

Estimating Neighborhood Choice Models: Lessons from a Housing Assistance Experiment[†]

By SEBASTIAN GALIANI, ALVIN MURPHY, AND JUAN PANTANO*

We use data from a housing-assistance experiment to estimate a model of neighborhood choice. The experimental variation effectively randomizes the rents which households face and helps identify a key structural parameter. Access to two randomly selected treatment groups and a control group allows for out-of-sample validation of the model. We simulate the effects of changing the subsidy-use constraints implemented in the actual experiment. We find that restricting subsidies to even lower poverty neighborhoods would substantially reduce take-up and actually increase average exposure to poverty. Furthermore, adding restrictions based on neighborhood racial composition would not change average exposure to either race or poverty. (JEL I32, I38, R23, R38)

Sorting models have been used extensively in economics to model household location decisions. Building on earlier theoretical work,¹ there has been a large recent empirical literature that employs the sorting framework to estimate preferences and the marginal willingness to pay for a host of public goods and amenities such as school quality, crime, pollution, and the attributes of one's neighbors.² These

*Galiani: Department of Economics, University of Maryland, Tydings Hall, College Park, MD 20742 (e-mail: galiani@econ.umd.edu); Murphy: Department of Economics, Arizona State University, PO Box 879801, Tempe, AZ 85287 (e-mail: Alvin.Murphy@asu.edu); Pantano: Department of Economics, Washington University in St. Louis, Campus Box 1208, One Brookings Drive, Saint Louis, MO 63130 (e-mail: pantano@wustl.edu). We thank Brian Feld, Raul Sanchez, Ivan Torre, Bo Wang, and Dan Weiss for their valuable research assistance. We also thank David Albouy, Pat Bayer, Kelly Bishop, Flavio Cuhna, Dennis Epple, Limor Golan, Hanming Fang, Fernando Ferreira, Maria Marta Ferreyra, Judy Geyer, Vernon Henderson, Paul Joice, Matt Kahn, Jeff Kling, Nick Kuminoff, Maurizio Mazzoco, Costas Meghir, Steve Ross, Holger Sieg, Paulo Somaini, Kerry Smith, Chris Taber, Steve Tadelis, Giorgio Topa, Matt Wiswall, Ken Wolpin, two anonymous referees, conference and seminar participants at UPenn, New York Fed, ITAM, Oxford, Vanderbilt, Collegio Carlo Alberto, Tsinghua University, Pompeu Fabra, Toulouse, Cowles Foundation, NBER, HULM, Universidad de San Andres, Rimini, Pittsburgh, UIC, Cleveland Fed, Arizona State University, Xiamen University, 2012 Summer Meetings of the Econometric Society in Evanston, IL, and 2013 ASSA Meetings in San Diego, CA for helpful comments. Funding from the Center for Social Development and the Weidenbaum Center, both at Washington University in St. Louis, is greatly appreciated. Data used in this article was obtained through restricted-access data use agreements with the Inter-University Consortium for Political and Social Research (ICPSR) and the US Department of Housing and Urban Development (HUD). The views in this article do not necessarily reflect those of HUD. The authors declare that they have no relevant or material financial interests that relate to the research described in this paper.

[†]Go to <http://dx.doi.org/10.1257/aer.20120737> to visit the article page for additional materials and author disclosure statement(s).

¹For important theoretical contributions, see Ellickson (1971); Epple, Filimon, and Romer (1984); Epple and Romer (1991); Epple and Romano (1998); and Nechyba (1999, 2000).

²See, among others, Epple and Sieg (1999); Sieg et al. (2004); Bayer, McMillan, and Rueben (2004); Bayer, Ferreira, and McMillan (2007); Ferreyra (2007); Walsh (2007); and Kuminoff (2012).

models have been used to evaluate policy as they allow researchers to quantify the benefits and costs of various policy interventions.

While the recent empirical literature has made many advances, we contribute to the literature by using experimental data to estimate and validate a location choice model.³ A key parameter in these models is the marginal utility of consumption, which is typically recovered as the coefficient on consumption, where consumption is naturally determined by price. This parameter is crucially important as it is necessary to estimate the marginal willingness to pay for amenities, as well as to evaluate many types of policy proposals. However, there exists a fundamental endogeneity problem as housing prices are typically correlated with a location's unobserved attributes. While the literature has developed many clever instrumentation strategies, these strategies are typically derived directly from the model.⁴

In this paper we estimate a model of neighborhood choice using data from the Moving to Opportunity (MTO) experiment. We use random variation in the rents which households face to estimate our model.⁵ The unique features of these data allow us to validate our model with out-of-sample measures of fit. Finally, we are able to decompose the effects of the policy experiment and simulate the effects of interesting alternative policies.

The starting point for our analysis is data from the MTO experiment. The MTO data provide details on the demographic characteristics and location choices made by households placed into one of three random assignment groups: a control group; a treatment group given mobility counseling and housing subsidies that were conditioned on moving to a low-poverty neighborhood; and a treatment group that was given unrestricted housing subsidies with no counseling. The MTO data have been previously used to estimate the effect of the MTO intervention on labor market and other outcomes, as well as estimating neighborhood effects.⁶ To our knowledge, we are the first to leverage these data to estimate a model of neighborhood choice.

The two-treatment experimental data from the MTO experiment provide a unique opportunity to pursue our research question. Usually when combining structural estimation with experimentally generated data, the econometrician may either exploit the rich experimental variation to identify and estimate the model's parameters, or estimate the model using the control group data only and then validate the model by predicting the outcomes observed in the treatment group data.⁷ As we have two separate treatment groups, we are able to do both; we use one treatment group (together

³The closest the urban literature has come to using experimental data is Wong (2013), who estimates ethnic preferences by cleverly exploiting ethnic housing quotas in Singapore as a natural experiment. Similarly, Bayer, Ferreira, and McMillan (2007) embed the Black (1999) regression discontinuity design in their sorting model to measure preferences for school quality. Using data from Michigan, Ferreyra (2009) uses a large nonexperimental policy change to validate a model of location and school choice.

⁴For example, Bayer, Ferreira, and McMillan (2007) use the equilibrium prices predicted by the model based only on exogenous attributes as an instrumental variable.

⁵As discussed in greater detail below, while randomization is important for identification of the parameter that characterizes the marginal utility of consumption, the identification of other parameters is similar to the existing literature.

⁶See, among others, Katz, Kling, and Liebman (2001); Kling, Ludwig, and Katz (2005); and Kling, Liebman, and Katz (2007).

⁷If the model does a successful job at reproducing the experimental data, the researcher can be more confident in using the model to simulate alternative policies.

with the control group) for estimation of the location-choice model and reserve the other treatment group for out-of-sample validation.⁸

We are then able to address various important policy questions. In particular, we are able to (i) disentangle the separate quantitative roles of two features of the actual experimental treatment; (ii) examine the impact of changing one of the key features of the experiment; and (iii) consider the consequences of adding race-based location constraints on the use of housing subsidies. Given the nature of our model, we can evaluate these alternative policies by simulating their associated neighborhood choice patterns and subsidy take-up rates.

In addition to having location restrictions on subsidy use, the experimental treatment group received mobility counseling, which among other things trained households so they could eventually have more successful interviews with landlords. Barring further experimentation, the effects of bundled randomized treatments, like the combination of mobility counseling and location restrictions, cannot be disentangled without relying on a model. Theoretically, location restrictions should reduce the subsidy take-up rate and mobility counseling should increase it. In the MTO data, the treatment group which receives both mobility counseling and location restrictions is approximately 13 percent less likely to use the subsidy compared with the group assigned the unrestricted subsidy and no mobility counseling. With our parameter estimates we can disentangle the two effects, and we find that location restrictions alone (i.e., not supplemented by counseling) would reduce subsidy take-up by 47 percent.

We find that changing the maximum allowed poverty rate of the destination neighborhood (in the restrictions for subsidy use) has a large impact on take-up rates. For example, only 16 percent of households would use the subsidy under a more stringent restriction that limits subsidy use to neighborhoods with a poverty rate under 5 percent. An important implication of this is that more stringent location constraints designed with the goal of exposing the target population to *lower* neighborhood poverty rates could end up backfiring. In our simulations, assigned households (including those who decide not to use the voucher) end up exposed, on average, to *higher* neighborhood poverty rates because of their lower subsidy take-up.

Finally, our desegregation experiment considers further limiting where households can move to based on the racial composition of the destination neighborhoods. We find that, compared with the MTO experimental subsidy, the alternative policy that supplements poverty-based constraints with race-based constraints would, on average, expose households to the same neighborhood characteristics but would lower the subsidy take-up rate.

⁸Structural estimation combined with (and disciplined by) experimentally generated data can be quite useful for policy evaluation. Indeed, one of the earliest applications of this approach was actually in the field of housing subsidies. Wise (1985) exploits a housing subsidy experiment to evaluate a model of housing demand. Todd and Wolpin (2006) estimate a structural model of school attendance using only control observations from the randomized evaluation of the PROGRESA intervention. They use the treatment group for validation purposes by examining whether simulation of treatment using the estimated model can replicate the observed pattern of behavior for the treatment group in the interventions. Attanasio, Meghir, and Santiago (2012) also use data from PROGRESA but argue that instead of using it for validation, it is important to exploit the exogenous variation induced by the experiment for estimation purposes. Another example of work combining a structural model and experimental data is Duflo, Hanna, and Ryan (2012).

The remainder of the paper proceeds as follows. Section I discusses the MTO program and the data. Our model is outlined in Section II and Section III describes the estimation strategy and results. We present the model fit and validation exercises in Section IV and policy evaluations in Section V. Finally, Section VI concludes.

I. Experimental Background and Data

A. *The Moving to Opportunity Experiments*

In the mid-1990s, the Department of Housing and Urban Development (HUD) along with the public housing authorities (PHAs) in five metropolitan areas (Baltimore, Boston, Chicago, Los Angeles, and New York City) carried out the Moving To Opportunity experiments. The main objective of MTO was to evaluate the role that neighborhoods play in shaping various outcomes for low-income households receiving housing assistance. Within each PHA's jurisdiction, eligible households living in public or project-based housing were allowed to enroll and participate in the experiment. These households were randomly assigned to one of three groups.

The first group was a pure control group that continued to receive public housing assistance in public housing projects. We refer to this group as the control group. The second group was an experimental treatment group that received restricted tenant-based Section 8 rental assistance. The Section 8 subsidies could only be used in areas with less than 10 percent poverty. This group also received training sessions which, among other things, helped them do a better job interviewing landlords about potential rental units outside the public housing project. We refer to this group as the experimental group. The third group was a treatment group that received the standard, unrestricted Section 8 subsidies. In this case the subsidies could be used without any location constraints. Like the control group, this group did not receive any mobility counseling. We refer to this group as the Section 8 group.

Random assignment of households started in 1994 and continued through 1998. Each household completed a baseline interview at time of random assignment. A follow-up was conducted in 2001. Upon receiving a subsidy offer, a household in the Section 8 group planning to use the subsidy was given 90 days to find an apartment and sign a lease. Households in the experimental group were given an additional month to find an apartment complying with the location constraint and were required to stay in the low-poverty area for at least one year. They were allowed to use the subsidy in an unrestricted way after that.⁹

Most of the research on the impact of the MTO experiments has focused on experimental-control comparisons and as such, has carefully estimated intent-to-treat (ITT) and treatment-on-the-treated (TOT) parameters. See, for example, Katz, Kling, and Liebman (2001); Ludwig, Duncan, and Hirschfield (2001); and Kling, Ludwig, and Katz (2005).¹⁰

⁹A dynamic model that captures this option value would be needed to formally capture this feature of the experimental subsidy. In order to keep the model tractable, we abstract away from this type of forward-looking behavior.

¹⁰In addition to published academic articles discussed here, excellent summaries and policy-oriented compilations of this research can be found in the volume edited by Goering and Feins (2003) on early site-specific findings and in the interim evaluation report by Orr et al. (2003).

In an earlier paper, Katz, Kling, and Liebman (2001) exploit the variation generated by the MTO experiment in Boston. They document that baseline characteristics were similar for all groups, indicating a successful randomization. A year after randomization however, those who had moved lived in very different areas than those who had not. They show that, as expected, the experimental treatment was more successful than the unrestricted Section 8 treatment in relocating poor families into low-poverty areas. On the other hand, the unrestricted Section 8 assistance was more effective in getting more families out of the public housing projects (i.e., unrestricted subsidies had a higher take-up rate).

Kling, Liebman, and Katz (2007) moved beyond estimation of ITT and TOT parameters and examined the question of estimating neighborhood effects using the MTO experiment. In particular, they examined the relationship between a neighborhood's poverty rate and various outcomes.¹¹ They found that a neighborhood with lower poverty rates improves mental health outcomes and has gender-specific effects on youth risky behavior (with reductions for females and increases for males).¹²

An important feature of the experiments is the take-up rate of the subsidies. Shroder (2003) documents that the rate at which the subsidy was actually used by the experimental group was lower than the one from the unrestricted Section 8 group, despite the fact that experimental households received mobility counseling. Shroder concludes that location constraints had strong effects and trumped the positive effects of counseling.¹³ Below we use our structural model to disentangle the separate roles of counseling and location restrictions.

B. *The Data*

Our main datasets contains information for all households in the interim evaluation sample from the MTO experiment. This information was collected in a follow-up survey conducted in 2001. In addition, we have information collected at baseline for each of these MTO households.¹⁴ In this paper, we focus on data from Boston. The MTO microdata provide us with initial location, neighborhood choice, household demographic characteristics (e.g., race, household size, marital status), household income, random assignment group, subsidy take-up decision, and indicators of propensity to move out of the public housing project (e.g., whether they are dissatisfied with the neighborhood, whether the household had moved at least three times in the last five years, whether the household had applied for Section 8 subsidies in the past). For households who use the subsidy offered through MTO, we

¹¹ Given the endogeneity problems induced by residential choices, the model was estimated by two-stage least squares (2SLS) using a full set of site-by-treatment interactions as the excluded instruments for the neighborhood poverty rate in the first stage.

¹² See also Aliprantis (2011) for a reanalysis of these findings. Beyond MTO, Jacob and Ludwig (2012) analyzed a different housing voucher experiment and found that housing assistance has a negative effect on labor supply and earnings.

¹³ Shroder (2003) pooled data from all the five MTO sites and introduced site effects in his logit models of take-up. Even when some sites like Boston had only one counseling agency, the effect of counseling could then be identified in Shroder (2003) by allowing for a parametric relationship between the intensity of counseling services and the probability of voucher use. Baltimore and Boston only had one counseling agency, whereas the larger sites (Chicago, Los Angeles, and New York) each had two. See also Feins, McInnis, and Popkin (1997) for ratings of counseling intensity across MTO agencies.

¹⁴ For more details on the MTO data, see Orr (2011).

TABLE 1—MTO DATA DESCRIPTIVE STATISTICS

	Control	Experimental	Section 8	Total
White	0.07	0.11	0.10	0.09
Household income (\$1,000s)	11.8	11.9	11.3	11.7
Never married	0.63	0.64	0.69	0.65
Household size	3.38	3.07	3.26	3.23
Applied to Section 8 before	0.56	0.52	0.61	0.56
Moved three times before	0.15	0.15	0.15	0.15
Dissatisfied with neighborhood	0.30	0.33	0.27	0.30
Observations	165	204	172	541

Notes: Final analysis sample from Boston. Single headed households enrolled in the MTO demonstration. Variables in the table are measured at baseline. Annual household income in 1,000s of 1997 US\$ includes welfare payments for those on welfare and estimated labor income for those working. See text for details.

observe the neighborhood where they use the subsidy and, for those observations, we treat this as our measure of neighborhood choice. For those households who do not use a subsidy, we use the neighborhood of residence in 2001. One of the key features of the subsidy is the fair market rent (FMR) which determines the amount of rent a household must pay.¹⁵

After cleaning the data, we end up with a final sample of 541 observations. Appendix A provides more details regarding the variables used in our analysis and the selection of our final sample. Table 1 presents descriptive statistics. As can be seen in the table, the data from our final analysis sample retain good covariate balance, preserving the value of randomization across control, experimental, and Section 8 groups.

We also exploit data from the 2000 population census. In particular, we use Summary File 3 data to create neighborhood characteristics (percent white and distance to jobs).¹⁶ From Summary File 4, we obtain data on renter counts by income bracket and race for each neighborhood. By reweighting these data, we can compute neighborhood shares for a population with characteristics similar to the MTO sample, based on renter status, race, and income.¹⁷ These shares are a key input to our estimation strategy that controls for unobserved neighborhood attributes. To obtain neighborhood poverty rates, we rely on 1990 population census data at the six-digit census tract level. These were the numbers that were checked against to verify compliance with the poverty-based location constraint.

The other data sources we use are DataQuick transactions data on housing sales in Boston, which we use to compute neighborhood level price indexes, and the

¹⁵Since 1995, FMR is set at the fortieth percentile of the rents in the metropolitan area. The effective FMR is different for different households depending on their characteristics because there are different FMRs for housing units with different number of bedrooms. We use the Boston FMRs from 1997 and assign the two-bedroom FMR to two- or three-person households, the three-bedroom FMR to four-person households, and the four-bedroom FMR to households with five or more members.

¹⁶We also retrieve the median rent in each of the neighborhoods in which MTO households were originally located. Since in our model households can choose different quantities of housing, we do not use these median rents as measures of market rent. We do use them to normalize the housing endowment, (i.e., the quantity of housing consumed at baseline for MTO households living in public housing projects).

¹⁷See Appendix A for reweighting details.

5 percent census micro-level data from IPUMS for Boston, which we use to compute the share of income spent on housing.¹⁸

For our model and estimation approach, we define neighborhoods as six-digit census tracts and the choice set includes 585 six-digit census tracts in the Boston primary metropolitan statistical area.¹⁹ Many of these census tracts are not chosen by MTO participants. For Boston, the post-treatment distribution of households across census tracts is very dispersed. MTO households ended up scattered over 137 census tracts in Boston. Initially, however, they were distributed in a more narrow set of 25 census tracts, essentially corresponding to the census tracts in which the targeted public housing projects were located.

Before going to the model we briefly document the patterns of take-up in the sample. Table 2 presents the results from estimating the following linear probability model of take-up where D_i denotes take-up (i.e., use of the subsidy), G_i denotes assignment group, and \mathbf{Z}_i denotes demographic characteristics of the households,

$$(1) \quad D_i = \alpha_0 + \alpha_1\{G_i = \text{Experimental}\} + \alpha_2\{G_i = \text{Section 8}\} + \mathbf{Z}_i'\beta + u_i.$$

As can be seen in Table 2, the take-up rate for the Section 8 group is substantially higher than for the experimental group. There is an 8-percentage-point gap (55 percent versus 63 percent) in take-up rates.²⁰ This suggests that restrictions on location outweigh any positive effect that mobility counseling may have had. Note, however, that we are only able to observe their combined effects and cannot identify their independent magnitudes.

Finally, to appreciate the value of imposing structure, it is worthwhile considering what data would be needed otherwise. With an infinite budget for experiments, we would want to create several experimental groups, each with varying restrictions on the destination neighborhoods and with different experimental arms with and without counseling. This would allow us to estimate take-up rates separately for each possible unique treatment. Without access to these ideal data, we alternatively specify a structural model of neighborhood choice, estimate the structural parameters of the model with data from the control group and the experimental group, and externally validate the model with data from the Section 8 group. With estimates of the structural model in hand, we can simulate the effect of other policies not implemented during the experiment.

Our contribution lies in emphasizing a rather unexplored use of the experimental data generated by MTO. Our aim is to leverage the data to credibly estimate parameters that are the key inputs to a set of counterfactual policy experiments. Our counterfactual simulations ultimately allow us to get a sense of what the effect of other feasible policies would be without incurring the cost and time involved in running new experiments.

¹⁸For details on IPUMS data, see Ruggles et al. (2010).

¹⁹Defining neighborhoods at the census tract level is the natural choice as this was the definition used by the MTO program officers to determine eligibility for the experimental subsidy. We only include counties for which we can construct the house price index. These counties are Suffolk (which includes the city of Boston), Norfolk, Middlesex, and Essex. Appendix A provides more details on the construction of the choice set.

²⁰Given randomization, controlling for covariates in the second column makes no difference to the results.

TABLE 2—SUBSIDY TAKE-UP

	(1)	(2)
Experimental	0.554*** (0.0349)	0.540*** (0.0348)
Section 8	0.634*** (0.0368)	0.624*** (0.0374)
White		0.142** (0.0557)
Household income		0.0174 (0.0253)
Never married		0.0227 (0.0367)
Household size		-0.0253* (0.0147)
Applied to Section 8 before		0.110*** (0.0351)
Moved three times before		0.0969** (0.0451)
Dissatisfied with neighborhood		0.127*** (0.0363)
Constant		-0.0734 (0.0716)
Observations	541	541

Notes: Boston MTO final analysis sample. The dependent variable is equal to 1 if the household uses the subsidy, and equal to zero otherwise. Control group observations are the omitted category but they were not given subsidies so their dependent variable is always zero, and the regression without covariates in column 1 goes through the origin. Robust standard errors in parentheses.

***Significant at the 1 percent level.

**Significant at the 5 percent level.

*Significant at the 10 percent level.

II. The Model

Our model falls into the broad framework of empirical urban sorting models. We use a discrete choice approach that allows for unobserved attributes for each neighborhood.²¹ While this literature has been well established, the use of these models to study either renter behavior or housing assistance policy is in its infancy. An example is Geyer (2011), who uses data from Pittsburgh to study housing assistance policy.²² The primary difference between the approach taken here and previous sorting models is our use of experimental data.

²¹Following earlier work by McFadden (1974), the literature on discrete choice significantly gained in popularity after Berry (1994) and Berry, Levinsohn, and Pakes (1995) hereafter, BLP showed how to allow for unobserved product characteristics and conduct estimation using aggregate shares of the chosen characteristics. In recent papers, Berry and Haile (2010, 2014) have clarified the conditions for identification of these BLP-type models for cases in which the econometrician has only access to aggregate data and/or microdata. Among other possibilities, they emphasize the need for price instruments such as those used in Waldfogel (2003) for identification. Our work has the potential to contribute to this literature by showing that experimental variation in the price of the alternatives can be exploited to achieve identification.

²²See Geyer and Sieg (2013) for a model which focuses on public-housing assistance rather than subsidy-based assistance, which we focus on here.

Given random assignment $G_i \in \{0, 1, 2\}$ into either the control ($G_i = 0$), experimental ($G_i = 1$), or Section 8 ($G_i = 2$) groups, our model considers households' simultaneous choice of residential neighborhood and consumption of non-housing and housing services.²³ The treated households ($G_i \in \{1, 2\}$) will be simultaneously considering a decision D_i of whether to use the assigned subsidy or not. Households make a neighborhood choice $d_i = j$ according to their preferences for neighborhood characteristics \mathbf{X}_j and household's characteristics \mathbf{Z}_i .²⁴ Within each potential neighborhood, a household must optimally choose how much of their income to allocate to the consumption of housing services. Household utility is maximized subject to both the corresponding budget constraint and the other constraints associated with the rules for subsidy use. Neighborhoods in the model are heterogeneous in both observable and unobservable ways.

Household i 's utility depends on overall household consumption, C_i , which is comprised of nonhousing consumption, Q_i , and housing services consumption, H_i . Utility also depends on observable and unobservable neighborhood attributes, respectively \mathbf{X}_j and ξ_j , household characteristics, \mathbf{Z}_i , and unobserved household-specific taste shocks for each neighborhood, ϵ_{ij} . We denote the vector of preference parameters by θ .

Household i maximizes utility by choosing a neighborhood $d_i = j \in \{1, \dots, J\}$ among the available neighborhoods, including the option of staying in the same public housing unit ($j = j_{t-1}$), at the same time as choosing optimal levels of non-housing and housing services.²⁵ Our model assumes away any problem of lack of information about neighborhood characteristics, which is in common with all previous papers in the literature of residential choice. Households assigned to either the experimental or the Section 8 treatment groups also effectively choose whether to use the subsidy ($D_i = 1$).²⁶

Therefore, households are solving

$$(2) \quad \max_{\{d_i, Q_i, H_i\}} U(C(Q_i, H_i), \mathbf{X}_j, \xi_j, \mathbf{Z}_i, j_{i,t-1}, \epsilon_{ij}, \theta),$$

subject to a budget constraint where the price of nonhousing services is normalized to 1, the out-of-pocket rent payment for housing services is given by R_{ij} , and income is denoted by I_i ,

$$(3) \quad Q_i + R_{ij} = I_i.$$

The out-of-pocket rent function takes as its arguments housing services, H_i ; treatment group assignment, G_i ; an indicator S_i for whether the household receives the housing assistance subsidy in the form of a voucher (v) or a certificate (c);²⁷

²³ As discussed in Section I, a neighborhood is defined as a six-digit census tract.

²⁴ Based on results in Kling, Liebman, and Katz (2007), we assume households anticipate no income differences across neighborhoods.

²⁵ This model follows closely the models presented in Bayer, Keohane, and Timmins (2009) and Geyer (2011).

²⁶ Given the structure of our model, subsidy use D_i maps directly into neighborhood choice. Whenever a Section 8 group household moves, it does so using the subsidy. Therefore, for these households $D_i = 1$ whenever $j \neq j_{i,t-1}$. An experimental group household uses the subsidy whenever it moves to a low-poverty neighborhood, so $D_i = 1$ whenever $j \neq j_{i,t-1}$ and $Pov_j < 10$ percent.

²⁷ See Olsen (2003) for a comprehensive discussion of different housing assistance subsidies.

neighborhood choice (including its rental price of a unit of housing services, r_j , and its poverty rate, Pov_j), j ; baseline neighborhood choice, j_{t-1} ; household income, I_i ; and features of the subsidy program, (σ, ρ_i, τ) . In addition to its format, S_i , the actual subsidy depends on the share of household income that must be paid, σ ; the subsidy cap, ρ_i ;²⁸ and the restriction on a neighborhood's poverty rate, τ .²⁹ For those receiving the voucher, the subsidy amount is equal to $\rho_i - \sigma I_i$.³⁰ This is subject to a constraint that out-of-pocket rent cannot be negative. For those receiving a certificate, the out-of-pocket rent is σI_i as long as the market rent, $r_j H_i$, is less than or equal to the fair market rent cap, ρ_i . If the market rent exceeds the cap, the household must pay the full market rent. For those in the experimental group who move to a neighborhood where poverty exceeds the poverty cutoff, τ , the household must pay the full market rent, regardless of subsidy format, S_i . Figure 1 in Appendix C illustrates the out-of-pocket rent function:

$$(4) \quad R_{ij} \left\{ \begin{array}{ll} \sigma I_i & \text{if } j = j_{i,t-1}, \text{ all } G_i \\ r_j H_i & \text{if } j \neq j_{i,t-1}, G_i = \text{Control} \\ \max\{0, r_j H_i - [\rho_i - \sigma I_i]\} & \text{if } j \neq j_{i,t-1}, G_i = \text{Sec 8}, S_i = v \\ r_j H_i & \text{if } j \neq j_{i,t-1}, G_i = \text{Exp}, S_i = v, Pov_j > \tau \\ \max\{0, r_j H_i - [\rho_i - \sigma I_i]\} & \text{if } j \neq j_{i,t-1}, G_i = \text{Exp}, S_i = v, Pov_j < \tau \\ \sigma I_i & \text{if } j \neq j_{i,t-1}, G_i = \text{Sec 8}, S_i = c, r_j H_i \leq \rho_i \\ r_j H_i & \text{if } j \neq j_{i,t-1}, G_i = \text{Sec 8}, S_i = c, r_j H_i > \rho_i \\ r_j H_i & \text{if } j \neq j_{i,t-1}, G_i = \text{Exp}, S_i = c, Pov_j > \tau \\ \sigma I_i & \text{if } j \neq j_{i,t-1}, G_i = \text{Exp}, S_i = c, Pov_j < \tau, r_j H_i \leq \rho_i \\ r_j H_i & \text{if } j \neq j_{i,t-1}, G_i = \text{Exp}, S_i = c, Pov_j < \tau, r_j H_i > \rho_i \end{array} \right.$$

We parameterize the direct utility function for household i associated with choosing neighborhood j as

$$(5) \quad U_{ij} = C_i^{\beta^C} e^{\mathbf{X}'_j \beta_i^X + \lambda_{ij} 1\{j \neq j_{i,t-1}\} + \xi_j + \epsilon_{ij}}$$

where overall consumption, C_i , is parameterized as

$$(6) \quad C_i = Q_i^{(1-\beta^H)} H_i^{\beta^H}$$

where ϵ_{ij} is i.i.d. Type 1 extreme value, and where $1\{x\}$ is an indicator function that equals 1 whenever x is true and equals zero otherwise. The parameter β^C controls how strongly the household trades off consumption of (Q_i, H_i) against neighborhood

²⁸ At the beginning of the actual MTO implementation the cap ρ_i was the forty-fifth percentile of the distribution of rents in the metropolitan area. Since 1995, the cap is set at the fortieth percentile. These numbers are the fair market rents (FMR). σ is set at 30 percent.

²⁹ We assume that a household in the control group faces the market rental price of housing services, $\{r_j^m\}_{j=1}^J$, if they choose to move. We ignore transfers or reassignments to public housing projects located in different neighborhoods. See Appendix A for more details.

³⁰ In our data, ρ_i is always greater than σI_i .

amenities (\mathbf{X}_j, ξ_j) . Noting that \mathbf{X}_j is a vector of K attributes, we specify the household-specific parameters β_i^X as

$$(7) \quad \beta_{i,k}^X = \beta_{0,k}^X + \beta_{1,k}^X \mathbf{Z}_i,$$

where β_1^X captures how the utility parameters vary with household demographic characteristics, \mathbf{Z}_i .

The moving cost λ_{ij} is specified as

$$(8) \quad \lambda_{ij} = \lambda_0 + \lambda_1 \text{Dist}_{j,j_{i,t-1}} + \lambda_2 1\{G_i = 1\}$$

and it is only paid if the household moves (i.e., if $j \neq j_{i,t-1}$). It is allowed to vary with $\text{Dist}_{j,j_{i,t-1}}$ which is the distance in miles from the original neighborhood, $j_{i,t-1}$, to any alternative neighborhood, j . As those in the experimental group ($G_i = 1$) receive mobility counseling, we allow their moving cost to differ (by amount λ_2) from the baseline moving costs ($\lambda_0 + \lambda_1 \text{Dist}_{j,j_{i,t-1}}$) faced by the other groups.³¹

Conditional on choosing a neighborhood j , agents choose housing services optimally by maximizing (6) subject to the budget constraint (3). Let H_{ij}^* denote household i 's optimal choice of housing services. Optimal consumption is then given by

$$(9) \quad C_{ij}^* = (I_i - R_{ij}(j, G_i, H_{ij}^*))^{(1-\beta^H)} (H_{ij}^*)^{\beta^H}.$$

Plugging in optimal consumption and taking logs allows us to define the log indirect utility function:

$$(10) \quad v_{ij} = \beta^C \log(C_{ij}^*) + \mathbf{X}'_j \beta_i^X + \lambda_{ij} 1\{j \neq j_{i,t-1}\} + \xi_j + \epsilon_{ij}.$$

Employing the definition of the household-specific utility parameters and collecting neighborhood-level effects into the fixed effect, δ_j , allows us to rewrite v_{ij} as

$$(11) \quad v_{ij} = \beta^C \log(C_{ij}^*) + \delta_j + \mathbf{X}'_j \beta_1^X \mathbf{Z}_i + \lambda_{ij} 1\{j \neq j_{i,t-1}\} + \epsilon_{ij},$$

where δ_j is given by

$$(12) \quad \delta_j = \mathbf{X}'_j \beta_0^X + \xi_j.$$

To complete the model, we need to solve for the optimal level of housing services, H_{ij}^* . If households stay in their existing neighborhood, they have no choice over the level of H_{ij} to consume and must consume their endowment, H_{ij}^e , regardless of assignment category. For control group movers and those in the experimental group who move to a neighborhood where the poverty rate, Pov_j , exceeds the

³¹ We think of mobility counseling as providing mobility skills. While we cannot identify the exact mechanisms by which mobility counseling operates to reduce moving costs, our aim is to account for mobility counseling in order to avoid making incorrect inference about the importance of the location constraints.

allowable threshold, τ , the relevant budget constraint is $R_{ij} = r_j H_i$. In this case, maximizing (6) subject to (3) yields the standard Cobb-Douglas optimal level of housing services,

$$(13) \quad H_{ij}^* = \beta^H \frac{I_i}{r_j}.$$

For experimental group movers who comply with the restriction ($\text{Pov}_j < \tau$) and for all Section 8 group movers, the relevant budget constraint when receiving the subsidy in the form of a voucher is $R_{ij} = \max\{0, r_j H_i - [\rho_i - \sigma I_i]\}$. In this case, optimal housing services are given by

$$(14) \quad H_{ij}^* = \max \left\{ \beta^H \frac{(1 - \sigma) I_i + \rho_i}{r_j}, \frac{\rho_i - \sigma I_i}{r_j} \right\}.$$

Finally, for those receiving a certificate, it will always be optimal for the household to choose housing services such that the rent is exactly equal to the certificate value; choosing lower levels does not reduce rent (i.e., it remains at σI_i) and choosing higher levels results in forfeiture of the subsidy. Therefore,

$$(15) \quad H_{ij}^* = \frac{\rho_i}{r_j}.$$

The (log) indirect utility functions can then be calculated by combining the optimal housing service expressions with (9) and (11). These indirect utility functions are provided in Appendix B. Overall, the model provides a rich representation of household residential mobility decisions and captures how those mobility decisions may be influenced by housing assistance policy parameters.

III. Estimation

A. Estimation Overview and Identification Strategy

To estimate the model, we develop a novel estimation approach that makes use of both the experimental data provided by MTO and the large-sample nature of US census data. This approach allows us to identify the marginal utility of consumption using the experimental data while still controlling for unobserved neighborhood attributes using the census data. We first discuss the identification of the marginal utility of consumption parameter, β^C , where our approach differs from the existing literature and where we rely heavily on the randomization created by the experiment. We then discuss the identification of the other parameters of the model, where we follow closely standard existing approaches.³²

³² See, for example, Berry (1994); Berry, Levinsohn, and Pakes (1995, 2004); Epple and Sieg (1999); Sieg et al. (2004); Bayer, Ferreira, and McMillan (2007); Ferreyra (2007); Walsh (2007); Kuminoff (2012); Bayer, Keohane, and Timmins (2009); Bayer, McMillan, and Rueben (2004); Geyer (2011); and Wong (2013).

As it is variation in consumption levels that identifies the marginal utility of consumption parameter, β^C , it is instructive to consider what generates variation in consumption. We interpret the MTO randomization as providing purely random variation in the out-of-pocket rental prices that households face across neighborhoods. When considering moving, households in the control group face the market rent in each neighborhood. The experimental group faces a reduced rent in some neighborhoods (i.e., the ones that satisfy the location constraint). The experiment therefore generates variation in consumption across assignment groups. Holding household characteristics fixed, for any neighborhood with poverty higher than the threshold, consumption is the same across assignment groups. However, for any neighborhood with poverty lower than the threshold, the subsidy ensures that consumption is higher for the experimental assignment. This random variation in rents (and therefore consumption) is key to identifying the key parameter β^C without relying on the typical model-based exclusion restrictions that are necessary to form instruments. As R_{ij} is randomly different for the control and experimental group participants, their neighborhood-specific consumption levels will differ and we would expect the two groups to make different location decisions. This difference in location decisions identifies the marginal utility of consumption coefficient, β^C .

It is worth noting that consumption does also vary within assignment groups. We can decompose this within-group variation in consumption into two sources. The first is the variation in log consumption across neighborhoods, conditional on moving. This variation is not used for identification as, conditional on moving, log consumption factors into an individual-specific constant (which does not affect location decisions) and a neighborhood-specific price term (which is perfectly colinear with the neighborhood fixed effect).³³ The second within-group source of variation is the fact that consumption will differ between remaining in the initial location and moving to another neighborhood. This variation in log consumption does help identify β^C , however, this source is available only because we observe both the initial location and the subsequent choice of location.³⁴ If we did not observe this feature of the data and used only within-group variation, β^C would be subsumed into the fixed effect.³⁵

Turning to the other parameters of the model, the MTO microdata reveal how the location decisions vary with demographic characteristics. Therefore, we are able to identify β_1^X , the parameters governing how individual characteristics affect preferences for neighborhood attributes.

The propensity to move in the control group identifies the baseline moving cost parameter λ_0 and how the propensity to move varies with distance identifies λ_1 . As we observe a different propensity to move across assignment groups, we can also

³³This can be seen in equations (30) and (33) in Appendix B.

³⁴More specifically, the dependence of the decision to move on income and housing-service levels in the initial location helps identify β^C . When we only use this variation for estimation (i.e., by estimating using the control group only), we get an imprecise and insignificant estimate for β^C , which suggests that the experimentally generated variation is the key source of identifying power. See Section IIIC for more details.

³⁵Therefore, if our data were less rich, β^C (along with neighborhood housing rental prices) would be subsumed into the fixed effect and we could decompose the fixed effect in exactly the same way as the previous literature. However, we can take advantage of the fact that given our data, β^C is not subsumed into the fixed effect. As such, our data provide an alternative way to estimate the marginal utility of consumption parameter. Of course, nothing about this logic invalidates the existing approaches for estimating the marginal utility of consumption.

identify how moving costs differ for the experimental group, which is captured by the parameter λ_2 .

To control for unobserved neighborhood attributes, we rely on a data strategy that combines the MTO microdata with US census aggregate data. The census data provide the joint distribution of demographic attributes and neighborhood choices among renters in the Boston metropolitan area. A key component of the estimation is that the location shares predicted by the model must match the empirical shares found in the census, which identifies δ_j .³⁶

B. Estimation Details

Optimal housing services depend on β^H , which is the relative utility weight on housing services in the Cobb-Douglas specification for overall consumption given in (6). We follow Bayer, Keohane, and Timmins (2009) and set this parameter equal to the median share of income spent on housing services. To do this, we use the 5 percent census microdata to create a sample of households located in our choice set with the same distribution of income and race as in the MTO sample. We set β^H equal to 0.312, which is the median share of income spent of housing in this sample.

To estimate r_j , we follow the approach of Sieg et al. (2002). Letting P_n denote sales price, letting \mathbf{W}_n denote a vector of house attributes for house n , and parameterizing $\log(H_n) = \mathbf{W}_n' \gamma$, we recover r_j by estimating the following equation using a dataset containing housing transactions in the Boston primary metropolitan statistical area (PMSA) between 1988 and 2009:³⁷

$$(16) \quad \log(P_{nj}) = \log(r_j) + \mathbf{W}_n' \gamma + \nu_{nj}.$$

As the level of housing services that must be consumed if a household doesn't move, H^e , is not observed, we assume that it is proportional to the median level of housing services in that neighborhood, where the median level of housing services can be recovered as $\frac{\text{medianrent}_j}{r_j}$.³⁸

The main estimation routine then proceeds in two steps. In the first step, the parameter vector, θ , is chosen to maximize the log-likelihood of observing the MTO data, subject to a constraint that the model's predicted shares must match those found in the census. Note that in addition to β^C , β_1^X , λ_0 , λ_1 , λ_2 , the vector of location-specific fixed effects, δ , is estimated in this initial step. In the second step, these δ are decomposed into a function of the observable neighborhood characteristics as given by equation (12), which allow us to recover the remaining parameters, β_0^X .

Letting N denote the number of MTO observations, the probability that household i chooses location j when receiving housing subsidies s is given by π_{ij}^s .

³⁶ As discussed below, we estimate β_0^X by using instrumental variables (IV) to decompose δ_j . This is done similar to Bayer, Ferreira, and McMillan (2007) and Geyer (2011).

³⁷ \mathbf{W}_n includes age, lot size, house size, number of bathrooms, number of bedrooms, number of stories, number of units, and year dummies.

³⁸ We normalize the constant of proportionality to 0.5.

$$(17) \quad \pi_{ij}^s = \frac{\exp(\beta^C \log(C_{ij}^*(s)) + \delta_j + \mathbf{X}'_j \beta_1^X \mathbf{Z}_i + \lambda_{ij} 1\{j \neq j_{i,t-1}\})}{\sum_{k=1}^J \exp(\beta^C \log(C_{ik}^*(s)) + \delta_k + \mathbf{X}'_k \beta_1^X \mathbf{Z}_i + \lambda_{ik} 1\{k \neq j_{i,t-1}\})}$$

where $C_{ij}^*(s)$ is the optimal consumption that household i would receive from choosing neighborhood j when receiving the subsidy with format s . Recall that $s = v$ for vouchers and $s = c$ for certificates. Since the format of housing assistance is unobserved, we integrate over it by letting $\pi_{ij} = \pi_{ij}^v \Pr\{S = v\} + \pi_{ij}^c \Pr\{S = c\}$.³⁹ The first estimation step finds the vector $\theta = (\beta^C, \beta_1^X, \lambda_0, \lambda_1, \lambda_2, \{\delta_j\}_{j=1}^J)$ that solves the following problem:

$$(18) \quad \max_{\theta} \sum_{i=1}^N \sum_{j=1}^J \log(\pi_{ij}) 1\{d_i = j\},$$

subject to

$$(19) \quad \pi_j^{cen}(\theta) = \pi_j^{cen} \quad \forall j,$$

where π_j^{cen} is the empirical share of households who choose neighborhood j in the census data and $\pi_j^{cen}(\theta)$ is the model prediction for this share based on a given parameter guess θ . To compute π_j^{cen} we reweight the census microdata to ensure that the distribution of race and household income matches that of MTO households.⁴⁰

For each trial of $(\beta^C, \beta_1^X, \lambda_0, \lambda_1, \lambda_2)$, the constraint fully determines the value of $\{\delta_j\}_{j=1}^J$. Finding the values of $\{\delta_j\}_{j=1}^J$ that satisfy the constraint can be done quickly using the following contraction mapping

$$(20) \quad \delta_j^{m+1} = \delta_j^m + \log(\pi_j^{cen}) - \log(\pi_j^{cen}(\delta^m)) \quad \forall j,$$

where for a given δ , the predicted share of neighborhood j is given by the model as

$$(21) \quad \pi_j^{cen}(\delta) = \sum_{\mathbf{z}} \pi_j^{cen}(\delta, \mathbf{z}) \Pr(\mathbf{Z} = \mathbf{z}).$$

The probability of household i choosing a neighborhood j , $\pi_{ij}^{cen}(\theta)$, is formed in a similar way to equation (17). However, we interpret the census shares as coming from a long-run model and set λ_{ij} to zero. We also assume the households in the census data face the market rent. In order to calculate the predicted shares, we need the joint distribution of the demographic characteristics, $\Pr(\mathbf{Z})$. We have constrained the census data to match the MTO distribution of race and income and additionally assume that the distribution of other attributes conditional on race and income is the same in MTO and census.

While our estimation strategy is similar to Berry, Levinsohn, and Pakes (1995) and Bayer, McMillan, and Rueben (2004), as we use a contraction mapping, there

³⁹ Certificates and vouchers were themselves randomly assigned. Two-thirds of MTO households assigned to either the Section 8 group or experimental group received vouchers and one-third received certificates. Therefore in estimation we use $\Pr\{S = v\} = \frac{2}{3}$ and $\Pr\{S = c\} = \frac{1}{3}$.

⁴⁰ We compute the neighborhood shares by reweighting and aggregating the neighborhood choices made by households observed in the census. Appendix A provides more details on the reweighting procedure.

is one important difference. In our estimation strategy, we are able to consistently estimate β^C in a first step as we have household-level variation in rental prices, R_{ij} and therefore C_{ij} ; as we include ξ_j in δ_j , the variation in C_{ij} is random and therefore uncorrelated with ϵ_{ij} .

The clean identification of β^C in the first stage using the experimental variation in out-of-pocket rent means that we do not have to find instruments for rent. As discussed above, finding appropriate instruments in BLP-style models can be difficult and has typically required clever, but explicit, use of the model's assumptions in the urban literature.

Knowing $\theta = (\beta^C, \beta_1^X, \lambda_0, \lambda_1, \lambda_2, \{\delta_j\}_{j=1}^J)$ is sufficient to conduct the counterfactual policy simulations described in Section V. However, to obtain measures of willingness to pay for neighborhood attributes requires estimating β_0^X in the second stage:

$$(22) \quad \hat{\delta}_j = \mathbf{X}_j' \beta_0^X + \xi_j.$$

Since two of our attributes (*Percent White_j* and *Pov_j*) partially reflect the aggregate characteristics of fellow renters in the same choice model, the sorting model suggests that these characteristics will likely be correlated with ξ_j .⁴¹ As such, estimating using IV is appropriate and consequently we require instruments for these two characteristics. For each neighborhood j , we use the average percentage of white neighbors and the average poverty rate in neighborhoods similar to j (excluding j from the average).⁴² Finally, it is worth noting that the experimental variation only contributes to the identification of the first stage and not the second stage where the unit of observation is a neighborhood.

C. Estimation Results

We consider a parsimonious specification of our model. For household attributes, \mathbf{Z}_i , we include household size as well as dummy variables for whether the household head was white, was never married, had previously applied to Section 8, had previously moved three times, or was very dissatisfied with their neighborhood. For neighborhood characteristics, \mathbf{X}_j , we include the poverty rate, the percentage white—two attributes which play a critical role in the design of the housing assistance programs we analyze—as well as measures of school quality and distance to jobs.

Table 3 presents the point estimates for the structural parameters of the neighborhood-choice model. As expected, the estimate of β^C is positive, meaning that reducing rental prices, and therefore increasing consumption, increases utility. With regard to the moving cost parameters, we find that λ_0 is negative and, as such, moving reduces utility. We find λ_1 is also negative, so that moving costs increase in

⁴¹The two neighborhood attributes (*Percent White_j* and *Pov_j*) are computed across all households in neighborhood j , regardless of whether they are renters or owners.

⁴²To define which neighborhoods are similar, we follow Geyer (2011) and use the k -means algorithm to partition the choice set into clusters of similar neighborhoods, using the other two attributes (school quality and distance to jobs) along with the median year in which the rental units in the neighborhood were built to assess the degree of similarity.

TABLE 3—ESTIMATED PARAMETERS

	Consumption		Mobility costs			
	Coefficient	SE	Coefficient	SE		
	β^C	4.33	0.48	λ_0	-6.60	0.26
				λ_1	-0.06	0.02
				λ_2	0.68	0.33
		Marginal utility from percent white		Marginal utility from poverty rate		
		Coefficient	SE	Coefficient	SE	
White		6.95	3.10		3.19	2.23
Never married		-0.82	0.76		1.11	1.13
Household size		-0.07	0.30		0.87	0.43
Applied to Section 8 before	β_1	-1.16	0.75	β_1	-3.16	1.12
Moved three times before		0.03	0.94		-0.81	1.84
Very dissatisfied		0.80	0.76		-1.68	1.23
Baseline	β_0	-3.38	1.34	β_0	-11.86	2.11
		Marginal utility from distance to jobs		Marginal utility from school quality		
		Coefficient	SE	Coefficient	SE	
White		-0.05	0.03		0.03	0.07
Never married		0.04	0.02		0.12	0.05
Household size		-0.0002	0.01		0.001	0.02
Applied to Section 8 before	β_1	-0.04	0.02	β_1	0.02	0.05
Moved three times before		0.01	0.04		0.03	0.06
Very dissatisfied		0.02	0.02		-0.02	0.05
Baseline	β_0	-0.07	0.04	β_0	-0.19	0.10

Notes: Standard errors computed using bootstrap. The table shows the maximum likelihood estimates for the first-stage structural parameters characterizing the marginal utility of (aggregate) consumption of goods and housing services, as well as moving costs and parameters of the marginal utility from neighborhood characteristics (poverty rate, percent white, distance to jobs, and school quality). The parameters associated with the six observable household characteristics represent utility interaction effects between such characteristics and the corresponding neighborhood characteristic. Also shown in the table are the baseline marginal utilities estimated in the second-stage decomposition using IV. Estimation sample includes only control group ($G = 0$) and experimental group ($G = 1$) observations. Section 8 held out for out-sample validation. Distance to jobs is measured in minutes using public transportation. School quality is measured by average fourth grade math and language test scores in the census tract. The effect of distance on moving costs is measured in miles.

distance. Finally, we find that λ_2 is positive, indicating a significant effect of mobility counseling for the experimental group in reducing moving costs.

The results in Table 3 have no direct interpretation in dollar values, however, marginal willingness to pay measures are easily interpretable. The annual marginal willingness to pay for attribute k of household i is given by $\frac{\beta_{i,k}^X I_i}{\beta^C}$. For example, we find nonwhite households have an average annual willingness to pay of $-\$122.10$ for a 1-percentage-point increase in the number of white neighbors.⁴³

⁴³We require the second-stage estimate for the parameter β_0^X associated with the neighborhood characteristic percent white to compute WTP for this characteristic. However, as noted above, for all of the policy analysis conducted in Section V, we do not need to decompose δ and only require the first-stage estimates of $\theta = (\beta^C, \beta_1^X, \lambda_0, \lambda_1, \lambda_2, \{\delta_j\}_{j=1}^J)$.

The negative estimate of WTP shows that these households would have to be compensated to accept this change in neighborhood characteristics and likely reflects preferences for neighbors of the same race. The differential average willingness to pay for percent white between whites and nonwhites is \$190.90. As a point of comparison, using a somewhat different model and a mixture of homeowners and renters in the Bay Area, Bayer and McMillan (2012) estimate a differential willingness to pay for percent white (between whites and blacks) of \$117.60. Finally, as mentioned above, one could estimate the model without using the experimental variation by using the control group only. When we do this, the parameter results are quite different. In particular, β^C is poorly identified using the control group only. Relative to the results above, the coefficient falls by about 50 percent, the standard error more than doubles, and the parameter is not significant at the 5 percent level, suggesting that the experimentally generated variation is the key source of identifying power.

IV. Model Validation

In this section, we provide evidence for how well our model fits the data, using both in-sample and out-of-sample exercises. To do this, we compare key empirical moments observed in the MTO data with the corresponding moments predicted by the model. In both cases, we find strong validation of our model and estimation approach.⁴⁴

The first moment we consider is the (ex post) mean exposure to given neighborhood characteristics, \mathbf{X} , conditional on assignment to a given group, $E[\mathbf{X}|G = g]$. We calculate this moment for the four neighborhood characteristics: poverty rate, percent white, distance to jobs, and school quality.

The second moment that we try to match is the subsidy take-up rate conditional on group assignment, $E[D|G = g]$. That is, the proportion of households who move using the subsidy, conditional on treatment status. To compute the model prediction for take-up, we compute, for each treatment group, the probability of moving to neighborhoods where the subsidies could be used. This method of computing the model's prediction of take-up assumes households behave rationally and, for a given neighborhood, would take advantage of a subsidy if a subsidy were available. For Section 8 households, there is no reason to not use the subsidy if moving. Therefore, the probability of moving is equal to the probability of moving using the subsidy. For experimental households, the subsidy take-up rate equals the rate at which these households moved to low-poverty neighborhoods.

Finally, we consider an alternative version of the moments relating to exposure to neighborhood attributes \mathbf{X} where we condition on subsidy take-up (as well as conditioning on treatment assignment) $E[\mathbf{X}|G = g, D = 1]$.⁴⁵ As before, we do this for the neighborhood attributes of poverty rate, percent white, distance to jobs, and school quality.

⁴⁴To compute empirical moments, we take averages across the appropriate MTO households. To compute model-predicted moments, we use the estimated model to compute the corresponding moments, by integrating over \mathbf{Z} , ϵ , and S .

⁴⁵Since we are conditioning on take-up (i.e., conditioning on moving using the subsidy), these conditional moments are not defined for the control group $G_i = 0$.

TABLE 4—WITHIN-SAMPLE FIT

	Data			Model		
	All C+E	0 Control	1 Exp	All C+E	0 Control	1 Exp
<i>Unconditional on move using the subsidy</i>						
Percent who move	0.49	0.29	0.65	0.49	0.29	0.65
Mean poverty rate	0.28	0.37	0.20	0.27	0.34	0.21
Mean percent white	0.41	0.32	0.48	0.44	0.37	0.50
School quality	33.7	31.9	35.1	34.3	33.1	35.3
Distance to jobs	43.2	40.3	45.5	41.4	40.1	42.4
<i>Percent who move using the subsidy</i>						
		0	0.55		0	0.47
<i>Conditional on move using the subsidy</i>						
Mean poverty rate		—	0.06		—	0.07
Mean percent white		—	0.69		—	0.75
School quality		—	38.2		—	38.9
Distance to jobs		—	48.3		—	43.7
Observations	369	165	204			

Notes: Empirical moments computed directly from final analysis sample of MTO households. Within-sample fit evaluated only on observations used in estimation (control and experimental groups only). Control group observations are not assigned subsidies so none of them move using the subsidy. Note that moments computed conditional on subsidy take-up are not defined for the control group.

Table 4 shows the quality of fit within the estimating samples of the control and experimental groups. As can be seen in the table, the model does a very good job of matching key features of the MTO data. Our model is able to replicate well the behavior of MTO participants in these two groups.

With the exception of moving costs, all of the model's parameters are assumed to be constant across group assignment. Therefore, we find the fact that we match typical exposure to neighborhood attributes separately for the control and experimental groups encouraging, particularly given that the exposure levels are very different across these groups in the actual data. Table 4 illustrates this point. In the data, the mean ex post exposure to poverty is 37 percent in the control group and 20 percent in the experimental group; a similar pattern of large differences between control group and experimental group exposure can be seen for percentage white, school quality, and distance to jobs. The model somewhat over predicts the exposure to percentage white and matches the other neighborhood moments very well, even though the respective utility parameters are constant across groups. Furthermore, we have similar success at matching these moments when we additionally condition on subsidy take-up. The moment we have most difficulty predicting is the percentage who move using the subsidy.

With access to a second treatment group, we also provide external validation of our model following the strategy in Todd and Wolpin (2006). That is, we can see how the model performs when applied to a sample that was randomly assigned to different moving incentives, but was not used in estimation.⁴⁶ For our test of

⁴⁶As Keane and Wolpin (2007) point out, randomized holdout samples that are experimentally assigned to different incentives provide one of the most convincing model validation strategies for structural models. Keane

TABLE 5—OUT-OF-SAMPLE FIT

	Section 8	
	Data	Model
<i>Unconditional on move using the subsidy</i>		
Mean poverty rate	0.27	0.25
Mean percent white	0.34	0.42
School quality	32.7	34.2
Distance to jobs	41.7	41.1
<i>Percent who move/percent who move using the subsidy</i>	0.63	0.70
<i>Conditional on move using the subsidy</i>		
Mean poverty rate	0.20	0.19
Mean percent white	0.38	0.50
School quality	33.6	35.2
Distance to jobs	42.4	41.0
Observations	172	

Notes: Subsample of Section 8 households held out for external model validation. Empirical moments computed directly from final analysis sample of MTO households. Out-of-sample fit evaluated on observations not used in estimation (Section 8 group only).

out-of-sample fit, we assess whether the model is able to replicate the neighborhood choice patterns of the Section 8 group that was offered an unrestricted subsidy. These observations (which were not used in estimation), faced different incentives as they were given no mobility counseling and had no location restrictions on subsidy use.⁴⁷

As may be seen in Table 5, the model is successful at matching the behavior of observations in the Section 8 group. The model over predicts exposure to white neighbors and the take-up rate, but matches the other six moments almost exactly. The success of the model is noteworthy given that the decisions made by the Section 8 group, as well as the incentives, are quite different from either the control or experimental groups.⁴⁸

V. Counterfactual and Policy Experiments

With strong evidence of external validation, we consider various counterfactual experiments using our model. Specifically, we look at (i) disentangling the effects of mobility counseling and location constraints; (ii) varying τ , the poverty-based

and Wolpin also propose a compelling nonrandom holdout sample approach for situations in which an experimentally generated validation sample is not available.

⁴⁷Ideally, one would also like to see if the model predicted well the location decisions of MTO participants in other sites. However, an important feature of our model and estimation approach is that we control for unobserved neighborhood attributes ξ_j , which precludes making predictions about neighborhood choices for MTO participants outside of Boston.

⁴⁸The numbers in Tables 4 and 5 provide the necessary ingredients to compute various ITT and TOT effects. One can compute and compare the empirical treatment effects estimated from the raw data with the ones based on the corresponding predictions from the structural model. An ITT effect is just the difference in a mean outcome (e.g., mobility rate or exposure to certain neighborhood attribute) between a treatment group (either experimental or Section 8) and the control group. An estimate of the corresponding TOT effect is given by the ITT effect divided by the subsidy take-up rate.

location constraint faced by the experimental group; and (iii) supplementing this poverty-based constraint with additional race-based constraints.

A. *Disentangling Counseling and Locations Constraints*

Recall that the take-up rates for the two treatment groups were very different. The two features of the experimental treatment influence households in opposite directions: mobility counseling encourages moving whereas location restrictions on subsidy use discourage moving. Using the mean difference in take-up between the two treatment groups, we can only conclude that location restrictions dominate counseling but cannot identify their separate magnitudes. To disentangle the two effects, we simulate moving behavior for the experimental treatment group without mobility counseling by setting $\lambda_2 = 0$. In our simulation, the location restrictions alone reduce take-up from 70 percent to 37 percent. When we add the mobility counseling, simulated take-up increases back up to 47 percent. This is consistent with work by Shroder (2003), who finds that the experimental group would need to be exposed to an extremely large counseling intensity to make up for the negative effects of the location constraint on take-up.

B. *Stringency of Location Constraints and Take-Up*

We also explore alternative policies where we vary the stringency of the location constraint τ . The experimental group faced a constraint of $\tau = 10\%$. For our simulations, we consider the following different values for τ :

$$\tau \in \{2.5, 5, 7.5, 10, 15, 20\}.$$

We then focus on take-up, and the change in exposure to neighborhood characteristics that these policies generate. The idea is to see whether more stringent location restrictions are successful in changing exposure to certain neighborhood characteristics, such as a low poverty rate in the neighborhood of residence. Of course, a lower (i.e., more stringent) poverty threshold τ for the location constraint would mechanically reduce exposure to poverty among those households that still decide to use the restricted subsidy. However, this positive effect could be outweighed by reduced take-up resulting from the more stringent location constraint associated with the subsidy.

As can be seen in Table 6, changing the restrictions on the maximum allowed poverty rate of the destination neighborhood (τ) changes the take-up rate substantially. When $\tau = 2.5\%$, the take-up rate is only 4 percent, whereas with a less stringent $\tau = 20\%$ it goes up to 63 percent. These simulations illustrate how binding the location constraints on subsidy use really are. The mean exposure to poverty resulting from these alternative policies actually declines with increases in τ .

As we reduce τ , exposure to poverty is reduced conditional on subsidy take-up. However, as we reduce τ , the subsidy take-up rate also falls. For the range of values that we consider, this second effect is stronger and reducing τ leads to higher overall exposure to poverty. The minimum unconditional average exposure to poverty for the experimental group is 20 percent and it is achieved at $\tau = 20\%$. Note,

TABLE 6—ALTERNATIVE NEIGHBORHOOD POVERTY RATE CUTOFFS (Percent)

τ (1)	Take-up (2)	Mean poverty rate (given take-up) (3)	Mean poverty rate (unconditional) (4)	Mean percent white (given take-up) (5)	Mean percent white (unconditional) (6)	WTP relative to MTO (7)
2.5	4	2	30	90	37	−\$1,494
5	16	4	27	87	42	−\$1,177
7.5	32	5	24	77	46	−\$647
10	47	7	21	75	50	\$0
15	57	9	20	70	51	\$503
20	63	11	20	65	50	\$878

Notes: Column 1 indexes counterfactual housing assistance policies that would introduce a more stringent ($\tau < 10\%$) or lenient ($\tau > 10\%$) location constraint relative to that implemented in MTO ($\tau = 10\%$). Column 2 shows what the take-up rate for the experimental group would be under each of the policies. Columns 3 and 5 display the resulting exposure to neighborhood characteristics (poverty rate and percent white) for those experimental households who decide to use the subsidy under each policy. Columns 4 and 6 show the unconditional exposures for the experimental group, by taking also into account the residential outcomes of those households that do not take up the subsidy offer. Column 7 measures annual willingness to pay in 1997 US\$ for each of the alternative policies (relative to the specific MTO policy). See text for details on the computation of WTP. All counterfactual policies in this table include counseling services. MTO policy allowed some households to move to places with poverty rate over 10 percent but still below 11 percent.

however, that the unconditional poverty exposure induced by the actual MTO policy ($\tau = 10\%$) is just one percentage point higher (21 percent) and that the pattern is fairly flat between $\tau = 10\%$ and $\tau = 20\%$. An alternative way of gauging the strength of the location constraints exploits our estimate of the marginal utility of consumption and calculates the willingness to pay (WTP) for alternative policies. In particular, we compute the average WTP for an alternative policy (relative to the MTO policy) as $\frac{1}{N_i} \sum_{i=1}^{N_i} WTP_i^\tau$ where the willingness to pay of experimental household i to get alternative policy τ is denoted by WTP_i^τ and is defined as follows:

$$(23) \quad E \left[\max_j \{ v_{ij}(\tau, I_i - WTP_i^\tau) \} \right] = E \left[\max_j \{ v_{ij}(\tau^{MTO}, I_i) \} \right].$$

These measures of willingness to pay make use of our estimate of β^C , and show that households in the experimental group are willing to pay \$878 per year to relax the location constraint from $\tau^{MTO} = 10\%$ to $\tau = 20\%$. Similarly, households in the experimental group are willing to pay \$1,494 per year to avoid changing the location constraint from $\tau^{MTO} = 10\%$ to $\tau = 2.5\%$.⁴⁹

C. A Desegregation Experiment

Finally, we explore what would have happened if the location restrictions regarding low poverty had been supplemented with a restriction on the racial composition of the destination neighborhood, similar in spirit to what the Gautreaux project

⁴⁹These average willingness to pay figures include households that would not move in either treatment. If we conditioned the willingness to pay measure on moving in both treatments or on moving in one of the treatments, the willingness to pay figures would naturally be larger.

implemented.⁵⁰ The Gautreaux program included a location constraint that only allowed subsidy use in neighborhoods in which no more than 30 percent of the households were black.⁵¹ We use our model to simulate the implications of an additional race-based location constraint for subsidy take-up and the resulting exposure to neighborhood characteristics.

We use a threshold of 30 percent as in Gautreaux, however, as our data only reveal white and nonwhite we impose the restriction on nonwhite rather than on black. Table 7 presents the results for nonwhite households, those most likely affected by the new constraint. As can be seen in the table, the additional location restrictions based on race substantially reduce take-up. Implementing a Gautreaux-style restriction (percent Nonwhite_{*j*} ≤ 30 percent) on top of the original restriction (Pov_{*j*} ≤ 10 percent) would have reduced the subsidy take-up rate among nonwhite experimental households in Boston from 44.1 percent to 35.4 percent.⁵² Interestingly, this combined policy is not successful at further reducing exposure to poverty, beyond what can be achieved with the MTO policy. The ex post unconditional exposure to poverty rate is essentially the same (22.0 percent under MTO versus 23.6 percent under the combined policy).

Moreover, despite its focus on race, a Gautreaux-like restriction would not significantly change exposure to other minority households (i.e., nonwhite experimental households end up exposed, on average, to neighborhoods with 46.7 percent white households under MTO and 47.1 percent under the combined policy). While the average racial composition of the neighborhood of residence changes substantially for those who do take up the subsidy with the two restrictions (percent white increases from 72.8 percent to 83.6 percent), the take-up rate is much smaller and therefore many more households remain in the public housing projects in highly segregated neighborhoods. The end result is that the neighborhood racial composition would be, on average, the same for this population whether or not we supplement the basic MTO location constraint with a race-based location constraint.

VI. Conclusion

We use data from the MTO experiment to estimate a model of neighborhood choice. The experimentally generated data are used for both estimation and out of sample validation. We rely on data from the control group and the experimental treatment group for estimation while holding out data from the unrestricted Section 8 treatment group for out-of-sample validation. The experimental variation is shown to be a powerful source of identification for one of the model's key structural parameters. The estimated model is successful in replicating the mobility and neighborhood choice patterns of low-income households receiving housing assistance. Model fit is good within the estimating sample and the model is also successful at replicating

⁵⁰ As discussed in Cutler and Glaeser (1997), racial segregation may theoretically have either positive or negative effects. However, they find empirically that decreasing segregation would significantly improve outcomes for black households.

⁵¹ See Rosenbaum (1995) for more details about the Gautreaux project and its results.

⁵² Note that the location restriction embodied in a Gautreaux-like intervention is relatively easy to comply with in the Boston metropolitan area. This is because the vast majority of neighborhoods in Boston are predominantly white. Therefore, take-up rate and WTP for this type of policy could be even lower in other cities where fewer neighborhoods satisfy the race-based constraint.

TABLE 7—ADDING RACE-BASED LOCATION CONSTRAINTS TO MTO (Percent)

Take-up	Mean poverty rate (unconditional)	Mean poverty rate (given take-up)	Mean percent white (unconditional)	Mean percent white (given take-up)
<i>MTO (experimental subsidy)</i>				
44.1	22.0	7.2	46.7	72.8
<i>MTO + Race-based location constraint</i>				
35.4	23.6	6.9	47.1	83.6

Notes: Simulations in this table are for nonwhite households. First row shows take-up and exposure to neighborhood characteristics (conditional on take-up and unconditionally) for the experimental subsidy as implemented in MTO. This is similar to the fourth row in Table 6, but for nonwhite households only. The second row shows the impact of adding a race constraint to the poverty-constrained, counseling-assisted MTO subsidy given to the experimental group. The race constraint resembles that used in Gautreaux by conditioning subsidy use to neighborhoods with more than 70 percent white households.

the behavior of the Section 8 group, a random subset of households not used in estimation and experimentally exposed to different moving incentives.

We use the estimated model to separate the quantitative importance of the two bundled features of treatment for the experimental group. We find that the effects of counseling and poverty-based location constraints are both large and that the location constraints end up dominating, which explains the lower take-up for the experimental group. We also show that subsidy take-up is sensitive to the particular design of the location constraint, with very stringent constraints inducing very low take-up. In particular, we show that due to reduced subsidy take-up rates, restricting subsidy use to very low (i.e., lower than what was required by MTO) poverty neighborhoods would actually increase average exposure to poverty. Finally, we show that supplementing the MTO intervention with a Gautreaux-style race-based location constraint would not change the average unconditional exposure to neighborhood characteristics in the population assigned to the experimental treatment.

APPENDIX A: DATA DETAILS

Neighborhood Attributes.—Our model includes four observable attributes: poverty rate, percent white, distance to jobs, and school quality. The poverty rate at the census tract level was computed using data from 1990 decennial population census, as this was the relevant rate used to verify whether a given apartment was located in a neighborhood that satisfied the experimental constraint. While our choice set is defined using census tract from the 2000 census, we use a crosswalk between census tract numbers for 2000 and 1990 to recover the appropriate poverty rate whenever a tract splits or two tracts merge between the two censuses. Percent white and distance to jobs were taken from the 2000 census. Percent white is just the number of white persons in the census tract divided by the total population in the census tract. Distance to jobs is measured as the average (within the census tract) number of minutes it takes to get to their jobs for those using public transportation. We also use the 2000 census to obtain the median age of rental units in the census tract, which helps us create instruments. Finally, to get a proxy measure of school quality, we used the 2001 Massachusetts Comprehensive Assessment System (MCAS) results. All children from third grade up to high school sit for this exam, but the

subject examined varies across grades (it can be either math or English language arts (ELA), or both). The main criterion for school eligibility is proximity (one-half of the seats in each school are reserved for students who live at one mile or less from the school). Although we are not able to individualize each child, we have the raw scores for each test taken at every school. Using data on each school's location from the Massachusetts Department of Elementary and Secondary Education, we can then average the scores at the zip code level. We then use a zip-code-to-census tract crosswalk to assign an average test score for each census tract using the corresponding zip codes. We use the average math and ELA test scores for fourth grade.

Using Census Data to Compute Neighborhood Shares.—Since the MTO data are too small to reliably compute shares for the approximately 600 neighborhoods in the choice set, we rely on data from the 2000 census. Of course a random household in the census renting within our choice set cannot be expected to be representative of an MTO household. Therefore, we compute the neighborhood shares by reweighting and aggregating the neighborhood choices made by these census households. While we do not have access to census microdata, we use publicly available tabulations with counts of renters by race and income bracket for each census tract. Using these counts, we construct a dataset that contains hundreds of thousands of renters residing in the census tracts that comprise our choice set. We then have neighborhood choice d , income bracket \tilde{I} , and race w for these N^{CEN} households: $\{d_n, \tilde{I}_n, w_n\}_{n=1}^{N^{CEN}}$. The empirical share for neighborhood j is then given by

$$(24) \quad \pi_j^{cen} = \frac{\sum_{n=1}^{N^{CEN}} \alpha(\tilde{I}_n, w_n) I\{d_n = j\}}{\sum_{n=1}^{N^{CEN}} \alpha(\tilde{I}_n, w_n)},$$

where the weights are constructed using

$$(25) \quad \alpha(\tilde{I}, w) = \frac{p^{MTO}(\tilde{I}, w)}{p^{CEN}(\tilde{I}, w)}$$

and $p^{MTO}(\tilde{I}, w)$ and $p^{CEN}(\tilde{I}, w)$ are the empirical joint probability mass functions for race ($w = 1$ if white, 0 otherwise) and household income (bracket) observed in MTO and census, respectively; $p^{MTO}(\tilde{I}, w)$ is generated by bracketing the total household income reported in 2001 by an MTO household, using the same income brackets for which we have renter counts by race in the 2000 census. By construction, this procedure makes sure that the (reweighted) census households have the same joint distribution of (\tilde{I}, w) as the households enrolled in MTO.

Final Analysis Sample.—Our final analysis sample includes 541 MTO households from the Boston area. To arrive at our final sample we focus on the 604 households who (i) had only one adult and at least one child at baseline; (ii) had valid information on the census tract of their original and, if moved, subsequently chosen residential location; (iii) had original and, if moved, subsequent locations within our choice set; and (iv) had a basis for baseline household income imputation (i.e., the household head was either on welfare and/or working at baseline). Of these 604,

we drop approximately 10 percent of households whose neighborhood choice cannot be easily rationalized within our simple model and one outlier whose estimated baseline income was too high.

Household-Level Variables.—We use seven individual-level household variables. We have indicators for race (=1 if white, =0 otherwise) and for marital status (=1 if household head had never married, =0 otherwise). Our measure of household size is an integer larger than or equal to 2, as we focus on households with just one adult and at least one child at baseline. This variable is used to determine welfare benefits and the appropriate FMR to be used when computing the housing subsidy that corresponds to the household. We also use three dummy variable indicators (=1 if statement is true, =0 otherwise) for whether (i) the household had moved three times before; (ii) it had applied for Section 8 assistance before enrolling in MTO; and (iii) it reported being very dissatisfied with the neighborhood. Regarding household income at baseline note that we focus on households with only one adult. When that adult is on welfare at baseline, we use the welfare benefits prevailing in Massachusetts in 1997 for a household with the appropriate number of children. Welfare benefits increase with the number of children. If the adult is working at baseline, we impute annual predicted labor earnings deflated back to 1997 using a regression of log weekly earnings (reported by working MTO adults in 2001) on age, age squared, and education. If the adult is working and is on welfare, we take the sum of the two.

Finally, to be consistent with our model we must be careful when constructing the neighborhood choice variable. First, in the model we assume that all control group moves pay market rent. However, a few control households report receiving Section 8 subsidies and living in a private unit at follow-up. An analogous issue arises with experimental households who didn't use the location-restricted subsidy offered through MTO. We drop these households. Another set of control group households, who moved to a new unit, report receiving public housing or project-based housing assistance at the time of follow-up. We consider these control households are reassignments within the Boston public housing system and treat them as stayers in the original tract, which given our model, is the most natural option as we don't model reassignments. We apply the same rule to experimental households who didn't use the location-restricted subsidy offered through MTO. In addition, to avoid further reduction in sample size we treat the few for whom we don't know about housing assistance receipt at the time of follow-up the same as those who report no assistance.

A couple of additional issues arise when dealing with some experimental households. First, we drop any experimental household who is recorded as having moved using the MTO subsidy, but who is observed to be living in a census tract whose 1990 poverty rate is above the cutoff, as this violates the location constraint. Similarly, in the model an experimental household who moves to a low-poverty census tract should always use the subsidy. While this is the case for the majority of experimental households, we drop those observations who have not used the subsidy but are living in a low-poverty area at the time of follow-up.

Finally, in the model we assume that whenever a Section 8 household moves to any neighborhood, it finds it in its interest to use the subsidy. Yet, a few Section 8 households are identified as not using the subsidy but observed to have moved into

a different census tract by the time of follow-up. If these households report being in public housing or receiving project-based assistance, we treat them as stayers, as this more accurately reflects their choice to decline the Section 8 subsidy at the time of random assignment. Some few other Section 8 households initially declined to use the subsidy offered through MTO, but subsequently made a move and report receiving Section 8 assistance at the time of follow-up. We treat these households as actually having taken up the Section 8 subsidy offered through MTO. We also drop those that were originally identified as having not used the subsidy, moved to a different census tract, and where we either don't know whether they are receiving or know they are not receiving housing assistance. Finally, we also drop the few Section 8 households who moved using the subsidy but did not change census tracts.

It is worth reiterating that, as explained above, despite all the cleaning steps required to smoothly integrate the neighborhood choice data into our model, we lose only 10 percent of the data due to all these considerations.

Neighborhood Choice Set.—The neighborhood choice set for our model includes 585 census tracts in four counties within the Boston primary metropolitan statistical area for which we are able to compute an index for the price of housing services. These counties are Suffolk (which includes the city of Boston), Norfolk, Middlesex, and Essex. The intersection of the Boston PMSA and these four counties contains 624 tracts. Based on real estate transactions data, we are able to compute a housing price index for all but five of them. We lose 32 additional tracts due to lack of school quality data and two tracts due to lack of access to jobs data. The choice set used in the model includes the remaining 585 tracts for which we have all four neighborhood attributes as well estimated prices for housing services.

APPENDIX B: INDIRECT UTILITIES

For a given decision of H_{ij}^* the indirect utility function can be found by first solving for optimal consumption C_{ij}^* using (9) and then plugging optimal consumption into (11).

If households stay in their existing neighborhood, they have no choice over the level of H_{ij} to consume and must consume their endowment, H_{ij}^e , regardless of assignment category. Housing services, overall (log) consumption, and (log) indirect utilities are given by

$$(26) \quad H_{ij}^* = H_{ij}^e$$

$$(27) \quad \log(C_{ij}^*) = (1 - \beta^H) \log((1 - \sigma)I_i) + \beta^H \log(H_{ij}^e)$$

$$(28) \quad v_{ij} = \beta^C ((1 - \beta^H) \log((1 - \sigma)I_i) + \beta^H \log(H_{ij}^e)) + \delta_j + \mathbf{X}'_j \beta_1^X \mathbf{Z}_i + \epsilon_{ij}.$$

For control group movers and those in the experimental group who move to a neighborhood where the poverty rate (Pov_j) exceeds the allowable threshold, τ ,

there is no subsidy. Therefore, the relevant budget constraint is $R_{ij} = r_j H_i$. Housing services, overall (log) consumption, and (log) indirect utilities are given by

$$(29) \quad H_{ij}^* = \beta^H \frac{I_i}{r_j}$$

$$(30) \quad \log(C_{ij}^*) = \Psi + \log(I_i) - \beta^H \log(r_j)$$

$$(31) \quad v_{ij} = \beta^C (\Psi + \log(I_i) - \beta^H \log(r_j)) + \delta_j + \mathbf{X}'_j \beta_1^X \mathbf{Z}_i + \lambda_{ij} + \epsilon_{ij},$$

where $\Psi = (1 - \beta^H) \log(1 - \beta^H) + \beta^H \log(\beta^H)$.

For experimental group movers who comply with the restriction and for all Section 8 group movers, the relevant budget constraint when the subsidy is given as a voucher is

$$R_{ij} = \max\{0, r_j H_i - [\rho_i - \sigma I_i]\}.$$

Housing services, overall (log) consumption, and (log) indirect utilities are given by

$$(32) \quad H_{ij}^* = \max \left\{ \beta^H \frac{(1 - \sigma) I_i + \rho_i}{r_j}, \frac{\rho_i - \sigma I_i}{r_j} \right\}$$

$$(33) \quad \log(C_{ij}^*) = \min \left\{ \Psi + \log((1 - \sigma) I_i + \rho_i) \right. \\ \left. - \beta^H \log(r_j), (1 - \beta^H) \log(I_i) + \beta^H \log\left(\frac{\rho_i - \sigma I_i}{r_j}\right) \right\}$$

$$(34) \quad v_{ij} = \beta^C \left(\min \left\{ \Psi + \log((1 - \sigma) I_i + \rho_i) - \beta^H \log(r_j), (1 - \beta^H) \log(I_i) \right. \right. \\ \left. \left. + \beta^H \log\left(\frac{\rho_i - \sigma I_i}{r_j}\right) \right\} \right) + \delta_j + \mathbf{X}'_j \beta_1^X \mathbf{Z}_i + \lambda_{ij} + \epsilon_{ij}.$$

Finally, for those using a certificate, it will always be optimal for the household to choose housing services such that the rent is exactly equal to the certificate value. Housing services, overall (log) consumption, and (log) indirect utilities are given by

$$(35) \quad H_{ij}^* = \frac{\rho_i}{r_j}$$

$$(36) \quad \log(C_{ij}^*) = (1 - \beta^H) \log((1 - \sigma) I_i) + \beta^H \log\left(\frac{\rho_i}{r_j}\right)$$

$$(37) \quad v_{ij} = \beta^C \left((1 - \beta^H) \log((1 - \sigma) I_i) + \beta^H \log\left(\frac{\rho_i}{r_j}\right) \right) \\ + \delta_j + \mathbf{X}'_j \beta_1^X \mathbf{Z}_i + \lambda_{ij} + \epsilon_{ij}.$$

APPENDIX C: OUT-OF-POCKET RENT FUNCTION

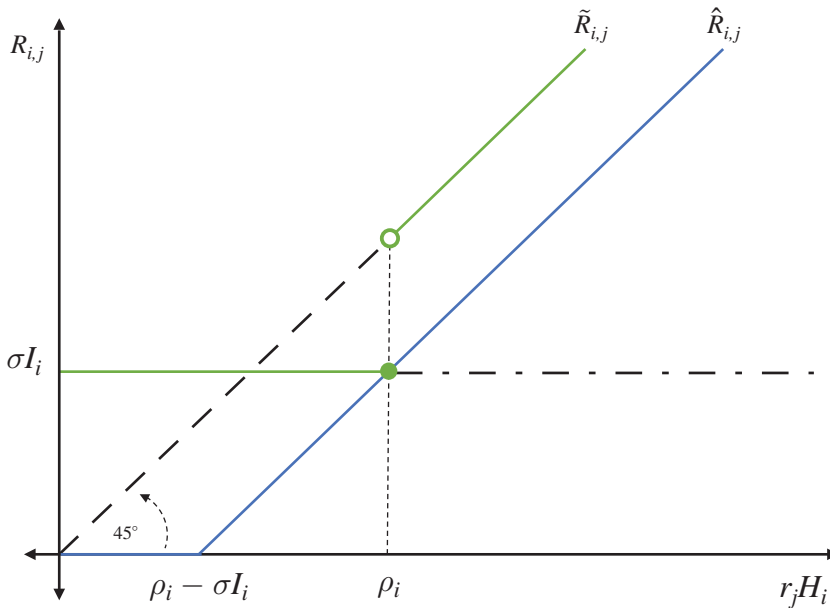


FIGURE C1

The experiment affected household behavior through the out-of-pocket rent function as given in (4). The function can be seen graphically in Figure C1. In the vertical axis the figure shows R_{ij} , the actual out-of-pocket rent that household i pays in a given neighborhood j . The horizontal axis displays $r_j H_i$, the market cost of renting H_i units of housing in neighborhood j . The line \tilde{R}_{ij} denotes the out-of-pocket rent function for certificate recipients (from either Section 8 group for any j or experimental group whenever $Pov_j \leq 10\%$). The line \hat{R}_{ij} denotes the out-of-pocket rent function for voucher recipients (from either Section 8 group for any j or experimental group whenever $Pov_j \leq 10\%$). Any nonmover who remains in the public housing projects pays σI_i . Finally, the 45-degree line characterizes the out-of-pocket rent function for control group movers as well as experimental group movers in neighborhoods that do not satisfy the location constraint ($Pov_j > 10\%$).

REFERENCES

- Aliprantis, Dionissi. 2011. "Assessing the Evidence on Neighborhood Effects from Moving to Opportunity." Federal Reserve Bank of Cleveland Working Paper 11–22.
- Attanasio, Orazio P., Costas Meghir, and Ana Santiago. 2012. "Education Choices in Mexico: Using a Structural Model and a Randomized Experiment to Evaluate PROGRESA." *Review of Economic Studies* 79 (1): 37–66.
- Bayer, Patrick, Fernando Ferreira, and Robert McMillan. 2007. "A Unified Framework for Measuring Preferences for Schools and Neighborhoods." *Journal of Political Economy* 115 (4): 588–638.
- Bayer, Patrick, Nathaniel Keohane, and Christopher Timmins. 2009. "Migration and Hedonic Valuation: The Case of Air Quality." *Journal of Environmental Economics and Management* 58 (1): 1–14.

- Bayer, Patrick, and Robert McMillan.** 2012. "Tiebout Sorting and Neighborhood Stratification." *Journal of Public Economics* 96 (11–12): 1129–43.
- Bayer, Patrick, Robert McMillan, and Kim Rueben.** 2004. "An Equilibrium Model of Sorting in an Urban Housing Market." National Bureau of Economic Research Working Paper 10865.
- Berry, Steven T.** 1994. "Estimating Discrete-Choice Models of Product Differentiation." *RAND Journal of Economics* 25 (2): 242–62.
- Berry, Steven T., and Philip A. Haile.** 2010. "Nonparametric Identification of Multinomial Choice Demand Models with Heterogeneous Consumers." Cowles Foundation for Research in Economics Discussion Paper 1718.
- Berry, Steven T., and Philip A. Haile.** 2014. "Identification in Differentiated Products Markets Using Market Level Data." *Econometrica* 82 (5): 1749–97.
- Berry, Steven, James Levinsohn, and Ariel Pakes.** 1995. "Automobile Prices in Market Equilibrium." *Econometrica* 63 (4): 841–90.
- Berry, Steven, James Levinsohn, and Ariel Pakes.** 2004. "Differentiated Products Demand Systems from a Combination of Micro and Macro Data: The New Car Market." *Journal of Political Economy* 112 (1): 68–105.
- Black, Sandra E.** 1999. "Do Better Schools Matter? Parental Valuation of Elementary Education." *Quarterly Journal of Economics* 114 (2): 577–99.
- Cutler, David M., and Edward L. Glaeser.** 1997. "Are Ghettos Good or Bad?" *Quarterly Journal of Economics* 112 (3): 827–72.
- Duflo, Esther, Rema Hanna, and Stephen P. Ryan.** 2012. "Incentives Work: Getting Teachers to Come to School." *American Economic Review* 102 (4): 1241–78.
- Ellickson, Bryan.** 1971. "Jurisdictional Fragmentation and Residential Choice." *American Economic Review* 61 (2): 334–39.
- Epple, Dennis, Radu Filimon, and Thomas Romer.** 1984. "Equilibrium among Local Jurisdictions: Toward an Integrated Treatment of Voting and Residential Choice." *Journal of Public Economics* 24 (3): 281–308.
- Epple, Dennis, and Richard E. Romano.** 1998. "Competition between Private and Public Schools, Vouchers, and Peer-Group Effects." *American Economic Review* 88 (1): 33–62.
- Epple, Dennis, and Thomas Romer.** 1991. "Mobility and Redistribution." *Journal of Political Economy* 99 (4): 828–58.
- Epple, Dennis, and Holger Sieg.** 1999. "Estimating Equilibrium Models of Local Jurisdictions." *Journal of Political Economy* 107 (4): 645–81.
- Feins, Judith D., Debra McInnis, and Susan Popkin.** 1997. *Counseling in the Moving to Opportunity Demonstration Program*. HC-5953 Task Order 3. Report prepared for the US Department of Housing and Urban Development, Washington, DC.
- Ferreira, Maria Marta.** 2007. "Estimating the Effects of Private School Vouchers in Multidistrict Economies." *American Economic Review* 97 (3): 789–817.
- Ferreira, Maria Marta.** 2009. "An Empirical Framework for Large-Scale Policy Analysis, with an Application to School Finance Reform in Michigan." *American Economic Journal: Economic Policy* 1 (1): 147–80.
- Galiani, Sebastian, Alvin Murphy, and Juan Pantano.** 2015. "Estimating Neighborhood Choice Models: Lessons from a Housing Assistance Experiment: Dataset." *American Economic Review*. <http://dx.doi.org/10.1257/aer.20120737>.
- Geyer, Judy.** 2011. "Housing Demand and Neighborhood Choice with Housing Vouchers." Unpublished.
- Geyer, Judy, and Holger Sieg.** 2013. "Estimating a Model of Excess Demand for Public Housing." *Quantitative Economics* 4 (3): 483–513.
- Goering, John, and Judith D. Feins, eds.** 2003. *Choosing a Better Life? Evaluating the Moving to Opportunity Social Experiment*. Washington, DC: Urban Institute Press.
- Jacob, Brian A., and Jens Ludwig.** 2012. "The Effects of Housing Assistance on Labor Supply: Evidence from a Voucher Lottery." *American Economic Review* 102 (1): 272–304.
- Katz, Lawrence F., Jeffrey R. Kling, and Jeffrey B. Liebman.** 2001. "Moving to Opportunity in Boston: Early Results of a Randomized Mobility Experiment." *Quarterly Journal of Economics* 116 (2): 607–54.
- Keane, Michael P., and Kenneth I. Wolpin.** 2007. "Exploring the Usefulness of a Nonrandom Hold-out Sample for Model Validation: Welfare Effects on Female Behavior." *International Economic Review* 48 (4): 1351–78.
- Kling, Jeffrey R., Jeffrey B. Liebman, and Lawrence F. Katz.** 2007. "Experimental Analysis of Neighborhood Effects." *Econometrica* 75 (1): 83–119.

- Kling, Jeffrey R., Jens Ludwig, and Lawrence F. Katz.** 2005. "Neighborhood Effects on Crime for Female and Male Youth: Evidence from a Randomized Housing Voucher Experiment." *Quarterly Journal of Economics* 120 (1): 87–130.
- Kuminoff, Nicolai V.** 2012. "Partial Identification of Preferences in a Dual-Market Sorting Equilibrium." <http://www.public.asu.edu/~nkuminof/K12.pdf> (accessed September 17, 2015).
- Ludwig, Jens, Greg J. Duncan, and Paul Hirschfield.** 2001. "Urban Poverty and Juvenile Crime: Evidence from a Randomized Housing-Mobility Experiment." *Quarterly Journal of Economics* 116 (2): 655–79.
- McFadden, Daniel.** 1974. "Conditional Logit Analysis of Qualitative Choice Behavior." In *Frontiers in Econometrics*, edited by P. Zarembka, 105–42. New York: Academic Press.
- Nechyba, Thomas J.** 1999. "School Finance Induced Migration and Stratification Patterns: The Impact of Private School Vouchers." *Journal of Public Economic Theory* 1 (1): 5–50.
- Nechyba, Thomas J.** 2000. "Mobility, Targeting, and Private-School Vouchers." *American Economic Review* 90 (1): 130–46.
- Olsen, Edgar O.** 2003. "Housing Programs for Low-Income Households." In *Means-Tested Transfer Programs in the United States*, edited by Robert A. Moffitt, 365–441. Chicago: University of Chicago Press.
- Orr, Larry.** 2011. "Moving to Opportunity (MTO) for Fair Housing Demonstration: Interim Impacts Evaluation, Tier 1 Restricted Access Data, 1994–2001 [United States]." ICPSR31661-v1. Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor]. <http://doi.org/10.3886/ICPSR31661.v1>.
- Orr, Larry, Judith D. Feins, Robin Jacob, Erik Beecroft, Lisa Sanbonmatsu, Lawrence F. Katz, Jeffrey B. Liebman, and Jeffrey R. Kling.** 2003. *Moving to Opportunity Interim Impacts Evaluation*. Final report prepared for the US Department of Housing and Urban Development, Office of Policy Development & Research, Washington, DC.
- Rosenbaum, James E.** 1995. "Changing the Geography of Opportunity by Expanding Residential Choice: Lessons from the Gautreaux Program." *Housing Policy Debate* 6 (1): 231–69.
- Ruggles, Steven J., Trent Alexander, Katie Genadek, Ronald Goeken, Matthew B. Schroeder, and Matthew Sobek.** 2010. Integrated Public Use Microdata Series: Version 5.0 [Machine-readable database]. Minneapolis, MN: Minnesota Population Center [producer and distributor], 2010.
- Shroder, Mark.** 2003. "Locational Constraint, Housing Counseling, and Successful Lease-Up." In *Choosing a Better Life? Evaluating the Moving to Opportunity Social Experiment*, edited by John Goering and Judith D. Feins, 59–80. Washington, DC: Urban Institute Press.
- Sieg, Holger, V. Kerry Smith, H. Spencer Banzhaf, and Randy Walsh.** 2002. "Interjurisdictional Housing Prices in Locational Equilibrium." *Journal of Urban Economics* 52 (1): 131–53.
- Sieg, Holger, V. Kerry Smith, H. Spencer Banzhaf, and Randy Walsh.** 2004. "Estimating the General Equilibrium Benefits of Large Changes in Spatially Delineated Public Goods." *International Economic Review* 45 (4): 1047–77.
- Todd, Petra E., and Kenneth I. Wolpin.** 2006. "Assessing the Impact of a School Subsidy Program in Mexico: Using a Social Experiment to Validate a Dynamic Behavioral Model of Child Schooling and Fertility." *American Economic Review* 96 (5): 1384–1417.
- Waldfogel, Joel.** 2003. "Preference Externalities: An Empirical Study of Who Benefits Whom in Differentiated-Product Markets." *RAND Journal of Economics* 34 (3): 557–68.
- Walsh, Randy.** 2007. "Endogenous Open Space Amenities in a Locational Equilibrium." *Journal of Urban Economics* 61 (2): 319–44.
- Wise, David.** 1985. "A Behavioral Model versus Experimentation: The Effects of Housing Subsidies on Rent." In *Methods of Operations Research 50*, edited by P. Brucker and R. Pauly, 441–89. Verlag Anton Hain.
- Wong, Maisy.** 2013. "Estimating Ethnic Preferences Using Ethnic Housing Quotas in Singapore." *Review of Economic Studies* 80 (3): 1178–1214.