Estimating Neighborhood Choice Models: Lessons from a Housing Assistance Experiment*

Sebastian Galiani
Dept. of Economics
Washington University
in St. Louis

Alvin Murphy
Olin Business School
Washington University
in St. Louis

Juan Pantano
Dept. of Economics
Washington University
in St. Louis

June 5, 2012

Abstract

We use data from a housing assistance experiment to estimate a model of neighborhood choice. The experimental variation, which effectively randomizes the rents that households face, allows us to identify the model's structural parameters. Access to two randomly-selected treatment groups, in addition to a control group, allows for the out-of-sample validation of the model using a group of households who were not used in estimation and who faced a separate set of incentives. We use our estimated model to simulate the effects of changing the subsidy-use constraints implemented in the actual experiment and find that restricting subsidies to even lower poverty neighborhoods substantially reduces take-up. As a result, average exposure to poverty actually increases under these more restrictive subsidies. We also simulate the effect of adding additional subsidy restrictions based on neighborhood racial composition and find that this policy does not change the average household exposure to either neighborhood racial composition or poverty.

*We thank Dan Weiss, Raul Sanchez, and Bo Wang for their valuable research assistance. We also benefited from comments by Pat Bayer, Kelly Bishop, Flavio Cuhna, Dennis Epple, Limor Golan, Hanming Fang, Fernando Ferreira, Maria Marta Ferreyra, Vernon Henderson, Paul Joice, Maurizio Mazzoco, Holger Sieg, Paulo Somaini, Steve Tadelis, Giorgio Topa, Ken Wolpin and seminar participants at UPenn, New York Fed, ITAM, Oxford, Vanderbilt, Collegio Carlo Alberto, Tsinghua University, Pompeu Fabra and Toulouse.

1 Introduction

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 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, no paper to date has used experimental data to either estimate or 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 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 that households face to estimate our model. 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 provides details on the demographic characteristics and location choices made by households placed into

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

[2] See among others, Epple and Sieg (1999), Sieg, Smith, Banzhaf and Walsh (2004), Bayer, Ferreira, and McMillan (2007), Ferreyra (2007), Walsh (2007), Kuminoff (2008) and Bayer, McMillan, and Rueben (2011)

[3] The closest the urban literature has come to using experimental data is Wong (2010) 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 non-experimental policy change to validate a model of location and school choice.

[4] For example, Bayer, Ferreira, and McMillan (2007) use the equilibrium prices predicted by the model based only on exogenous attributes as an instrumental variable and Bayer, Keohane, and Timmins (2009) use the model's prediction of the share of income spent on housing to calibrate the price parameter

one of three random assignment groups: a control group, a treatment group given mobility counseling and housing subsidies that were restricted to low poverty neighborhoods, and a treatment group that was given unrestricted housing subsidies with no counseling. The MTO data has been previously used to estimate the effect of the MTO intervention on labor market and other outcomes, as well as estimating neighborhood effects.[5] 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 provides 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.[6] As we have two separate treatment groups, we are able to do both; we use one treatment group (together with the control group) for estimation of the location-choice model and reserve the other treatment group for out-of-sample validation.[7]

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 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 to help in the search process for a new apartment outside the public

[5] See, among others, Katz, et al.(2001), Kling et al.(2005) and Kling et al.(2007).

[6] If 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.

[7] Structural 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) exploited 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 (2011) 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 structural model and experimental data is Duflo, Hanna and Ryan (2007).

housing project. 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 that receives both mobility counseling and location restrictions is approximately 20 percent less likely to use the subsidy compared with the group that was 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 55 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 13% of households would use the subsidy under a more stringent restriction that limits subsidy use to neighborhoods with a poverty rate under 5%. 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 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.

The remainder of the paper proceeds as follows: Section 2 discusses the MTO program and the data. Our model is outlined in Section 3 and Section 4 describes the estimation strategy and results. We present the model fit and validation exercises in Sections 5 and policy evaluations in Section 6. Finally, Section 7 concludes.

2 Experimental Background and Data

2.1 The Moving to Opportunity Experiments

Five public housing authorities (Baltimore, Boston, Chicago, Los Angeles, and New York City) administered Department of Housing and Urban Development contracts under the MTO demonstration. Within each authority's jurisdiction, eligible households who volunteered and who were living in assisted public housing projects were randomly selected and placed into 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 ten percent poverty. This group also received housing counseling to help them find appropriate locations and successfully use the subsidy. 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. A household offered a subsidy had 90 days to find an apartment if it was in the Section 8 group. Households in the Experimental group were given an additional month. Experimental group families 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.[8]

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 Katz, Kling, and Ludwig (2005).[9]

[8] A 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.

[9] In addition to published academic articles discussed here, excellent summaries and policy oriented compilations of this body of 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 are similar for all groups, indicating a successful randomization. A year after randomization however, those who had moved lived in strikingly different areas than those who had not and the difference persisted even four years after that. They also show that households in both treatment groups were more likely to live in substantially wealthier neighborhoods one year after the intervention. As expected, the Experimental treatment was more successful than the unrestricted Section 8 treatment in relocating poor families into low-poverty and suburban neighborhoods. However, the unrestricted Section 8 assistance was more effective in getting a larger share of families out of the most distressed communities (i.e., unrestricted subsidies had a higher take-up rate). The changes in neighborhoods induced by MTO did not appear to have significantly affected employment rates, earnings, or welfare receipt.

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.[10] 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).[11]

An important feature of the experiments is the take up rate of the subsidies. Schroder (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. Schroder concludes that location constraints had strong effects and trumped the positive effects of counseling.[12] Below we use our structural model to

[10] Given the endogeneity problems induced by residential choices, the model was estimated by 2SLS using a full set of site-by-treatment interactions as the excluded instruments for the neighborhood poverty rate in the first stage.

[11] See also Aliprantis (2011) for a re-analysis 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.

[12] Schroder (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 Schroder (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.

disentangle the separate roles of counseling and location restrictions.

2.2 The Data

The dataset contains data for all adults and children 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 household. 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,[13] 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 household has moved at least 3 times in the last five years, whether the household had applied for Section 8 vouchers in the past). For households who use the subsidy offered through MTO, we observe the neighborhood where they use the subsidy and, for those observations, we treat this as our measure of neighborhood choice.[14] 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 analysis sample of 614 observations.[16] Table 1 presents descriptive statistics. As can be seen in the table, the data show good covariate balance, confirming successful 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 (poverty rate and percent white) and neighborhood

[13] We restrict our analysis to those households with only one adult. When the adult is on welfare at baseline, we use the welfare benefits prevailing in Massachusetts in 1997. If the adult is working at baseline, we impute annual labor earnings deflated back to 1997 using a regression of 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 maximum of the two. Welfare benefits may vary by number of children.

[14] These moves occurred no longer than 90 (Section 8 group) or 120 (Experimental group) days from the date of random assignment.

[15] Since 1995, FMR is set at the 40^{th} 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 2-bedroom FMR to 2- or 3-person households, the 3-bedroom FMR to 4-person households, and the 4-bedroom FMR to households with 5 or more members.

[16] We only use observations in which we are sure only one adult is present in the houshold. 80 observations are lost as the information regarding their neighborhood choice is missing. Another 115 observations are discarded as their neighborhood choice involves a census tract not contained within the Boston PMSA.

Table 1: MTO Data Descriptive Statistics

	Control	Experimental	Section 8	Total
White	0.10	0.12	0.11	0.11
Household Income (in 1,000s)	10.1	10.3	9.5	10.0
Ever Married	0.62	0.64	0.67	0.64
Household Size	3.25	3.14	3.14	3.18
Applied to Section 8 Before	0.56	0.55	0.59	0.56
Moved 3 Times Before	0.13	0.14	0.14	0.14
Dissatisfied with Neighborhood	0.33	0.34	0.30	0.32
Observations	200	222	194	616

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 dollars includes welfare payments for those on welfare and estimated labor income for those working. See text for details.

rental price for apartment units. From Summary File 4, we obtain data reflecting the joint distribution of income and race for renters in each neighborhood. Using this data, we can form the neighborhood shares for a population with characteristics similar to the MTO sample, based on renter status, race and income.

For our model and estimation approach, we define neighborhoods as 6-digit Census Tracts and the choice set includes 591 6-digit Census tracts in the Boston primary metropolitan statistical area. 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 186 Census tracts in Boston. Initially, however, they were distributed in a more narrow set of 26 census tracts, essentially corresponding to the census tracts in which the targeted public housing projects were located. Finally, as a measure of market rent we use the median rent in each neighborhood, which we obtain from the Census summary files.

Before going to the model we briefly document the patterns of take-up rates 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 Z_i denotes demographic characteristics of the households.

$$D_i = \alpha_0 + \alpha_1 \{G_i = \text{Experimental}\} + \alpha_2 \{G_i = \text{Section 8}\} + Z_i'\beta + u_i \tag{1}$$

As can be seen in Table 2, the take-up rate for the Section 8 group is substantially higher than for the Experimental group.

Table 2: Voucher Take Up

	(1)	(2)
Experimental	0.468***	0.458***
	(0.0335)	(0.0325)
Section 8	0.582***	0.574***
	(0.0355)	(0.0359)
White		0.135***
		(0.0505)
Household Income		0.0013
		-0.0024
Ever Married		0.00402
		(0.0334)
Household Size		-0.0301*
		(0.0161)
Applied to Section 8 Before		0.129***
		(0.0323)
Moved 3 Times Before		0.149***
		(0.0455)
Dissatisfied with Neighborhood		0.105***
		(0.0339)
Constant	0	-0.0558
	(0)	(0.0675)
Observations	616	616

Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Boston MTO final analysis sample. The dependent variable is equal to one if the household uses the voucher, equals zero otherwise. Control group observations are the omitted category but they were not given vouchers so their dependent variable is always zero, and the regression without controls in column 1 goes through the origin.

There is an eleven percentage point gap (47% vs. 58%) in take-up rates.[17] This suggests that

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

restrictions on location outweigh any positive effect that housing counseling may have had. Note 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. This would allow us to estimate take up rates separately for each possible restriction. 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.

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.

3 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.[18] 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. We are only aware of the related work by Geyer (2011) that uses data from Pittsburgh to study housing assistance policy.[19] The primary difference between

[18] Following earlier work by McFadden (1974), the literature on discrete choice significantly gained in popularity after Berry (1994) and Berry, Levinsohn and Pakes (1995) showed how to allow for unobserved product characteristics and conduct estimation using aggregate shares of the chosen characteristics. In recent papers, Berry and Haile (2010a,2010b) 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.

[19] See Epple, Geyer, and Sieg (2011) for a model which focuses on public-housing assistance rather than voucher-based assistance, which we focus on here.

the approach taken here and previous sorting models is our use of experimental data.

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 choice of residential neighborhood.[20] 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, X_j, and household's characteristics Z_i.[21] 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 household consumption, C_i, observable and unobservable neighborhood attributes, respectively X_j and ξ_j, household characteristics, 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 $d_i = j \in \{1, \ldots, J\}$ among the available neighborhoods, including the option of staying in the same public housing unit $(j = j_{t-1})$.[22] Households assigned to either the Experimental or the Section 8 treatment groups also effectively choose whether to use the subsidy $(D_i = 1)$.

Therefore, households are solving:

$$\max_{\{d_i\}} U\left(C_i, X_j, \xi_j, I_i, Z_i, j_{i,t-1}, \epsilon_{ij}, \theta\right) \tag{2}$$

subject to the budget constraint:

$$C_i + R_{ij} = I_i \tag{3}$$

where R_{ij} denotes the out-of-pocket rent payment. The out-of-pocket rent is given by the function R, which takes as its arguments treatment group assignment, G_i; an indicator H_i for whether the household receives the housing assistance subsidy in the form of a voucher (v) or a certificate (c),[23]

[20] As discussed in Section 2, a neighborhood is defined as a 6 digit Census tract.

[21] Based on results in Kling et al (2007) we assume households anticipate no income differences across neighborhoods.

[22] We abstract from modeling the consumption of housing services within each neighborhood. See Wong (2011) for a similar specification of neighborhood choice.

[23] See Olsen (2003) for a comprehensive discussion of different housing assistance subsidies.

neighborhood choice (including its market rent, R_j^m, and its poverty rate), j; baseline neighborhood choice, j_{t-1}; household income, I_i; and features of the subsidy program,(σ, ρ, τ) . In addition to its format H, the actual subsidy depends on the share of household income that must be paid, σ; the subsidy cap, ρ;[24] and the restriction on a neighborhood's poverty rate, τ.

$$R_{ij} = R\left(G_i, H_i, j, j_{t-1}, R_j^m, I_i, \sigma, \rho, \tau\right) \tag{4}$$

The specific form of the out-of-pocket rent function depends on whether the assistance is location-restricted or not and, if restricted, whether the neighborhood being considered satisfies the restriction.[25]

$$R_{ij} = \begin{cases} \sigma I_i & \text{if } \ j = j_{t-1}, \ \text{all } G_i \\ R_j^m & \text{if } \ j \neq j_{t-1}, G_i = \text{Control} \\ \max\left\{0, R_j^m - [\rho - \sigma I_i]\right\} & \text{if } \ j \neq j_{t-1}, G_i = \text{Sec 8}, \ H_i = \text{v} \\ R_j^m & \text{if } \ j \neq j_{t-1}, G_i = \text{Exp}, \ H_i = \text{v}, \text{Pov. Rate}_j > \tau \\ \max\left\{0, R_j^m - [\rho - \sigma I_i]\right\} & \text{if } \ j \neq j_{t-1}, G_i = \text{Exp}, \ H_i = \text{v}, \text{Pov. Rate}_j < \tau \\ \sigma I_i & \text{if } \ j \neq j_{t-1}, G_i = \text{Sec 8}, \ H_i = \ \text{c}, \ R_j^m \leq \rho \\ R_j^m & \text{if } \ j \neq j_{t-1}, G_i = \text{Sec 8}, \ H_i = \text{c}, \ R_j^m > \rho \\ R_j^m & \text{if } \ j \neq j_{t-1}, G_i = \text{Exp}, \ H_i = \text{c}, \text{Pov. Rate}_j > \tau \\ \sigma I_i & \text{if } \ j \neq j_{t-1}, G_i = \text{Exp}, \ H_i = \text{c}, \text{Pov. Rate}_j < \tau, R_j^m \leq \rho \\ R_j^m & \text{if } \ j \neq j_{t-1}, G_i = \text{Exp}, \ H_i = \text{c}, \text{Pov. Rate}_j < \tau, R_j^m > \rho \end{cases} \tag{5}$$

We parameterize the conditional indirect utility function for household i associated with choosing neighborhood j as:

$$u_{ij} = \alpha_i X_j + \beta_i R_{ij} + \lambda_{ij} + \xi_j + \epsilon_{ij} \tag{6}$$

where ϵ_{ij} is distributed i.i.d. Type 1 Extreme Value. We specify $\lambda_{ij} = \lambda_i 1\left\{j \neq j_{i,t-1}\right\}$ where $1\left\{x\right\}$ is an indicator function that equals one whenever x is true and equals zero otherwise. Noting that

[24] At the beginning of the actual MTO implementation the cap ρ was the 45th percentile of the distribution of rents in the metropolitan area. Since 1995, the cap is set at the 40th percentile. These numbers are the fair market rents or FMRs. They vary not only across metropolitan areas and year, but also with the number of bedrooms in a unit.

[25] We assume that a household in the control group faces the market rents, $\left\{R_j^m\right\}_{j=1}^J$, if they choose to move. We ignore tranfers to another public housing project located in different neighborhoods.

X_j is a vector of K attributes, we specify the household-specific parameters (α_i, β_i, and λ_i) as:

$$\alpha_{i,k} = \alpha_{0,k} + \alpha_{1,k} Z_i \quad (7)$$

$$\beta_i = \beta_0 + \beta_I I_i + \beta_1 Z_i \quad (8)$$

$$\lambda_i = \lambda_0 + \lambda_1 1\{G_i = 1\} \quad (9)$$

λ_i is a moving cost function which, as expected, is only paid if the household moves (i.e. if $j \neq j_{i,t-1}$). As those in the Experimental group ($\mathrm{G}_i = 1$) receive mobility counseling, we allow their moving cost to differ (by amount λ_1) from the baseline moving costs (λ_0) faced by the other groups. β_I captures the price sensitivity for households with different incomes, while β_1 and α_1 capture how the utility parameters vary with household demographic characteristics, Z_i.

Employing the definition of the household-specific utility parameters, we can rewrite u_{ij} as:

$$u_{ij} = \alpha_0 X_j + \alpha_1 Z_i X_j + \beta_0 R_{ij} + \beta_I I_i R_{ij} + \beta_1 Z_i R_{ij} + \lambda_{ij} + \xi_j + \epsilon_{ij} \quad (10)$$

By adding and subtracting $\beta_0 R_j^m$ and collecting neighborhood-level effects into the fixed effect, δ_j, we rewrite the conditional indirect utility as:

$$u_{ij} = \delta_j + \alpha_1 Z_i X_j + \beta_0 (R_{ij} - R_j^m) + \beta_I I_i R_{ij} + \beta_1 Z_i R_{ij} + \lambda_{ij} + \epsilon_{ij} \quad (11)$$

where δ_j is given by:

$$\delta_j = \alpha_0 X_j + \beta_0 R_j^m + \xi_j \quad (12)$$

This model provides a rich representation of household residential mobility decisions and highlights how those mobility decisions may be influenced by housing assistance policy parameters.

4 Estimation

4.1 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 U.S. 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 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). This random variation in prices allows us to identify a key structural parameter of the neighborhood choice model (β_0) without relying on the typical model-based exclusion restrictions that are necessary to form instruments. One of the neighborhood attributes, rent, is randomly different for the Control and Experimental group participants and, as such, we would expect the two groups to make different location decisions. This difference in locations decisions identifies the coefficient on rent, β_0.

Additionally, as the MTO microdata reveal how the location decisions vary with demographic characteristics, we are able to identify how individual characteristics affect preferences for neighborhood attributes.

The propensity to move in the Control group identifies the baseline moving cost parameter λ_0. As we observe a different propensity to move across assignment groups, we can also identify how moving costs differ for the Experimental group, which is captured by the parameter λ_1.

To control for unobserved neighborhood attributes, we rely on a data strategy that combines the MTO microdata with U.S. 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.

4.2 Estimation Details

The estimation routine 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 $\alpha_1, \beta_0, \beta_I, \beta_1, \lambda_0, \lambda_1$, 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 the observable neighborhood characteristics as given by Equation 12, which allow us to recover the remaining parameters, α_0.

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

$$\pi^h_{ij} = \frac{\exp\left(\delta_j + \alpha_1 Z_i X_j + \beta_0(R^h_{ij} - R^m_j) + \beta_I I_i R_{ij} + \beta_1 Z_i R_{ij} + \lambda_{ij}\right)}{\sum_{k=1}^{J} \exp\left(\delta_k + \alpha_1 Z_i X_k + \beta_0(R^h_{ik} - R^m_k) + \beta_I I_i R_{ik} + \beta_1 Z_i R_{ik} + \lambda_{ik}\right)} \tag{13}$$

where R^h_{ij} is the out-of-pocket rent that household i would pay if choosing neighborhood j when receiving the subsidy with format h. Recall that $h = v$ for vouchers and $h = c$ for certificates. Since the format of housing assistance is unobserved we integrate over it by letting $\pi_{ij} = \pi^v_{ij} \Pr\{H = v\} + \pi^c_{ij} \Pr\{H = c\}$.[26] The first estimation step finds the vector $\theta = \left(\alpha_1, \beta_0, \beta_I, \beta_1, \lambda_0, \lambda_1, \{\delta_j\}_{j=1}^J\right)$ that solves the following problem:

$$\max_{\theta} \sum_{i=1}^{N} \sum_{j=1}^{J} \log(\pi_{ij}) 1\{d_i = j\} \tag{14}$$

subject to:

$$\pi_j(\theta) = \pi^{census}_j \quad \forall \quad j \tag{15}$$

where π^{census}_j is the empirical share of households who choose neighborhood j in the Census data and $\pi_j(\theta)$ is the model prediction for this share based on a given parameter guess θ.[27]

For each trial of $(\alpha_1, \beta_0, \beta_I, \beta_1, \lambda_0, \lambda_1)$, 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

$$\delta^{n+1} = \delta^n + \log(\pi^c_j) - \log(\pi^c_j(\delta^n)) \tag{16}$$

where the predicted share of neighborhood j is given by the model as:

$$\pi_j(\delta) = \int \pi_{ij}(\delta) dF(Z_i) \tag{17}$$

The probability of household i choosing a neighborhood j, π_{ij}, is formed in a similar way to

[26] 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 $\Pr\{H = v\} = \frac{2}{3}$ and $\Pr\{H = c\} = \frac{1}{3}$.

[27] Before computing π^{census}_j we resample the Census microdata to mimic the distribution of race and household income observed for MTO housholds.

Equation 13. However, we interpret the Census shares as coming from a long-run model and set λ_i to zero. As most of the households in the Census data are not receiving housing assistance, we assume they face the market rent. In order to calculate the predicted shares, we need the joint distribution of the demographic characteristics, $F(Z)$ which we can observe from the Census and MTO data.[28]

While our estimation strategy is similar to Berry, Levinsohn, and Pakes (1995) and Bayer McMillan, and Rueben (2011), as we use a contraction mapping, there is one important difference. In our estimation strategy, we are able to consistently estimate β_0 in a first step as we have household-level variation in rental prices, R_{ij}; as we include ξ_j in δ_j, the variation in $\mathrm{R}_{ij} - \mathrm{R}_j^m$ is random and therefore uncorrelated with ε_{ij}.

With β_0 consistently estimated in the first stage, a straightforward OLS approach is employed in the second stage to decompose the fixed effects described by Equation 12. This regression is given by:

$$\hat{\delta}_j - \hat{\beta}_0 R_j^m = \alpha_0 X_j + \xi_j \tag{18}$$

The clean identification of $\hat{\beta}_0$ in the first stage using the experimental variation in out-of-pocket rent means that we do not have to find instruments for R_j^m. 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.

4.3 Estimation Results

We consider a parsimonious specification of our model. For household attributes, Z, we include household size as well as dummy variables for whether the household was white, was ever married, had previously applied to Section 8, had previously moved three times, or was very dissatisfied with their neighborhood. For neighborhood characteristics, X, we include the poverty rate and the percentage white. These neighborhood attributes play a critical role in the design of the housing assistance programs we analyze. Finally, we use the market rent, R^m.

Table 3 presents the point estimates for the structural parameters, θ, of the neighborhood-

[28] The joint distribution of a subset of Z is observable in the Census data. To form the shares, we simply need to know the distribution of the attributes that are unobserved in the Census conditional on those that are observed. Fortunately, this information is readily observable in the MTO data.

choice model. As expected, the estimate of β_0 is negative, meaning that increasing rental prices reduces utility. Furthermore, the estimate of β_I is positive, suggesting that the sensitivity of utility to rental prices is lower for higher income households. With regard to the moving cost parameters, we find that λ_0 is negative and, as such, moving reduces utility. In addition, we find that λ_1 is positive, indicating a significant positive 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 $-\frac{\alpha_{i,k}}{\beta_i}$. For example, we find that a non-white household has an annual willingness to pay of \$-164.27 for a one percentage-point increase in the number of white neighbors (holding other demographic characteristics at their mean values).[29] The negative estimate of WTP shows that these households actually have to be compensated to consider this change in neighborhood characteristics and likely reflects preferences for neighbors of the same race.

5 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.

The first moment we consider is the (ex-post) mean exposure to a given neighborhood characteristic, X, conditional on assignment to a given group, $E\left[X|G=g\right]$. We calculate this moment for the neighborhood characteristics Poverty Rate and Percent White.

The second moment that we try to match is the subsidy take-up rate conditional on group assignment, $E\left[D|G=g\right]$. That is, the proportion of participants who move using the subsidy, conditional on treatment status.[30] To compute the model prediction for take up, we use the

[29] The second stage estimate for the parameter α_0 associated with the neighborhood characteristic Percent White is -1.943. We require this additional estimate to compute WTP for this characteristic. However, it is worth noting that for all of the policy analysis conducted in Section 6, we do not need to decompose δ and only require the first-stage estimates of $\theta = \left(\alpha_1, \beta_0, \beta_I, \beta_1, \lambda_0, \lambda_1, \{\delta_j\}_{j=1}^{J}\right)$.

[30] The take up rate is lower in the Experimental group relative to the Section 8 group. A priori, this is not necessarily obvious. The location constraint in the Experimental subsidy reduces the value of moving for many

Table 3: Estimated Parameters

		Mobility Costs	
		Coef.	SE
Baseline Mobility Cost	λ_0	-5.25	(0.027)
Experimental Group Interaction	λ_1	1.328	(0.026)

		Rental Price	
		Coef.	SE
Constant	β_0	-0.551	(0.014)
Annual Income (in 1,000s)	β_I	0.012	(0.001)
White	β_1	-0.036	(0.009)
Ever Married	β_1	0.088	(0.008)
Household Size	β_1	0.048	(0.004)
Applied S8 before	β_1	-0.093	(0.008)
Moved 3 times before	β_1	0.006	(0.009)
Very Dissatisfied	β_1	0.110	(0.008)

		Marginal Utility from Percent White			Marginal Utility from Poverty Rate	
		Coef.	SE		Coef.	SE
White	$\alpha_{1,WHITE}$	4.324	(0.123)	$\alpha_{1,POV}$	-0.135	(0.119)
Ever Married	$\alpha_{1,WHITE}$	-0.175	(0.061)	$\alpha_{1,POV}$	-0.308	(0.115)
Household Size	$\alpha_{1,WHITE}$	-0.564	(0.027)	$\alpha_{1,POV}$	0.097	(0.054)
Applied S8 before	$\alpha_{1,WHITE}$	-0.572	(0.052)	$\alpha_{1,POV}$	-2.571	(0.123)
Moved 3 times before	$\alpha_{1,WHITE}$	0.225	(0.067)	$\alpha_{1,POV}$	-3.025	(0.155)
Very Dissatisfied	$\alpha_{1,WHITE}$	0.978	(0.056)	$\alpha_{1,POV}$	-1.56	(0.120)

Standard errors in parentheses computed using bootstrap. The table shows the first stage structural parameters for price sensitivity, moving costs and parameters of marginal utility from neighborhood characteristics (poverty rate and % white). Rental price is annual rent measured in thousands of dollars. The parameters associated with the six observable household characteristics represent utility interaction effects between such characteristics and the corresponding neighborhood characteristic (rental price, poverty rate, % white). Estimation Sample includes only Control group (G=0) and Experimental Group (G=1) observations. Section 8 held out for out-sample validation.

Table 4: Within Sample Fit

	All	0	1	All	0	1
	C+E	Control	Exp	C+E	Control	Exp
	Data			Model		
Unconditional on Move Using the Subsidy						
% Who Move	0.54	0.35	0.71	0.54	0.35	0.71
Mean Poverty Rate	0.27	0.33	0.21	0.28	0.34	0.22
Mean % White	0.41	0.36	0.47	0.43	0.37	0.48
% Who Move Using the Subsidy		0	0.47		0	0.39
Conditional on Move Using the Subsidy						
Mean Poverty Rate		n/a	0.06		n/a	0.07
Mean % White		n/a	0.76		n/a	0.78
Observations	422	200	222			

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). See appendix for details about construction of moments predicted by the model. Control group observation are not assignedsubsidies so none of them move using the subsidy. Note that moments computed conditional on subsidy take up are not defined for the control group.

neighborhood choice probabilities predicted by the model and we sum these probabilities over neighborhoods in which 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.[31]

Finally, we consider an alternative version of the moments relating to exposure to neighborhood attributes X where we condition on voucher take-up (as well as conditioning on treatment assignment) $E\left[X|G=g,D=1\right]$.[32] As before, we do this for the neighborhood attributes of Poverty Rate and Percent White.

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

households. However, the Experimental group is subject to an additional treatment of mobility counseling, which should increase the take-up rate.

[31] Further details about how the moments are formed may be found in the Appendix.

[32] Since we are conditioning on take-up (i.e., conditioning on moving using the subsidy) these conditional moments are not defined for the Control group $\mathrm{G}_i = 0$.

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 33% in the Control group and 21% in the Experimental group; the respective figures for exposure to percentage white are 35% and 47%. The model predicts all four of these moments almost perfectly, even though the respective utility parameters are constant across groups. Furthermore, we closely match these moments when we additionally condition on subsidy take up. The only moment which is not predicted almost exactly by the model is the percentage who move using the subsidy.

Table 5: Out of Sample Fit

	Section 8	
	Data	Model
Unconditional on Move Using the Subsidy		
% Who Move	0.61	0.58
Mean Poverty Rate	0.27	0.28
Mean % White	0.38	0.40
% Who Move Using the Subsidy	0.58	0.58
Conditional on Move Using the Subsidy		
Mean Poverty Rate	0.21	0.19
Mean % White	0.40	0.50
Observations	194	

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). See appendix for details about construction of moments predicted by the model.

With access to a second treatment group, we also provide external validation of our model. That is, we can see how the model performs when applied to a sample that faces different moving incentives, but was not used in estimation. For our test of 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 restrictions on location.[33] As may be seen in Table 5, the model is very successful at matching the behavior of observations in the Section 8 group. The model overpredicts exposure to white neighbors conditional on using the subsidy, but matches the other five 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.

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

6.1 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 $\lambda_1 = 0$. In our simulation, the location restrictions alone reduce take up from 58% to 26%. When we add the mobility counseling, simulated take up increases back up to 39%. This is consistent with work by Shroder (2003) who finds that the experimental group

[33] Ideally, 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.

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.

6.2 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 voucher. 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 3%, whereas with a less stringent $\tau = 20\%$ it goes up to 58%. These simulations illustrate how binding the location constraints on subsidy use really are. The mean exposure to poverty resulting from these alternative policies actually <u>declines</u> 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 21.2% and it is achieved at $\tau = 20\%$. Note, however, that the unconditional poverty exposure induced by the actual MTO policy ($\tau = 10\%$) is just 1 percentage point higher (22.3%) 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 WTP for an

Table 6: Alternative Neighborhood Poverty Rate Cutoffs

(1)	(2)	(3)	(4)	(5)	(6)	(7)
τ	Take-up	Mean Poverty Rate (given take-up)	Mean Poverty Rate (unconditional)	Mean % White (given take-up)	Mean % White (unconditional)	WTP relative to MTO
2.5%	3%	2%	27.9%	93%	38%	-$1,476
5%	13%	3%	26.1%	90%	42%	-$1,127
7.5%	27%	5%	23.8%	78%	45%	-$579
10%	39%	7%	22.3%	78%	48%	$0
15%	50%	9%	21.4%	73%	48%	$504
20%	58%	11%	21.2%	66%	48%	$944

Column (1) indexes counterfactual voucher policies that would introduce more stringent (τ<10%) or lenient (τ>10%) location constraint relative to that implemented in MTO (τ=10%). Column (2) shows what the take up rate for the experimental group under each of the policies would be. Columns (3) and (5) display the resulting exposure to neighborhood characteristics (poverty rate and %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. Column (7) measures annual willingness to pay in 1997 dollars 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 slightly over 10% but still below 11%.

alternative policy (relative to the MTO policy) as:

$$\frac{1}{N_1}\sum_{i=1}^{N_1}\frac{E\left[\max_j u_{ij}\left(\tau\right)\right]-E\left[\max_j u_{ij}\left(\tau^{MTO}\right)\right]}{\beta_i} \tag{19}$$

These measures of willingness to pay make use of our estimate of β_0, and show that households in the experimental group are willing to pay \$944 per year to relax the location constraint from $\tau^{MTO} = 10\%$ to $\tau = 20\%$. Similarly, households in the experimental group are willing to pay \$1,476 per year to avoid changing the location constraint from $\tau^{MTO} = 10\%$ to $\tau = 2.5\%$.

6.3 A Desegregation Experiment

Finally, we explore what would have happened if the location restrictions regarding low poverty were supplemented with a restriction on the racial composition of the destination neighborhood,

Table 7: Adding Race-Based Location Constraints to MTO

Take-up	Mean Poverty Rate (unconditional)	Mean Poverty Rate (given take-up)	Mean % White (unconditional)	Mean % White (given take-up)
MTO (Experimental Voucher)				
35.1%	23.4%	7.11%	42.9%	75.5%
MTO + Race-Based Location Constraint				
28.0%	24.5%	6.82%	43.3%	87.3%

Simulations in this table are for non-white 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 4th row in Table 7 but for non-white 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 less than 30% minority households.

similar in spirit to what the Gautreaux program implemented.[34] The Gautreaux program included a location constraint that only allowed subsidy use in neighborhoods in which no more than 30% of the households were black.[35] 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% as in Gautreaux, however, as our data only reveal white and non-white we impose the restriction on non-white rather than on black. Table 8 presents the results for non-white 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 the actual Gautreaux restriction $\left(\tau^{NONWHITE} = 30\%\right)$ on top of the original restriction $\left(\tau^{POVERTY} = 10\%\right)$ would had reduced the voucher take up rate among non-white experimental households in Boston from 35.1% to 28%.[36] Interestingly, this combined

[34] 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.

[35] See Rosenbaum (1995) for more details about the Gautreaux Project and its results.

[36] 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

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 (23.4% under MTO vs. 24.5% 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. non-white experimental households end up exposed, on average, to neighborhoods with 42.9% white households under MTO and 43.3% under the combined policy). While the average racial composition of the neighborhood of residence changes substantially for those who do take up the voucher with the two restrictions (% White increases from 75.5% to 87.3%), 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.

7 Conclusion

We use data from the MTO experiment to estimate a model of neighborhood choice. The experimentally generated data is 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 the model's structural parameters. The estimated model is successful is 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 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

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.

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.

References

[1] Aliprantis, D. (2011), Assessing the Evidence on Neighborhood Effects from Moving to Opportunity, Federal Reserve Bank of Cleveland, Working Paper 11-22.

[2] Attanasio, O., Meghir, C. and Santiago, A (2011), Education Choices in Mexico: Using a Randomized Experiment and a Structural Model to evaluate PROGRESA", forthcoming Review of Economic Studies.

[3] Bayer, P., Ferreira, F. and McMillan, R. (2007), A Unified Framework for Measuring Preferences for Schools and Neighborhoods, Journal of Political Economy, vol. 115, no. 4

[4] Bayer, P., McMillan, R. and Rueben, K. (2005), An Equilibrium Model of Sorting in an Urban Housing Market, mimeo

[5] Black, Sandra, (1999) Do Better Schools Matter? Parental Valuation of Elementary Education, Quarterly Journal of Economics, May, Vol. 114, No. 2, Pages 577-599

[6] Berry, S. (1994), Estimating Discrete Choice Models of Product Differentiation, RAND Journal of Economics, Vol 25, No 2, Summer.

[7] Berry, S., Levinsohn, J. and Pakes, A. (1995), Automobile prices in market equilibrium, Econometrica, vol. 63, pp. 841–90.

[8] Berry, S. and Haile, P. (2010a) Identification in Differentiated Products Markets Using Market Level Data, CFDP 1744

[9] Berry, S. and Haile, P. (2010b) Nonparametric Identification of Multinomial Choice Demand Models with Heterogeneous Consumers, CFDP 1718

[10] Cutler, D. and Glaeser, E. (1997), Are Ghettos Good or Bad?, Quarterly Journal of Economics 112: 827-872.

[11] Duflo, E. Hanna, R. and Ryan, S. (2007), Incentives Work: Getting Teachers to Go to School, American Economic Review, forthcoming.

[12] Ellickson (1971) Jurisdictional Fragmentation and Residential Choice," American Economic Review, 61(2), 334-339.

[13] Epple, D. Filimon, R. and Romer, T. (1984) "Equilibrium among local jurisdictions: Toward an integrated treatment of voting and residential choice," Journal of Public Economics, 24(3), 281-308.

[14] Epple, D. and Romer, T. (1991) "Mobility and Redistribution," Journal of Political Economy, 99(4), 828-858.

[15] Epple, D. and Romano, R. (1998) "Competition between private and public schools, vouchers, and peer-group effects" American Economic Review, 88(1), 33-62

[16] Epple, D. and H. Sieg (1999), Estimating Equilibrium Models of Local Jurisdictions, Journal of Political Economy, 107, 645-681

[17] Epple, D. Geyer, J. and Sieg, H. (2011), Excess Demand and Rationing in Equilibrium in the Market for Public Housing, mimeo, CMU.

[18] Geyer, J. (2011) Housing Demand and Neighborhood Choice with Housing Vouchers, mimeo, CMU.

[19] Goering, J. and Feins, J. (2003) "Choosing a Better Life. Evaluating the Moving to Opportunity Social Experiment", The Urban Institute Press. Washington DC.

[20] Feins, J., McInnis, D. and Popkin, S. (1997) "Counseling in the Moving to Opportunity Demonstration Program" Abt Associates, Contract #HC-5953 Task Order 3, Report to HUD.

[21] Ferreyra, M.M. (2007) Estimating the Effects of Private School Vouchers in Multi-District Economies, American Economic Review, 97(3), 789-817.

[22] Ferreyra, M.M. (2009) An Empirical Framework for Large Scale Policy Analysis, with an application to School Finance Reform in Michigan, American Economic Journal.

[23] Jacob, B.A. and Ludwig, J. , (2012), The Effects of Housing Assistance on Labor Supply: Evidence from a Voucher Lottery, American Economic Review 102(1): 272–304

[24] Katz, L., Kling, J. R. and Liebman, J. B. (2001), Moving To Opportunity In Boston: Early Results Of A Randomized Mobility Experiment, Quarterly Journal of Economics, 116:607–54.

[25] Kling, J. R., Ludwig, J. and Katz, L., (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.

[26] Kling, J. R., Liebman, J. B. and Katz, L. (2007), Experimental Analysis of Neighborhood Effects, Econometrica

[27] Kuminoff, N. V. (2008), Recovering Preferences from a Dual-Market Locational Equilibrium," mimeo, Arizona State University.

[28] Ludwig, J. Duncan, G. and Hirschfield, P. (2001), Urban Poverty and Juvenile Crime, Evidence from a Randomized Housing-Mobility Experiment, Quarterly Journal of Economics, , Vol. 116, No. 2 (May), pp. 655-679

[29] Nechyba, T. J. (1999) School Finance Induced Migration and Stratification Patterns: The Impact of Private School Vouchers, Journal of Public Economic Theory, 1(1), 5-50.

[30] Nechyba, Y. J. (2000) Mobility, Targeting, and Private-School Vouchers, American Economic Review, 90(1), 130-146.

[31] Olsen, E. O. (2003), Housing Programs for Low Income Households, Chapter 6 in Means-Tested Transfer Programs in the United States, edited by Robert Moffitt.

[32] Orr, L. Feins, J. Jacob, R., Beecroft, E., Sanbonmatsu, L., Katz, L., Liebman, J.B, Kling, J.R. (2003), Moving to Opportunity Interim Impacts Evaluation. Report to HUD.

[33] Rosenbaum, J. E. (1995). Changing the geography of opportunity by expanding residential choice: Lessons from the Gautreaux program. Housing Policy Debate 6(1), 231–269

[34] Shroder, M. (2003) , Locational Constraints, Housing Counseling, and Successful Lease-Up", in Choosing a Better Life, Goering and Feins, Eds. The Urban Institute Press.

[35] Sieg, H., V. K. Smith, H. S. Banzhaf, and R. Walsh (2004): Estimating the General Equilibrium Benefits of Large Changes in Spatially Delineated Public Goods, International Economic Review, 45, 1047-1077

[36] Todd, P. and Wolpin, K. I (2006) "Using a Social Experiment to Validate a Dynamic Behavioral Model of Child Schooling and Fertility: Assessing the Impact of a School Subsidy Program in Mexico," American Economic Review, December

[37] Waldfogel, J. (2003) "Preference Externalities: An Empirical Study of Who Benefits Whom in Differentiated-Product Markets," RAND Journal of Economics, 34(3), 557–568.

[38] Walsh, R. (2007), Endogenous Open Space Amenities in a Locational Equilibrium, Journal of Urban Economics, 61(2), 319-344

[39] Wise, D., (1985). A behavioral model versus experimentation: The effects of housing subsidies on rent. In: Brucker, P., Pauly, R. (Eds.), Methods of Operations Research, vol. 50. Verlag Anton Hain, pp. 441–489.

[40] Wong, M. (2011): Estimating Ethnic Preferences Using Ethnic Housing Quotas in Singapore.

Appendix: Validation Moment Details

In this appendix we give further details about how we form the moments used in the validation exercises described in Section 5.

$$\begin{aligned} E\left[X|G=g\right] &= \sum_h \Pr\left(H=h|G=g\right) E\left[X|G=g,H=h\right] \\ & \sum_h \Pr\left(H=h\right) E\left[X|G=g,H=h\right] \end{aligned} \tag{20}$$

$$\begin{aligned} E\left[X|G=g,H=h\right] &= \sum_j X_j \Pr\left(d=j|G=g,H_i=h\right) \\ &= \sum_j X_j \left[\sum_z \Pr\left(d=j|G=g,H_i=h,z\right) p\left(z\right)\right] \\ &= \sum_j X_j \left[\frac{1}{N_g}\sum_{i=1}^{N_g} \Pr\left(d=j|G_i=g,H_i=h,Z_i\right)\right] \\ &= \sum_j X_j \left[\frac{1}{N_g}\sum_{i=1}^{N_g} \pi_{ij}^h\right] \end{aligned} \tag{21}$$

Regarding the take up rate we have

$$E\left[D|G=g\right] = \sum_h \Pr\left(H=h|G=g\right) E\left[D|G=g,H=h\right] \tag{22}$$

$$E\left[D|G=g,H=h\right] = \begin{cases} 0 & \text{if} \quad \text{G=0} \\ \Pr\left(d \neq j_{t-1}\ ;\ \text{Pov. Rate}_j < 10\%\ ;\ R_j^m < \rho|G=1,H=c\right) & \text{if} \quad \text{G=1, H=}c \\ \Pr\left(d \neq j_{t-1}\ ;\ \text{Pov. Rate}_j < 10\%|G=1,H=v\right) & \text{if} \quad \text{G=1, H=}v \\ \Pr\left(d \neq j_{t-1}\ ;\ R_j^m < \rho|G=2,H=c\right) & \text{if} \quad \text{G=2, H=}c \\ \Pr\left(d \neq j_{t-1}|G=2,H=v\right) & \text{if} \quad \text{G=2, H=}v \end{cases} \tag{23}$$

For Section 8 participants (G=2) who receive a voucher (H=v) we can use the fact that take up is equivalent to moving out of the public housing project so we have

$$
\begin{aligned}
E\left[D|G=2,H=v\right] &= \Pr\left(D=1|G=2,H=c\right) \\
&= \Pr\left(d\neq j_{t-1}|G=2,H=v\right) \\
&= 1-\Pr\left(d=d_{t-1}|G=2,H=v\right) \\
&= 1-\left[\sum_{z}\Pr\left(d=j_{t-1}|G=2,H=v,z\right)p\left(z\right)\right] \\
&= 1-\left[\sum_{i:G_i=2}\Pr\left(d_i=j_{t-1}|G_i=2,H=v,Z_i\right)\left(\frac{1}{N_2}\right)\right] \\
&= 1-\left[\frac{1}{N_2}\sum_{i=1}^{N_2}\Pr\left(d_i=j_{t-1}|G_i=2,H=v,Z_i\right)\right] \\
&= 1-\left[\frac{1}{N_2}\sum_{i=1}^{N_2}\pi^{v}_{i,j_{t-1}}\right]
\end{aligned}
\tag{24}
$$

where $N_2=\sum_i 1\left\{G_i=2\right\}$ is the number of MTO households assigned to the Section 8 group.

For Section 8 participants (G=2) who receive a certificate (H=c) take up is not necessarily equivalent to moving out of the public housing project. This is because a household may have a strong preference for moving to a neighborhood j for which the certificate does not qualify (i.e. $\mathrm{R}^m_j>\rho$) so we have

$$
\begin{aligned}
E\left[D|G=2,H=c\right] &= \Pr\left(D=1|G=2,H=c\right) \qquad (25)\\
&= \Pr\left(d\neq j_{t-1}\text{and } R_d^m<\rho|G=2,H=c\right)\\
&= \sum_z \Pr\left(d\neq j_{t-1} \text{ and } R_d^m<\rho|G=2,H=c,z\right)p\left(z\right)\\
&= \sum_{i:G_i=2} \Pr\left(d\neq j_{t-1} \text{ and } R_d^m<\rho|G=2,H=c,Z_i\right)\left(\frac{1}{N_2}\right)\\
&= \left(\frac{1}{N_2}\right)\sum_{i:G_i=2} \Pr\left(d\neq j_{t-1} \text{ and } R_d^m<\rho_i|G_i=2,H=c,Z_i\right)\\
&= \left(\frac{1}{N_2}\right)\sum_{i:G_i=2}\left\{\sum_j 1\left\{j\neq d_{i,t-1}\right\}\ 1\left\{R_j^m<\rho_i\right\}\Pr\left(d=j\left|\begin{array}{c} G_i=2,\\ H=c,Z_i\end{array}\right.\right)\right\}\\
&= \left(\frac{1}{N_2}\right)\sum_{i:G_i=2}\left\{\sum_j 1\left\{j\neq d_{i,t-1}\right\}\ 1\left\{R_j^m<\rho_i\right\}\pi_{i,j}^c\right\}
\end{aligned}
$$

For Experimental group participants (G=1) who received a voucher (H=v) we have

$$
\begin{aligned}
& E\left[D|G=1,H=v\right]\\
=\ & \Pr\left(D=1|G=1,H=v\right) \qquad (26)\\
=\ & \Pr\left(d\neq d_{t-1} \text{ and Pov. Rate}_d\leq 10\%|G=1,H=v\right)\\
=\ & \sum_z \Pr\left(d\neq d_{t-1} \text{ and Pov. Rate}_d\leq 10\%|G=1,H=v,z\right)p\left(z\right)\\
=\ & \sum_{i:G_i=1} \Pr\left(d\neq d_{t-1} \text{ and Pov. Rate}_d\leq 10\%|G_i=1,H=v,Z_i\right)\left(\frac{1}{N_1}\right)\\
=\ & \left(\frac{1}{N_1}\right)\sum_{i:G_i=1} \Pr\left(d\neq d_{i,t-1} \text{ and Pov. Rate}_d\leq 10\%|G_i=1,H=v,Z_i\right)\\
=\ & \left(\frac{1}{N_1}\right)\sum_{i:G_i=1}\left(\sum_j 1\left\{j\neq d_{i,t-1}\right\}\ 1\left\{\text{Pov. Rate}_j\leq 10\%\right\}\Pr\left(d=j\left|\begin{array}{c} G_i=1,\\ H=v,Z_i\end{array}\right.\right)\right)\\
=\ & \left(\frac{1}{N_1}\right)\sum_{i:G_i=1}\left(\sum_j 1\left\{j\neq d_{i,t-1}\right\}\ 1\left\{\text{Pov. Rate}_j\leq 10\%\right\}\pi_{ij}^v\right)
\end{aligned}
$$

Similarly, for Experimental group participants (G=1) who received a certificate (H=c) we have

$$
\begin{aligned}
& E\left[D|G=1,H=c\right] && (27)\\
= \; & \Pr\left(D=1|G=1,H=c\right) \\
= \; & \Pr\left(d\neq d_{t-1} \text{ and Pov. Rate}_d \;\leq 10\% \text{ and } \mathrm{R}_d^m < \rho|G=1,H=c\right) \\
= \; & \sum_z \Pr\left(d\neq d_{t-1} \text{ and Pov. Rate}_d \leq 10\% \text{ and } \mathrm{R}_d^m < \rho|G=1,H=c,z\right) p\left(z\right) \\
= \; & \sum_{i:G_i=1} \Pr\left(d\neq d_{t-1} \text{ and Pov. Rate}_d \leq 10\% \text{ and } \mathrm{R}_d^m < \rho_i|G_i=1,H=c,Z_i\right)\left(\frac{1}{N_1}\right) \\
= \; & \left(\frac{1}{N_1}\right)\sum_{i:G_i=1} \Pr\left(d\neq d_{i,t-1} \text{ and Pov. Rate}_d \leq 10\% \text{ and } \mathrm{R}_d^m < \rho_i|G_i=1,H=c,Z_i\right) \\
= \; & \left(\frac{1}{N_1}\right)\sum_{i:G_i=1}\left(\sum_j 1\left\{j\neq d_{i,t-1};\text{Pov. Rate}_j \leq 10\%;\mathrm{R}_j^m<\rho_i\right\}\Pr\left(d=j|\begin{array}{c} G_i=1, \\ H_i=c,Z_i\end{array}\right)\right) \\
= \; & \left(\frac{1}{N_1}\right)\sum_{i:G_i=1}\left(\sum_j 1\left\{j\neq d_{i,t-1};\text{Pov. Rate}_j \leq 10\%;\mathrm{R}_j^m<\rho_i\right\}\pi_{ij}^c\right)
\end{aligned}
$$

where $N_1 = \sum_i 1\left\{G_i = 1\right\}$ is the number of MTO households assigned to the Experimental group.

Regarding exposure conditional on take up we have

$$E\left[X|G_i=g,D_i=1\right] = \sum_h \Pr\left(H=h\right) E\left[X|G_i=g,D_i=1,H=h\right] \tag{28}$$

$$
\begin{aligned}
E\left[X|G_i=g,D_i=1,H=h\right] &= \sum_z E\left[X|G_i=g,D_i=1,H=h,Z_i\right]p\left(z|D_i=1\right) \\
&= \sum_{i:G=g,D_i=1}\left\{E\left[X|G_i=g,D_i=1,H=h,Z_i\right]\right\}\left(\frac{1}{N_{g,1}}\right) \\
&= \left(\frac{1}{N_{g,1}}\right)\sum_{i:G=g,D_i=1}\left\{E\left[X|G_i=g,D_i=1,H=h,Z_i\right]\right\} \\
&= \left(\frac{1}{N_{g,1}}\right)\sum_{i:G=g,D_i=1}\left\{\sum_j X\Pr\left(d_i=j\,\middle|\,\begin{array}{c}G_i=g,D_i=1,\\ H=h,Z_i\end{array}\right)\right\}
\end{aligned} \tag{29}
$$

where for $G=2$ and $H=v$ we have

$$
\Pr\left(d=j|G_i=2,D_i=1,H=v,Z_i\right)=\begin{cases} 0 & \text{if } j=d_{i,t-1} \\ \dfrac{\pi_{ij}^v}{\sum\limits_{k:k\neq d_{i,t-1}}\pi_{ik}^v} & \text{otherwise}\end{cases} \tag{30}
$$

for $G=2$ and $H=c$ we have

$$
\Pr\left(d=j|G_i=2,D_i=1,H=c,Z_i\right)=\begin{cases} 0 & \text{if } j=d_{i,t-1} \text{ or } \mathrm{R}_j^m>\rho_i \\ \dfrac{\pi_{ij}^c}{\sum\limits_{k:k\neq d_{i,t-1},\mathrm{R}_k^m<\rho_i}\pi_{ik}^c} & \text{otherwise}\end{cases} \tag{31}
$$

for $G=1$ and $H=v$ we have

$$
\Pr\left(d=j|G_i=1,D_i=1,H=v,Z_i\right)=\begin{cases} 0 & \text{if } \begin{cases} j=d_{i,t-1} \\ \text{or Pov. Rate}_j>10\% \end{cases} \\ \dfrac{\pi_{ij}^v}{\sum\limits_{k:k\neq d_{i,t-1};\text{Pov. Rate}_k<10\%}\pi_{ik}^v} & \text{otherwise}\end{cases} \tag{32}
$$

and for $G=1$ and $H=c$ we have

$$\Pr\left(d=j|G_i=1,D_i=1,H=c,Z_i\right)=\begin{cases} 0 & \text{if} \begin{cases} j=d_{i,t-1} \\ \text{or } \mathrm{R}_j^m>\rho_i \\ \text{or Pov. Rate}_j>10\% \end{cases} \\ \dfrac{\pi_{ij}^c}{\sum\limits_{\substack{k:k\neq d_{i,t-1} \\ k:\text{Pov. Rate}_k<10\%,\mathrm{R}_k^m<\rho_i}} \pi_{ik}^c} & \text{otherwise} \end{cases} \tag{33}$$