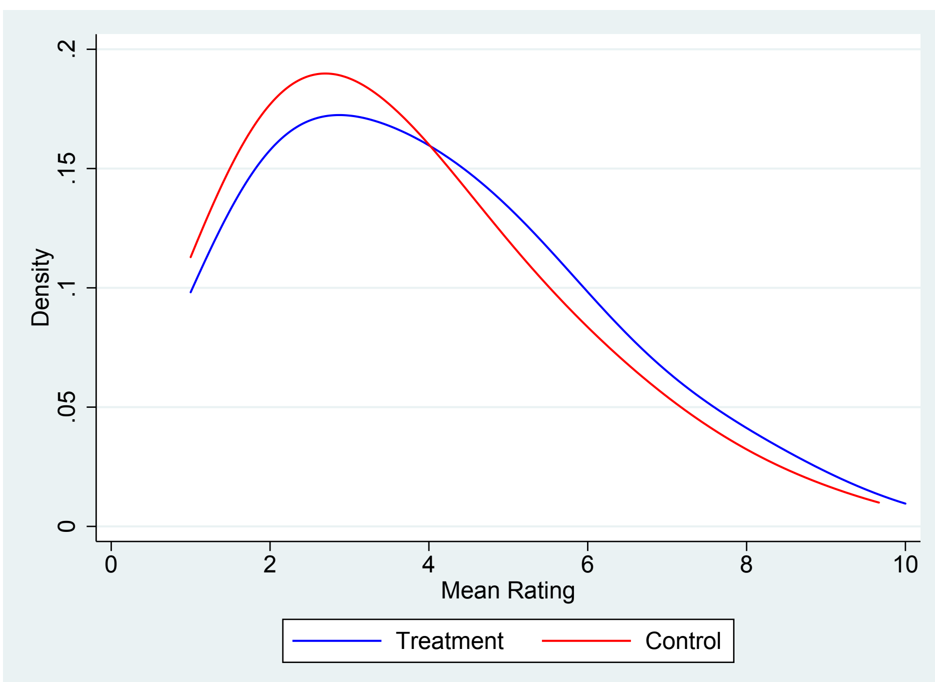


Figure 12: Poverty Density

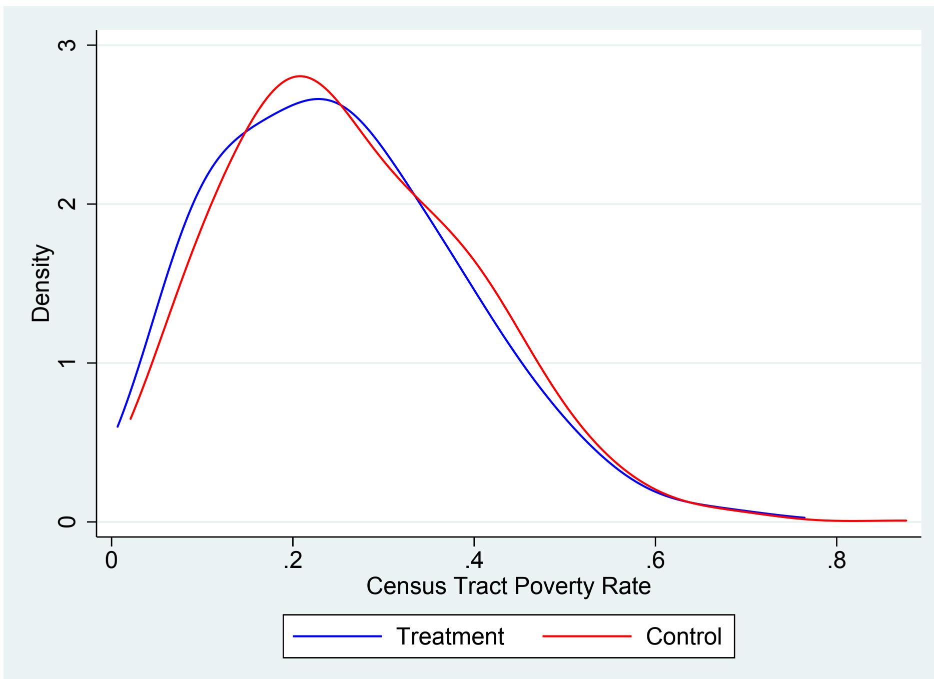
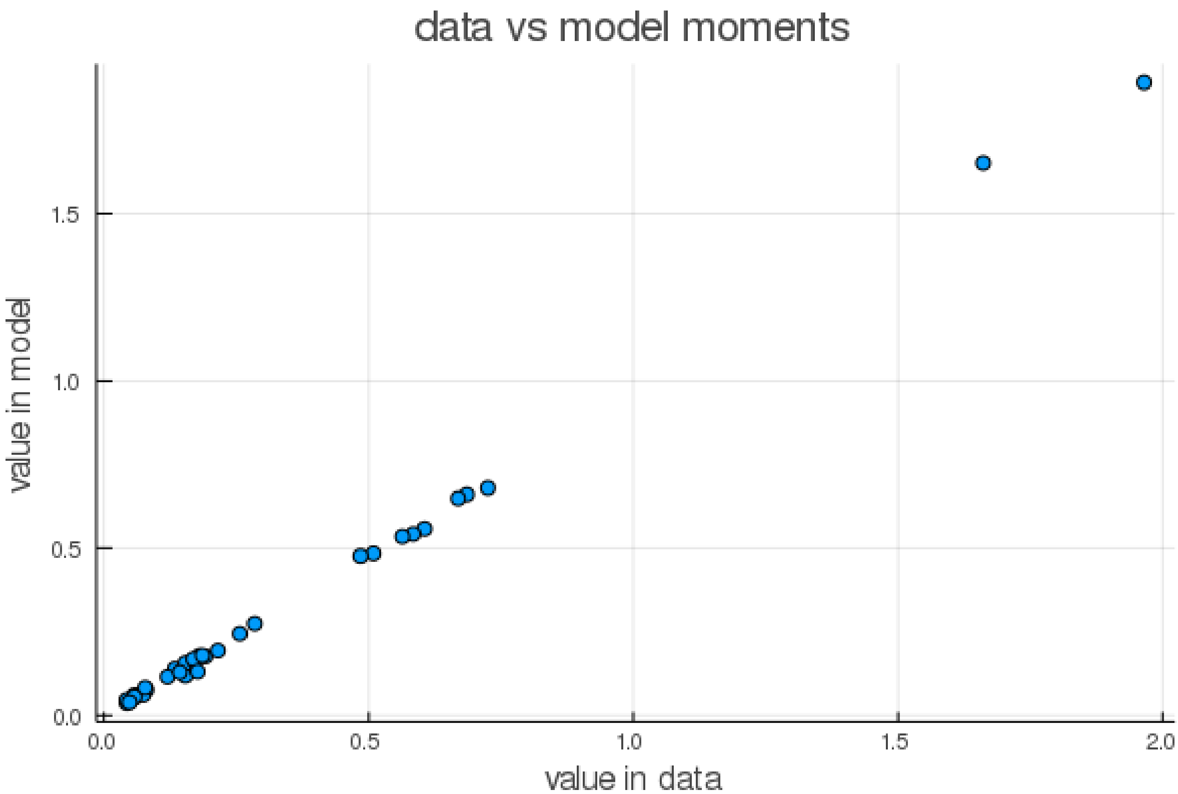
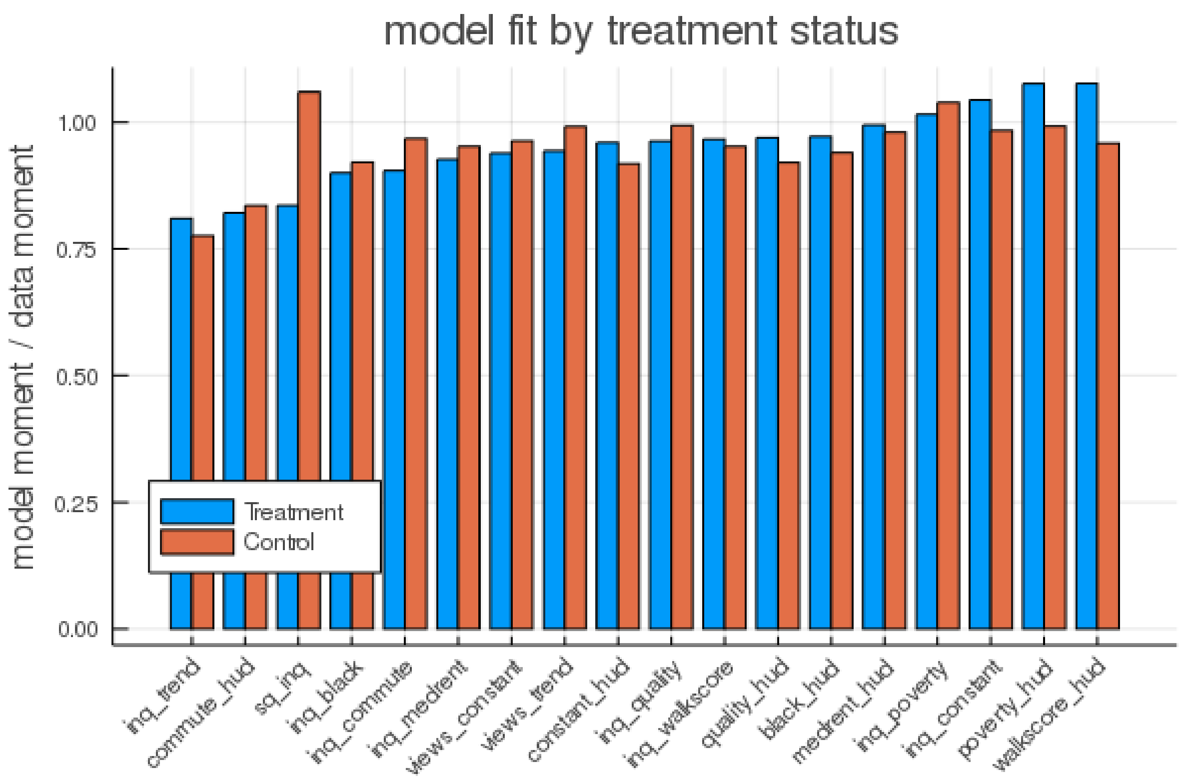


Figure 13: Model fit: all moments



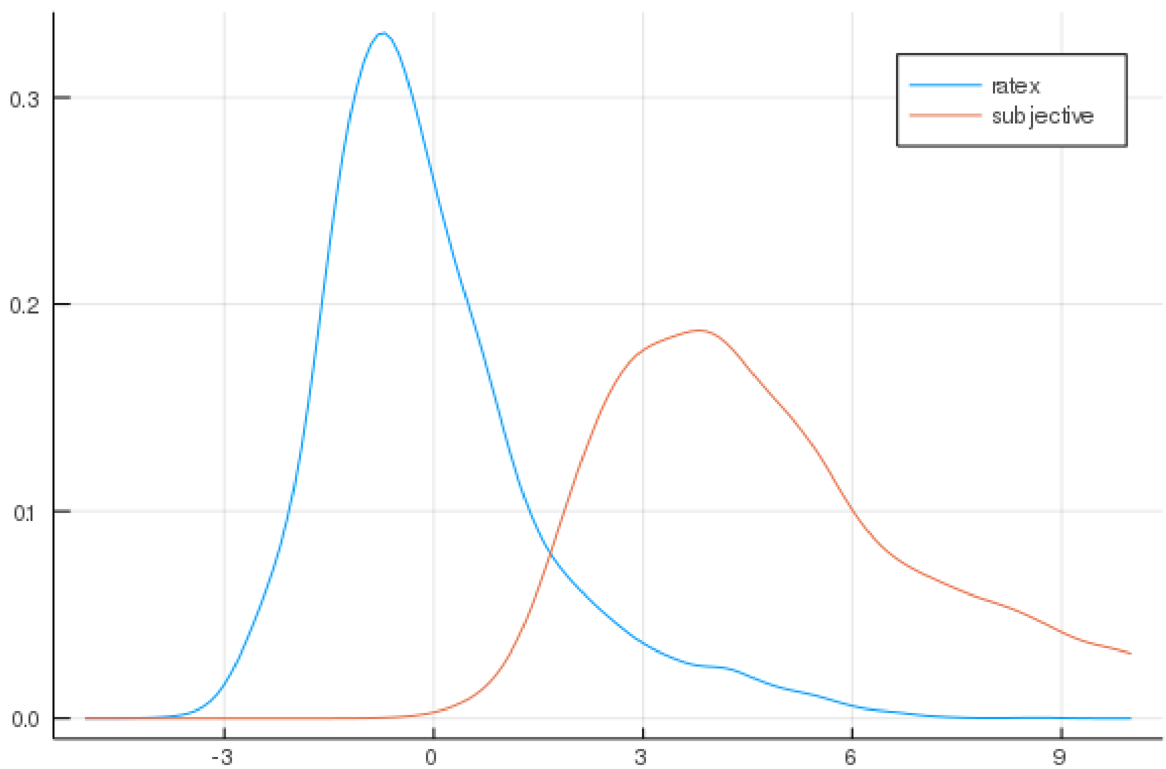
This figure compares the value of each of the targeted moments in the data (x-axis) to the value in our estimated model. Points along the line $y = x$ indicate good fit.

Figure 14: Model fit: all moments



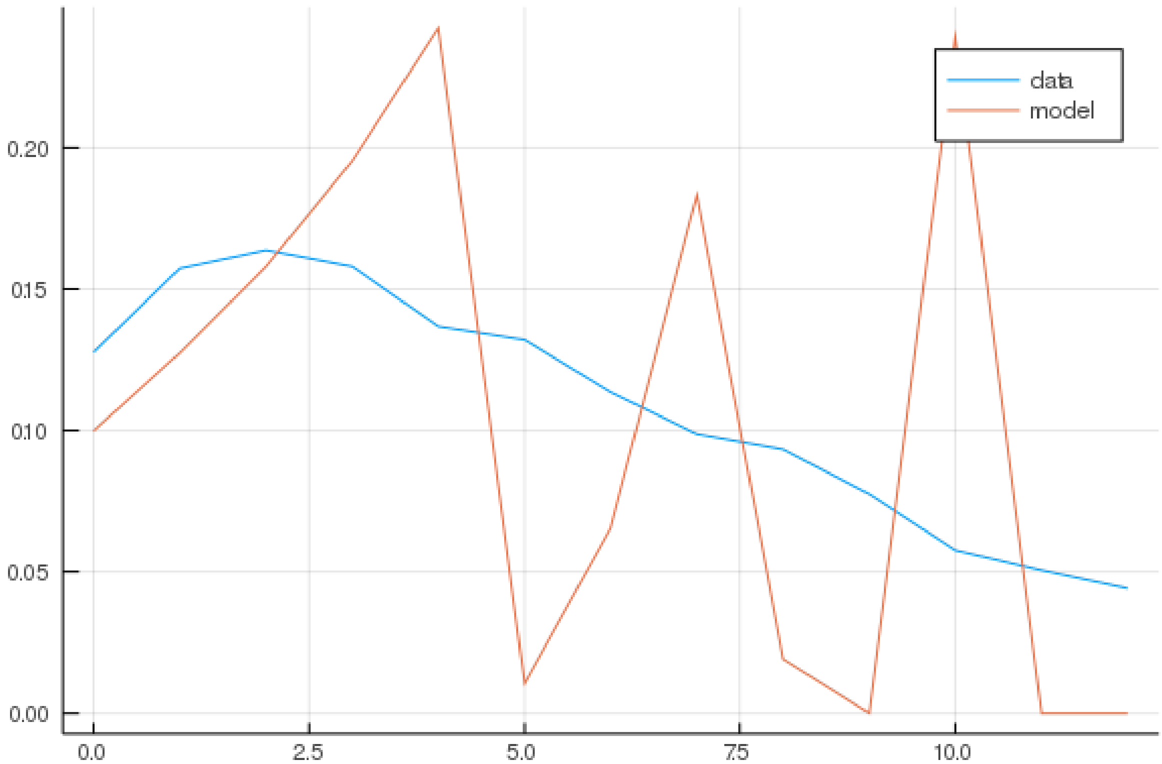
This figure compares the average over households in our dataset of each of the targeted moments to the value in our estimated model, presenting the same information as figure 13. Moments are sorted by ratio of model-predicted value to mean value over households in the data. Key: inq_x denotes the sum of $x \in$ {constant, median rent, poverty, share black, commute time, walkscore} over all inquiries made by a household. x_hud denotes the value of $x_j * 1(moved)$ where j is the unit that the household lives in at endline. sq_inq denotes the second moment of share of viewed units which receive direct inquiries within each household-week.

Figure 15: Kernel density of residuals $q - E(q|x)$



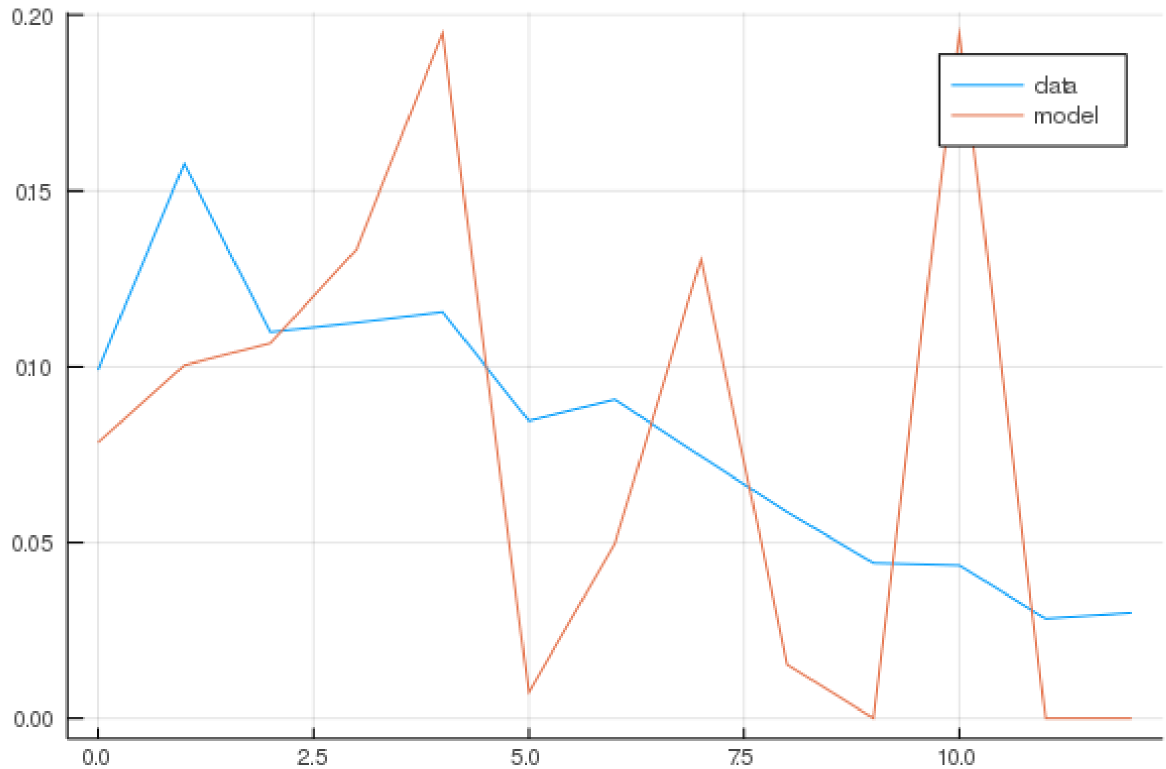
This figure displays the density of "residual quality" $q - E(q|x)$ according to OLS and subjective estimates of $E(q|x)$. This residual is the object that our information treatment provides to households when their signal η is completely uninformative.

Figure 16: Model fit: 1(use platform in week t)



This figure compares data means and model estimates of the probability of using the platform in each week as a function of time until intended move date. The x-axis denotes weeks remaining until the household's reported intended move date.

Figure 17: Model fit: number of inquiries in week t



This figure compares data means and model estimates of the mean number of direct inquiries made by each household in each week as a function of time until intended move date. The x-axis denotes weeks remaining until the household's reported intended move date.

Tables

Table 1: Demographic Characteristics: HUD (5% Random Sample) and Study Sample Comparison

Variable	go8 (matched) Mean	matched-HUD Difference	P-value	N (HUD)	N (matched)
HH female	0.87	0.03***	0.00	347,548	1,954
HH disabled	0.31	−0.16***	0.00	347,548	1,954
HH white	0.25	0.02**	0.02	347,543	1,949
HH Black	0.60	0.02**	0.05	347,545	1,951
HH Hispanic	0.14	−0.04***	0.00	347,548	1,954
Total household members	3.01	0.59***	0.00	347,548	1,954
Number of bedrooms	2.43	0.38***	0.00	347,544	1,954
Total annual income	14634.54	23.20	0.93	347,548	1,954
Rent to owner	1071.22	45.46***	0.00	347,548	1,954
Gross rent	1213.23	76.52***	0.00	347,548	1,954
Utility allowance	138.63	28.52***	0.00	347,548	1,954

Notes: All data from HUD. The abbreviation "HH" stands for "head of household." Data is a 5% random sample of heads of households restricted to those who ever completed a HUD Form 50058 with HUD between 2013-2017. For GoSection8 observations, it is restricted to only those who matched with the HUD database of voucher holders. *** p<0.01, ** p<0.05, * p<0.10

Table 2: Descriptive Comparison of School Characteristics

	US Elem Schools	GS8 Listings	HUD (5%)	Matched
GreatSchools Rating	5.78	5.28	4.71	3.27
Share Black	0.15	0.25	0.20	0.44
Share Hispanic	0.26	0.33	0.48	0.38
Share White	0.49	0.33	0.22	0.12
Share Asian	0.05	0.04	0.06	0.03
Share FRPL	0.52	0.63	0.71	0.84
Pupil-FTE Ratio	17.67	18.22	18.58	18.35
Observations (schools)	125,346	17,583	328,301	1,853

This table shows the means of the GreatSchools school rating and demographic data for the universe of US elementary schools, elementary schools associated with GoSection8 property listings, elementary schools associated with a 5% random sample of HUD properties, and elementary schools associated with participants' addresses that we were able to match to HUD data, respectively. Note that for the latter three columns, schools may be counted multiple times as they are matched to specific addresses. These columns can be interpreted as weighted means of ratings and characteristics. Also, a small share of schools do not have sufficient data to be rated, and thus are treated as missing. *** p<0.01, ** p<0.05, * p<0.10

Table 3: Experimental Sample Balance Table

Variable	Control Mean	Treatment-Control Difference	P-value	N
HH female	0.88	−0.02	0.31	1,921
HH Hispanic	0.15	−0.03**	0.03	1,921
HH Black	0.59	0.01	0.55	1,932
HH white	0.24	0.01	0.63	1,916
Total annual income	14,512.75	103.54	0.84	1,921
Total household members	3.05	−0.07	0.39	1,921
Moving to Work (MTW)	0.18	−0.04**	0.02	1,932
HUD count of children <18 years old	1.84	−0.06	0.43	1,932
Intend to move within 3 months	0.76	0.01	0.63	1,932
Moving for schools	0.31	0.02	0.29	1,932
HUD user matched to go8 user	0.66	−0.01	0.49	2,968
Omnibus Test P-value: 0.19				

Notes: All data from HUD, with the exception of intention to move within three months and moving for schools, which are derived from baseline survey data on GoSection8. Differences estimated from a regression of the baseline variable on a treatment indicator. Robust standard errors.

Table 4: Property Views

Variable	Control Mean	Treatment Effect	Std Error	P-value	N
Total vews made post treatment	33.57	1.80	4.21	0.67	1,932
Any views	0.77	−0.00	0.02	0.82	1,932
Number of Different Tracts Made Views In	14.36	0.11	1.26	0.93	1,932
Number of Views Rated 1 or Above	33.22	1.55	4.09	0.70	1,932
Number of Views Rated 2 or Above	24.73	2.72	3.29	0.41	1,932
Number of Views Rated 3 or Above	15.94	2.69	2.24	0.23	1,932
Number of Views Rated 4 or Above	10.27	1.88	1.56	0.23	1,932
Number of Views Rated 5 or Above	6.17	1.25	1.10	0.25	1,932
Number of Views Rated 6 or Above	3.90	0.97	0.82	0.24	1,932
Number of Views Rated 7 or Above	2.54	0.70	0.69	0.31	1,932
Number of Views Rated 8 or Above	1.03	0.03	0.21	0.88	1,932
Number of Views Rated 9 or Above	0.31	0.12	0.09	0.19	1,932
Number of Views Rated 10	0.07	0.12**	0.05	0.03	1,932

Notes: All outcome data from GoSection8.
*** p<0.01, ** p<0.05, * p<0.10

Table 5: Property Inquiries

Variable	Control Mean	Treatment Effect	Std Error	P-value	N
Total inquiries made post treatment	2.32	0.80*	0.45	0.07	1,932
Any Inquiries	0.34	0.03	0.02	0.20	1,932
Number of Different Tracts Made Inquiries In	1.58	0.28	0.22	0.21	1,932
Number of Inquiries Rated 1 or Above	2.30	0.81*	0.45	0.07	1,932
Number of Inquiries Rated 2 or Above	1.84	0.76**	0.36	0.04	1,932
Number of Inquiries Rated 3 or Above	1.27	0.67**	0.30	0.02	1,932
Number of Inquiries Rated 4 or Above	0.82	0.46**	0.20	0.03	1,932
Number of Inquiries Rated 5 or Above	0.52	0.19	0.14	0.18	1,932
Number of Inquiries Rated 6 or Above	0.32	0.09	0.10	0.34	1,932
Number of Inquiries Rated 7 or Above	0.20	0.05	0.07	0.53	1,932
Number of Inquiries Rated 8 or Above	0.10	−0.01	0.03	0.66	1,932
Number of Inquiries Rated 9 or Above	0.03	0.01	0.02	0.65	1,932
Number of Inquiries Rated 10	0.00	0.02*	0.01	0.10	1,932

Notes: All outcome data from GoSection8.
*** p<0.01, ** p<0.05, * p<0.10

Table 6: Mean "School Quality" Views and Inquiries

Variable	Control Mean	Treatment Effect	Std Error	P-value	N
Mean Quality Views (Next Rate)	3.53	−0.08	0.08	0.31	2,210
Mean Quality Inquiries (Next Rate)	3.51	0.02	0.13	0.87	1,011
Mean Quality Views (Mean Rate)	2.91	−0.12	0.08	0.12	2,793
Mean Quality Inquiries (Mean Rate)	3.58	0.08	0.11	0.48	1,038

Notes: All outcome data from GoSection8.
*** p<0.01, ** p<0.05, * p<0.10

Table 7: School Characteristics

Variable	Control Mean	Treatment Effect	Std Error	P-value	N
Average School Rating	3.65	0.26***	0.09	0.00	1,922
Maximum School Rating	4.82	0.31***	0.11	0.00	1,922
Minimum School Rating	2.64	0.18**	0.08	0.03	1,922
High School Rating	4.01	0.20*	0.11	0.06	1,821
Middle School Rating	3.52	0.36***	0.11	0.00	1,888
Primary School Rating	3.47	0.23**	0.11	0.03	1,864
Mean share FRPL	0.72	−0.02**	0.01	0.04	1,866
Fraction Black or Hispanic	0.66	−0.03**	0.01	0.02	1,866

Notes: All outcome data from HUD merged to school quality data from GreatSchools as shown on GoSection8.
*** p<0.01, ** p<0.05, * p<0.10

Table 8: School Characteristics by Subgroup

Variable	Control Mean	Treatment Effect	Std Error	P-value	N
Has Child 0-4					
Primary School Rating	3.26	0.29	0.18	0.11	573
Middle School Rating	3.43	0.19	0.19	0.33	579
High School Rating	3.98	0.02	0.20	0.91	557
Has Child 5-10					
Primary School Rating	3.51	0.36	0.26	0.16	386
Middle School Rating	3.40	0.93***	0.26	0.00	383
High School Rating	3.84	0.49**	0.24	0.04	378
Has Child 11-13					
Primary School Rating	3.37	0.19	0.29	0.53	236
Middle School Rating	3.61	−0.17	0.31	0.59	242
High School Rating	3.93	−0.09	0.29	0.77	237
Has Child 14-18					
Primary School Rating	3.65	0.05	0.26	0.85	376
Middle School Rating	3.62	0.22	0.24	0.36	388
High School Rating	4.35	−0.10	0.25	0.69	367
Average Next-Level Rating	3.65	0.31***	0.12	0.01	1,574

Notes: All outcome data from HUD merged to school quality data.
*** p<0.01, ** p<0.05, * p<0.10

Table 9: Endline Neighborhood Characteristics

Variable	Control Mean	Treatment Effect	Std Error	P-value	N
Percent Hispanic	0.23	−0.01	0.01	0.14	1,907
Percent White	0.50	0.01	0.01	0.51	1,907
Percent Black	0.34	−0.01	0.01	0.44	1,907
Percent Asian	0.04	0.00	0.00	0.40	1,907
Percent H.S. Graduates	0.79	0.01	0.00	0.11	1,907
Percent B.A. Graduates	0.19	0.00	0.01	0.45	1,907
Percent in Poverty	0.25	−0.01**	0.01	0.03	1,919
Percent on SNAP	0.56	−0.00	0.01	0.75	1,906
Walkscore	50.02	−3.87***	1.03	0.00	1,929
Commute to dwtn	15.36	1.89***	0.67	0.00	1,913
Median Gross Rent	980.06	13.15	12.31	0.29	1,917
Residual School Qlty	-0.09	0.18**	0.08	0.02	1,880
Predicted School Qlty	3.73	0.09**	0.04	0.04	1,887

Notes: All outcome data from HUD merged to GreatSchools school quality data and American Community Survey 5-year estimates (2012-2016) data.
*** p<0.01, ** p<0.05, * p<0.10

Table 10: Unit Characteristics (HUD Data)

Variable	Control Mean	Treatment Effect	Std Error	P-value	N
gross rent	1216.29	−18.08	21.24	0.39	1,921
number bedrooms	2.45	−0.03	0.05	0.51	1,921
Beds per Household Member	0.92	0.02	0.02	0.36	1,875

Notes: All outcome data from HUD.
*** p<0.01, ** p<0.05, * p<0.10

Table 11: Parameter Estimates

Variable	Estimate	SE
β(constant)	-3.1614	(0.3577)
β(medrent)	-0.4773	(0.1748)
β(poverty)	-0.1581	(0.1481)
β(black)	0.044	(0.1573)
β(walkscore)	-0.1458	(0.0784)
β(commute)	-0.3072	(0.0665)
β(quality)	0.2667	(0.0221)
β^{control}(constant)	-3.1793	(1.0043)
β^{control}(medrent)	-0.0827	(0.4179)
β^{control}(poverty)	-0.4169	(1.0026)
β^{control}(black)	0.0882	(0.1533)
β^{control}(walkscore)	-0.3235	(0.0786)
β^{control}(commute)	-0.3255	(0.5673)
β^{control}(quality)	0.1438	(0.0467)
μ_{search}	-2.1642	(0.2079)
σ_c(search)	0.0211	(0.0846)
μ_{inquire}	-4.3067	(0.5692)
σ_c(inquire)	1.0167	(1.3965)
σ_ϵ(treat)	0.4374	(0.0093)
σ_ϵ(control)	0.7042	(0.0847)
Pr(inquiry observed)	0.6453	(0.0589)
λ	1.0324	(0.9149)
pr(passive)	0.6062	(0.0372)

Table 12: Subjective Beliefs vs Rational Expectations

	Next School Quality	
	(OLS)	(Subjective)
γ(constant)	3.5381	0.1456
	(0.0657)	(7.4353)
γ(medrent)	1.5222	−3.2116
	(0.0416)	(4.2366)
γ(poverty)	−2.1992	2.106
	(0.0943)	(8.5542)
γ(black)	−1.8509	−0.3597
	(0.0322)	(2.0723)
γ(walkscore)	−0.6166	1.4462
	(0.0436)	(0.9601)
γ(commute)	−0.3428	0.1489
	(0.0501)	(4.6907)
resid. variance	2.6605	1.7558
$\sigma_\eta^2/(\sigma^2\eta + \sigma_q^2)$		0.2907

ratex: N=15355 obs. Standard errors in parentheses.

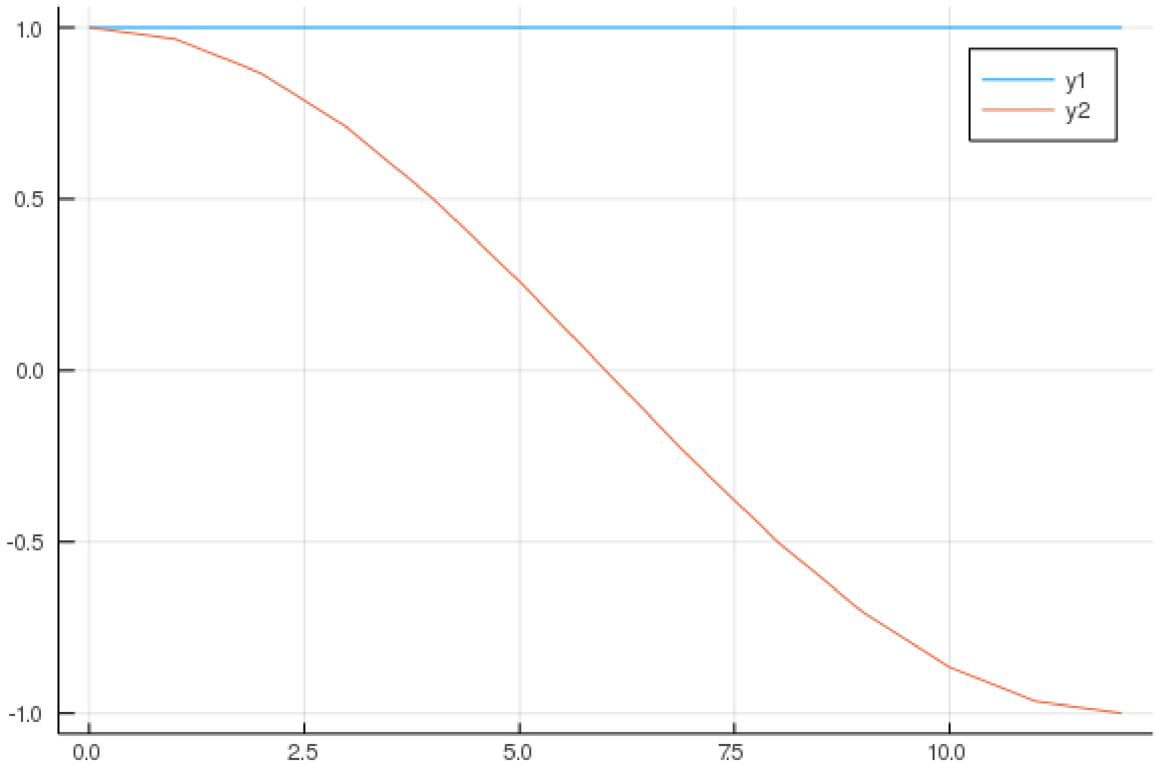
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Wald test $\gamma_{ols} == \gamma$: $\chi^2 = 57.3333$, $p = 0.0$

Appendices

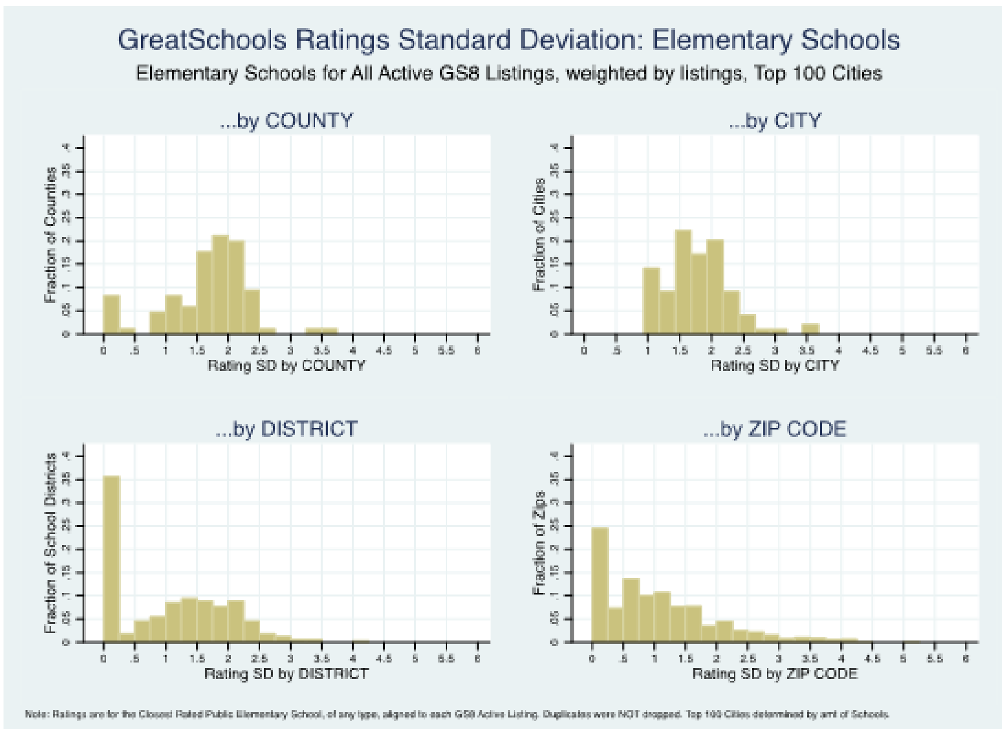
9 Additional tables and figures

Figure A.1: Basis functions for time-varying moments



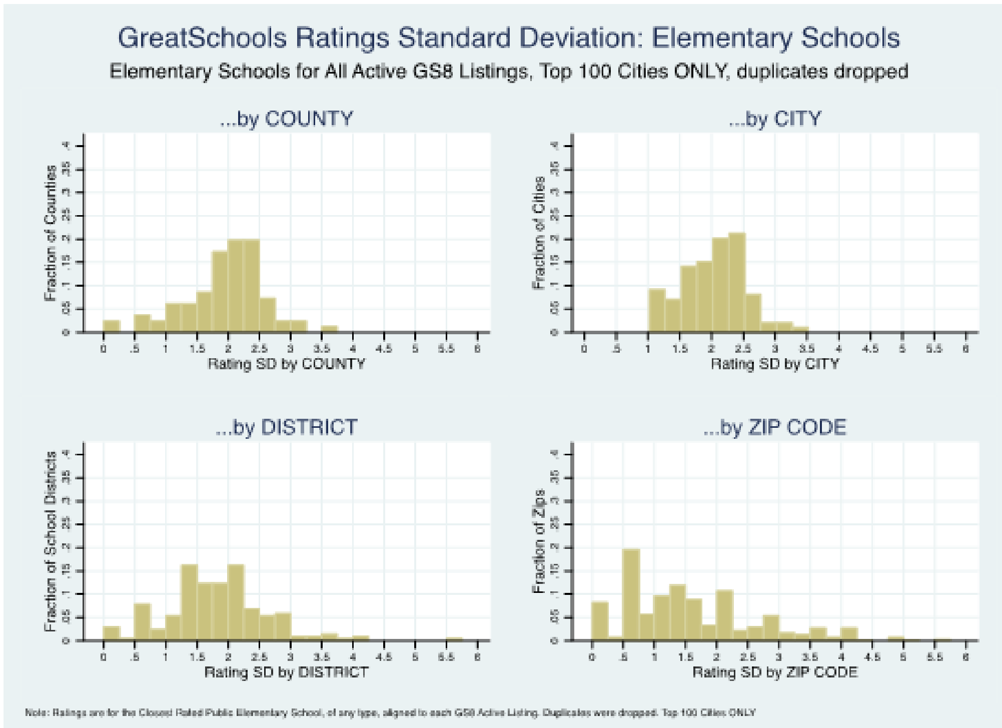
This figure shows the basis functions used to aggregate the time-varying moments, 1(used the platform in week t) and number of inquiries in week t.

Figure A.2: Elementary School Quality Variation associated with GoSection8 Listings, duplicated schools included.



This figure shows the variation in school quality for GoSection8 listings, within the county, city, district, and zip code levels. In general, users can search listings at the county, city, and zip code levels on the web site. This figure includes duplicated schools; for example, if a school is associated with multiple listings, it is included.

Figure A.3: Elementary School Quality Variation associated with GoSection8 Listings, duplicated schools excluded.



This figure shows the variation in school quality for GoSection8 listings, within the county, city, district, and zip code levels. In general, users can search listings at the county, city, and zip code levels on the web site. This figure excludes duplicated schools; for example, if a school is associated with multiple listings, it is not included

10 Computational details

This section describes our model solution and estimation procedure in detail.

To approximate the model moments, we draw a grid of 4000 sets of properties per time period $t = 1, \ldots, 13$ per decision in {use platform, don't}. (We have found that using 2000 draws per period rather than 4000 does not affect our estimates). For each grid point (m, t), we draw a set of on-platform units, and a set of off-platform units whose number is drawn $Poisson(\lambda_0)$. For each decision $d \in$ {search, don't}, this gives us a matrix $x_{m,t,d}$ and vector $q_{m,t,d}$. In addition, we draw a vector of iid shocks $e_{0,m,t,d} \sim N(0, 1)$.

We use these grids to construct a differentiable approximation to the value function. To describe how we do this, we begin with the value of a set of inquiries, then work backwards.

Drop i subscripts for convenience, and fix $V_{T+1} = 0$. Let V_t denote the value at start of period t. The Gumbel distribution of ϵ^1 implies that, conditional on set of inquiries, a household accepts unit j in period t with probability

$$P_{ijt} = \frac{\exp(v_{ij} - V_{t+1})}{1 + \sum_{\text{inquired at } j} \exp(v_{ij} - V_{t+1})}.$$

Moreover, it implies that the value of a set of inquiries is given by:

$$V_{t+1} + \log\left(1 + \sum_{\text{inquired at } j} \exp(v_{ij} - V_{t+1})\right).$$

To compute the value of discovering a set of properties with characteristics (x, q, e) given next-period value V_{t+1} and parameters θ which include the type-specific values of β_x, β_q, and σ_ϵ, we complete the following procedure:

1. Compute utilities

$$\{v_{ij}(\theta)\}_{j=1,\ldots,n_{m,t}} = x\beta_x + q\beta_q + \sigma_\epsilon e.$$

2. Sort them.

3. For each $k \leq n_{m,t}$ and decision d about platform use, compute the value of inquiring at the top k units:

$$U_{m,t,d}^{k}(\theta) = V_{t+1}(\theta) + \log\left(1 + \sum_{j=1,\ldots,k} \exp(v_{ij}(\theta) - V_{t+1}(\theta))\right).$$

4. Find cutoff values of $c_{inquire}^{k,k+1}$ at which $U_{m,t}^{k}(\theta) = U_{m,t}^{k+1}(\theta)$.

5. Let $c^{n_{m,t},n_{m,t}+1} = -\infty$.

6. The value of the choice set is given by:

$$U_{m,t,d}(\theta) = \sum_{k=1}^{n_{m,t,d}} \left(F_{inquire}(c^{k-1,k}(\theta) - F_{inquire}(c^{k,k+1}(\theta)\right) E\left(U_{m,t,d}^{k}(\theta) - c^{inquiry}|c^{inquiry} \in (c^{k,k+1}(\theta), c^{k-1,k}(\theta))\right)$$

In following these steps, we also compute the probability of inquiring at each set of k top units. We use these, together with the probabilities of accepting a unit, to compute differentiable expressions for the moment contributions of cell (m, t, d) conditional on accepting a unit, the moment contribution of this cell conditional on not accepting a unit, the probability of accepting a unit given draw (m, t, d), and the value $U_{m,t,d}$ of each draw of viewed properties (m, t) for $d \in 0, 1$, where 1 denotes platform use.

To compute the value in period t given platform-use decision d, we take the average of $U_{m,t,d}$ accounting for importance sampling in the number of off-platform listings discovered:

$$V_t^d(\theta) = \sum_m w_{m,t}(\theta) U_{m,t}^{(d)}(\theta), \quad w_{m,t} \propto F(n_{m,t}^{\text{off}}|\lambda)/F_0(n_{m,t}^{\text{off}}).$$

We use the same weights to aggregate the moment contributions from the various cells.

The value function at the start of period t is then given by:

$$V_t = V_t^0 + E(V_t^1 - V_t^0 - c_{search}|c_{search} < V_t^1 - V_t^0)F_{search}(c_{search}|\mu_{search}, \sigma_{search}^2).$$

When costs are lognormal, this expression has a (smooth) closed form. We induct backwards until we reach the initial period.

We start with an initial parameter vector θ_0, including an initial value of λ, denoted λ_0. We then draw the number of off-platform listings in each grid point according to

$$F_0(n^{\text{off}}_{m,t}) = Pr(Poisson(\lambda_0) = n^{\text{off}}_{m,t}).$$

We start with a diagonal weight matrix W_0, with diagonal elements proportional to the inverse variance of the corresponding moment in the data. (This choice effectively rescales all moments to have variance 1 in the data.)

We then obtain initial estimates θ_1, including an estimated arrival rate λ_1, which we use to redraw the grid. We estimate again, taking θ_1 as a starting value and again using weight matrix W_0, obtaining a new estimate θ_2 including λ_2.

We then redraw the grid using λ_2, and estimate the optimal weight matrix using the consistent estimate θ_2. Call this estimated weight matrix W_1. We compute the GMM estimator with weight matrix W_1, obtaining an estimate θ_3, including an arrival rate λ_3. We then recompute the optimal weight matrix (call this matrix W_2), redraw the grid using λ_3, and obtain an estimate, θ_4, which we report. (We find little difference between θ_3 and θ_4.)

To compute the GMM estimator in each step, we use the BOBYQA algorithm provided by the NLOpt software package. This is a gradient-free local optimization algorithm which uses past function evaluations to compute a local quadratic approximation to the objective function.