A Unified Empirical Framework to Study Neighborhood Segregation

Gregorio Caetano and Vikram Maheshri*

August 15, 2023

Abstract

We study the dynamics of race and income segregation in the San Francisco Bay area from 1990 to 2004. Our framework incorporates the endogenous feedback loop at the core of the seminal Schelling (1969) model of segregation into a dynamic model of neighborhood choice, which allows us to assess the relative importance of various sorting mechanisms that generate segregation (endogenous discriminatory sorting on the basis of the socioeconomic composition of neighbors, sorting on the basis of other neighborhood amenities, sorting on the basis of prices, initial conditions) and frictions that mediate these types of sorting (moving costs and incomplete information) in the short- and long-run. Identification of households' responses to prices and to the socioeconomic compositions of neighbors is facilitated by novel instrumental variables that exploit the logic of a dynamic choice model with frictions. Although racially discriminatory sorting is an important factor generating both racial and income segregation, its effects are substantially mitigated by frictions, especially moving costs. We also find that initial neighborhood conditions help explain segregation dynamics due to the presence of these frictions. Sorting on the basis of other amenities, including prices, have a sizable contribution to segregation, but discriminatory sorting on the basis of income plays little to no role in generating either racial or income segregation.

*University of Georgia and University of Houston. We gratefully acknowledge Patrick Bayer and Alvin Murphy for sharing their data and thank Milena Almagro, Patrick Bayer, Josh Kinsler, Tatjana Kleineberg, Robert McMillan, Maisy Wong, anonymous referees and various conference and seminar participants for their helpful comments. All errors are our own.

1 Introduction

Neighborhood demographics are often in a state of flux. In Figure 1, we present the evolution of the socioeconomic compositions of several San Francisco Bay Area neighborhoods over a fifteen year period. As suggested by these selected neighborhoods, there is rich heterogeneity in the trends of the race and income compositions across neighborhoods. What explains these trends? In his seminal work, Schelling (1969) proposed a concise answer to this question: the composition of neighborhoods may change due to the presence of discrimination, defined as households sorting on the basis of the race or income of their neighbors. If, for instance, Hispanic households prefer Hispanic neighbors relative to non-Hispanic ones, then an increase in the Hispanic share of a neighborhood might induce additional relative inflows of Hispanic households, which would in turn trigger further inflows of Hispanics in the future, and so on. This endogenous positive feedback loop could generate the observed serial correlation in socioeconomic composition we see in West Richmond (top left panel) by itself.

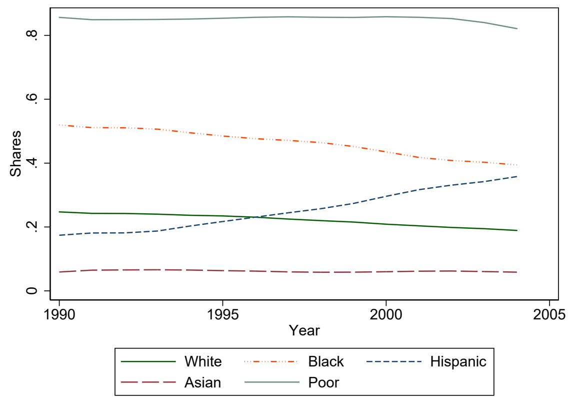
Meanwhile, a rich parallel literature on residential choice has developed to study sorting on the basis of local amenities (including the socioeconomic compositions of neighborhoods) and prices.[1] A common assumption in this literature is that neighborhoods are observed in *steady state*, i.e., in the absence of future amenity shocks, the demographic compositions of the neighborhoods will not change. This leaves no room for the endogenous mechanism discussed above, so the trends shown in Figure 1 would be attributed to serially correlated, exogenous changes in the amenities of these neighborhoods. For instance, we would conclude that some West Richmond amenity that Hispanics disproportionately like has gradually increased over the sample period in some manner outside of the model of residential choice.

This paper studies the determinants of segregation within a new framework that bridges these two literatures by incorporating into empirical models of residential choice the endogenous feedback loop that fuels the dynamics suggested by Schelling (1969). This allows us to study simultaneously the dynamic trajectories of all neighborhoods in a city toward steady state. This should be of particular interest to policymakers since any policy that influences the socioeconomic compositions of neighborhoods today may also impact the trajectories of their future socioeconomic compositions even if no other

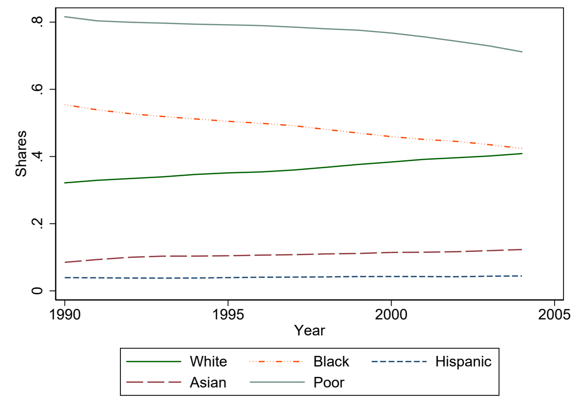
[1]See, for example, McGuire (1974); Epple, Filimon and Romer (1984); Kiel and Zabel (1996); Epple and Sieg (1999); Epple, Romer and Sieg (2001, 2003); Bayer, McMillan and Rueben (2004*a*,*b*); Bayer and Timmins (2005, 2007); Bayer, Ferreira and McMillan (2007); Bayer et al. (2016); Caetano (2019).

actions are taken later on, which implies that the short-run effect of such a policy may be completely different from its long-run effect.

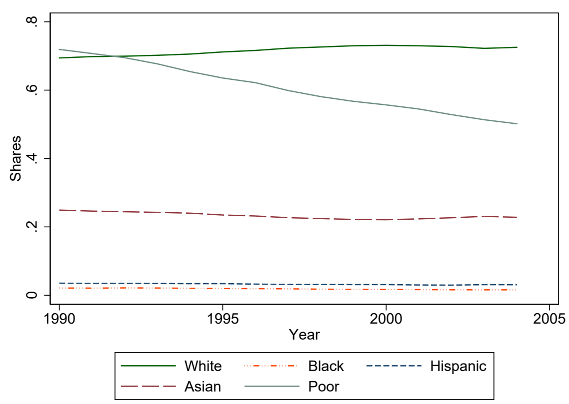
Figure 1: Socioeconomic Composition of Selected San Francisco Bay Area Neighborhoods Over Time, 1990-2004



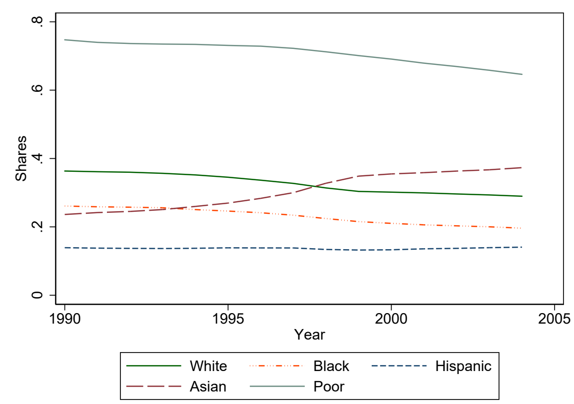
(a) West Richmond



(b) West Emeryville



(c) Allendale/SE Oakland



(d) Portola (SF)

Source: See Section 2.

We must modify certain standard assumptions in the residential choice literature to accommodate the endogenous Schelling mechanism. Specifically, we relax assumptions about households' expectations at the time they make their residential choices. Because these expectations impact the choices that households make, which in turn translate into future expectations of the socioeconomic compositions of neighborhoods, restrictions on how expectations are formed may influence the strength of the feedback loop. For instance, information frictions may attenuate the ability of discrimination to translate into segregation because of *ex-ante* uncertainty: each household may not individually respond as strongly to the socioeconomic compositions of neighborhoods because they may not be able to perfectly forecast the sorting decisions of others (leading to a coordination problem) or the extent to which amenities will change in the future due to changes in the socioeconomic compositions. This friction is reinforced by moving costs since it is costly to undo a choice that was made with an expectation that did not get realized.

Identifying this endogenous feedback loop requires us to isolate households' responses to the socioeconomic compositions of neighborhoods from their responses to other neighborhood amenities. To this end, we propose a new instrumental variable (IV) that follows the logic of a dynamic choice model with frictions. Our identification strategy relies on the assumption that information from the more distant past (e.g., two years ago) does not directly affect valuations of neighborhoods today conditional on valuations in the interim (e.g., in the past year). This translates into isolating the component of a neighborhood's socioeconomic composition that is due to *mismatched households*, i.e., those who currently reside in their neighborhood for reasons that are no longer relevant to new inflows; although they made optimal choices in the past given their expectations at that time, these households are now stranded in their homes because of moving costs in spite of that neighborhood having become less attractive to them in the meantime. We perform a detailed sensitivity analysis, including a Monte Carlo study, which together support the validity of the identification strategy in our application.

A key implication of our framework, from both a conceptual and practical standpoint, is that choice data alone is sufficient to study many important aspects of segregation dynamics. This follows directly from the fact that the demographic compositions of neighborhoods are the primitives of models of segregation (e.g., Schelling (1969); Becker and Murphy (2000); Bayer and Timmins (2005)), and they are ultimately determined by households' residential choices. Although neighborhood demographics may influence (and in turn be influenced by) other features of neighborhoods such as prices and amenities, there is information on these potentially complex mechanisms embedded in the choices that households' make. We offer empirical support for this implication by re-estimating a more complicated model with additional data on neighborhood prices and replicating our main findings.

We analyze a monthly data set of residential sales (Bayer et al. (2016)) across 224 neighborhoods in the San Francisco Bay Area from 1990-2004 that allows us to observe the heterogeneous sorting of eight socioeconomic groups over time: rich and poor Whites, Blacks, Hispanics and Asians. Our empirical framework combines estimation of a dynamic model of neighborhood choice with a simulation procedure that allows us to isolate specific determinants of segregation (and their interactions) by analyzing relevant counterfactuals. Our model can be summarized as follows: First, households form expectations about the endogenous characteristics of neighborhoods (their race and income compositions, and their expected effects on prices and other amenities) as well as unobserved neighborhood characteristics. Based on these expectations, they decide if they should move and, if so, to which neighborhood they will move. Households of different socioeconomic groups may sort heterogeneously on the basis of both the expected race and income compositions of their neighbors and unobserved amenities. All of this sorting is mediated by two frictions: moving costs and incomplete information. We estimate two key sets of parameters for each of the eight socioeconomic groups of households: their moving costs and their responses to different types of neighbors. Following Bayer et al. (2016), we identify moving costs from the decisions of households who chose to move instead of staying in their current houses. We identify the responses of households to their neighbors with the IV strategy discussed above. In an extension to our analysis, we also estimate two additional sets of parameters: households' responses to neighborhood prices, and the reduced form effect of how neighborhood demographics are capitalized into prices.

Given these estimated parameters and our model, we simulate what would happen endogenously to the socioeconomic compositions of neighborhoods under various counterfactuals that include: different initial allocations of households across neighborhoods, different responses to neighbors (e.g., no race and/or income discrimination), different levels of moving costs, and different price sensitivities. We explicitly consider the fact that sorting today affects the choices of households next month, which in turn may affect the choices of households in two months, etc., repeating this endogenous feedback loop indefinitely until a new steady state is reached. By simulating this entire dynamic resorting process, we can uncover the resulting trajectories of neighborhoods under these different counterfactuals. A comparison of trajectories across counterfactuals allow us to identify the relative roles of each factor in explaining segregation.

We find robust evidence of discriminatory sorting: households tend to respond positively to neighbors of the same race and income, though to different degrees depending on their race and their income. There is also substantial heterogeneity in the responses to neighbors of other types, and some of these responses are not reciprocated, which leads to complex dynamics. Racial discrimination is quantitatively meaningful, but income discrimination is not; we do find some evidence of heterogeneous sorting due to prices. We also find that responses to unobserved amenities contribute to segregation, though less so than racial discrimination. Nevertheless, all of these drivers of segregation, especially discrimination, are mitigated by frictions. Absent moving costs, there would be much more sorting across neighborhoods, which would dramatically reshape their socioeconomic compositions. This in turn would trigger further discriminatory sorting, which would grow in prominence in the long-run.

Relevant Literature

Our paper bridges two distinct but related literatures on residential choice and segregation. It also contributes to a growing literature on the causes of residential segregation. We briefly review some of the most relevant studies.

Empirical Models of Residential Choice and Neighborhood Sorting

Because segregation is an outcome of neighborhood sorting, we build upon the prolific literature on the determinants of residential choice.[2] This literature is largely interested in estimating the marginal willingness to pay for neighborhood amenities. Three papers in this literature are particularly related to our study. Bayer, McMillan and Rueben (2004*a*) develop a framework to estimate horizontal models of neighborhood choice by building on insights from the empirical industrial organization literature (Berry (1994); Berry, Levinsohn and Pakes (1995)). This framework has been widely applied and extended in this literature (e.g., Bayer, Ferreira and McMillan (2007); Bayer, Keohane and Timmins (2009); Ferreira (2010); Bayer and McMillan (2012); Ringo (2013); Bayer et al. (2016); Caetano (2019)). Bayer and Timmins (2005) study the existence and uniqueness of equilibrium in sorting models with endogenous amenities such as the demographic composition of a neighborhood; Bayer and Timmins (2007) discuss estimation in empirical models like these and suggest an IV approach for identification based on the logic of a static model of neighborhood choice.

Following this literature, we employ a discrete choice framework that enables us to study the relative importance of racial and income compositions, prices and unob-

[2]Kuminoff, Smith and Timmins (2013) provide a comprehensive review of the growing literature on neighborhood sorting.

served amenities in explaining the sorting patterns that lead to segregation. This also allows us to embed moving costs as an additional friction that prevents sorting. A key departure lies in our weakening of assumptions on households' expectations when residential decisions are made. This is crucial for our purpose, as it renders our approach compatible with residential choices being observed in the process of convergence toward steady state via the endogenous mechanism described above. Another related departure is that we take a different strategy to estimate a dynamic model of residential choice with moving costs. Although this is not the first paper to do so in the context of neighborhood choice (e.g., Bayer et al. (2016) and Caetano (2019)), we show that many standard assumptions in dynamic demand estimation can be avoided when the goal is to study segregation (as opposed to uncovering the value of amenities, as is typical in these studies). Finally, the IV approach that we develop is novel, and it follows from the logic of a dynamic model of neighborhood choice with frictions. These IVs can be created with no additional data requirements and they can also be used to identify price responses.

Dynamic Models of Segregation

A largely theoretical literature based on the seminal Schelling model (Schelling (1969, 1971)) has sought to explore how segregation can arise and evolve when households sort on the basis of the characteristics of their neighbors. In the Schelling model, heterogeneous agents select where to live by simple rules of thumb. Although this purely heuristic model is not explicitly based on the optimization of an objective, it generates a valuable insight into a fundamental social force that may drive segregation: agents of different types react systematically differently to the composition of their neighbors. Schelling also explicitly models a friction, namely myopia, to ensure that neighborhoods gradually evolve toward a steady state.

Subsequent theoretical papers have embedded this intuition into a more standard economic framework (e.g., Becker and Murphy (2000); Bayer and Timmins (2005)), and there have been some recent attempts to estimate these models of segregation in reduced-form and structural contexts (e.g., Card, Mas and Rothstein (2008a); Banzhaf and Walsh (2013); Caetano and Maheshri (2017, 2023)). Banzhaf and Walsh (2013) discuss the role of other amenities in generating segregation under no moving costs. Caetano and Maheshri (2017) and Caetano and Maheshri (2023) study school segregation in a framework that embeds the key insight of Schelling (1969) and discuss how policies may have completely different effects on segregation in the short- and the long-run because of the endogenous feedback loop. In this paper, we generalize and extend that framework in four directions. First, we analyze segregation along multiple dimensions simultaneously, studying both racial and income segregation. Second, we make fewer assumptions on households' expectations, thus imposing fewer restrictions on the way that race and income compositions of neighborhoods may evolve. Third, we explicitly model realistic frictions such as moving costs, which motivates novel IVs. Finally, we extend the framework to account for heterogeneous and endogenous responses to prices.

Causes of Residential Segregation

There is a large empirical literature on the causes of residential segregation with deep historical roots that span the social sciences. Massey and Denton (1993) build on the observations of Myrdal (1944) and articulate numerous causes of racial segregation in the United States, including explicit housing policies in the early twentieth century in both Northern and Southern cities, prejudice, and more modern structural biases in the real estate market (e.g., unequal mortgage access and discrimination by realtors). Indeed, a number of recent studies have found causal impacts on neighborhood segregation of specific institutional features such as historical zoning regulations and covenants (Shertzer, Twinam and Walsh (2021)), redlining (Aaronson, Hartley and Mazumder (2020)) and stringent credit standards (Ouazad and Rancière (2016)).[3] Cutler, Glaeser and Vigdor (1999) conclude that collective action on the part of Whites in developing both legal and extra-legal institutions explained much of segregation in the mid 20th century, though by the end of the century, individual action in the form of decentralized sorting decisions was the main driver of urban segregation.

This paper encompasses all of these potential mechanisms at a broader level and places them in a framework in which their interactions and dynamic implications can be assessed. Our findings that moving costs and incomplete information are key impediments to speedy (de)-segregation can explain why institutions that have in many cases been eliminated may still play an important role in current segregation levels. Our finding of a modest role for current discrimination (in contrast to sorting on the basis of other amenities as well as initial conditions, which may themselves have been brought up by former discrimination) is in line with Ihlanfeldt and Scafidi (2002), who find that

[3] Li (2023) considers the effects of all historical constraints broadly defined.

prejudice is significant but unimportant in explaining Black residential segregation.

The rest of the paper proceeds as follows. In Section 2, we describe our data. In Section 3 we present an empirical model of neighborhood segregation, articulate the specific assumptions required for identification, and discuss the estimation and simulation of this model. We present our baseline results in Section 4 and consider different counterfactuals in order to assess the importance of various determinants of segregation in Section 4.2. In Section 5, we extend our framework to explicitly incorporate prices and other amenities and present additional results before concluding in Section 6.

2 Data

We use a monthly sample of all San Francisco Bay Area neighborhoods from January 1990 to November 2004. We define the San Francisco Bay Area as the six core counties (Alameda, Contra Costa, Marin, Santa Clara, San Francisco and San Mateo counties) that comprise the major cities of San Francisco, Oakland and San Jose and their surroundings, which are divided into neighborhoods by merging contiguous Census tracts until each resulting neighborhood contains approximately 10,000 households. Those neighborhoods with fewer than six annual home sales in our sample period are dropped leaving a total of 224 neighborhoods.

For each neighborhood in each month, we compute estimates of their race and income composition following the approach described in Bayer et al. (2016). Because high frequency data on the socioeconomic composition of neighborhoods is unavailable from standard sources (e.g., the Census) we must merge information from two main sources in order to construct these variables. The first source is Dataquick Information Services, a national real estate data service. Dataquick provides a detailed listing of all real estate transactions in the Bay Area including buyers' and sellers' names, buyer's mortgage information and property locations. The second source is a a dataset on mortgage applications published in accordance with the Home Mortgage Disclosure Act (HMDA) of 1975. Notably, HMDA data contains demographic information on mortgage applicants and the locations of properties that the applicants are buying. By linking these datasets on buyer's mortgage information and property locations, we can estimate how the demographics of neighborhoods change with each real estate transaction. With neighborhood-level estimates of the flows of households of different groups, we estimate the actual socioeconomic composition of each neighborhood by anchoring our flow estimates to the actual socioeconomic composition of each neighborhood per the 1990 US Census.[4]

We classify households into eight groups on the basis of four races (Whites, Blacks, Hispanics and Asians) and two income designations (rich or poor, depending on whether household income is greater than $50,000 in 1990 dollars).[5] For expositional simplicity, we refer to Hispanics as a race rather than an ethnicity, and the other three racial groups include only non-Hispanic households.[6] For each race-income group g, neighborhood j and month t, we observe the total number of homeowners, the total numbers of homeowners who moved into a new house, and the total number of homeowners who stayed in the same home since last month.[7] We also observe the total number of households of each group who chose to exit the Bay Area homeownership market in each month.[8] Finally, we compute monthly neighborhood prices by averaging the sales prices of all transactions observed in the HMDA data.

We summarize our data in Table 1. The majority of homeowners in the Bay Area are White, although there are sizable Asian and Hispanic populations as well. Roughly 47% of homeowners in the Bay Area are classified as rich, though this share is much smaller for Blacks and Hispanics. The socioeconomic compositions of neighborhoods also change over time in our sample as reflected in monthly inflow rates ranging from 0.1% for poor Whites to 0.7% for rich Asians.

The high variance in the average number of homeowners of each group reflects substantial cross-sectional heterogeneity in the socioeconomic composition of neigh-

[4]Bayer et al. (2016) report the results of several diagnostic tests that ensure the validity of this estimation procedure.

[5]We obtain the race and income of the original stock of households as of 1990 from the 1990 Census. From 1990 onward, all changes in the income of the neighborhood are measured based on income data from HMDA deflated to 1990 levels. We chose an income threshold of $50,000 because it resulted in the most balance of rich and poor among all available thresholds in the 1990 Census.

[6]We are unable to observe populations at the race-ethnicity-income group-tract level in the 1990 Census. Instead, we are able to observe populations at the race-income group-tract level, at the ethnicity-income group-tract level, and at the race-ethnicity-tract level. As such, our raw counts of rich and poor Whites, Blacks and Asians in each neighborhood include Hispanics. To address this, we reweight each group uniformly across neighborhoods to ensure that the number of rich Whites plus the number of poor Whites is equal to the number of non-Hispanic Whites (and do the same for Blacks and Asians), and we uniformly reweight each group to ensure that the number of rich Hispanics plus the number of poor Hispanics is equal to the total number of Hispanics. Our results are effectively unchanged if we assume all Hispanics to be White and adjust the population numbers accordingly.

[7]Households who move between houses within the same neighborhood counted as inflows (but not stayers).

[8]They are the households who are observed to move out of some neighborhood in $t-1$ but not observed to move into any neighborhood in t.

borhoods, i.e., segregation. We calculate the dissimilarity index for each of the eight socioeconomic groups defined by race and income and summarize it in Table 1.[9] We choose this widely used measure of segregation because it is easy to interpret. For instance, a rich White dissimilarity index of 0.29 indicates that 29% of rich Whites would have to be relocated (holding all other households' locations fixed) in order to distribute them evenly across all Bay Area neighborhoods (i.e., to ensure that the share of rich Whites was the same in all neighborhoods). The index ranges from zero to one, and a higher value means that households of a given socioeconomic group are more concentrated in certain neighborhoods. Blacks are the most concentrated racial group, followed distantly by Asians, Hispanics and Whites. While rich Whites and Asians tend to be more concentrated than their poor counterparts, the opposite is true for rich Blacks and Hispanics.

Finally, the average neighborhood price in our sample is \$329,000 with a standard deviation of \$232,000. There is considerable appreciation over our sample period, as the average price rises from \$248,000 in 1990 to \$564,000 in 2004 (all prices in constant November 2004 dollars).

---

[9] If N_{gj} is the total number of group g households residing in neighborhood j, then the dissimilarity index for group g households is defined as $\frac{1}{2}\sum_j \left|\frac{N_{gj}}{\sum_k N_{gk}} - \frac{N_j - N_{gj}}{\sum_k (N_k - N_{gk})}\right|$ where $N_j = \sum_g N_{gj}$. Note that a group may corresponds to a race-income combination (e.g., rich Whites), a race (e.g., rich Whites plus poor Whites) or an income level (e.g., poor households of all races).

Table 1: Summary Statistics

	White		Black		Hispanic		Asian	
	Rich	Poor	Rich	Poor	Rich	Poor	Rich	Poor
Share of Homeowners	0.38	0.33	0.02	0.04	0.04	0.06	0.08	0.05
	(0.24)	(0.22)	(0.02)	(0.04)	(0.04)	(0.06)	(0.07)	(0.05)
Average Num. of Homeowners	2,196	1,879	108	244	217	340	471	315
	(2,089)	(1,161)	(145)	(370)	(280)	(330)	(342)	(342)
Average Monthly Inflows	8.71	2.36	0.38	0.30	1.26	0.90	3.58	1.13
	(10.57)	(3.36)	(0.92)	(0.78)	(2.71)	(1.73)	(6.60)	(2.02)
Average Monthly Stayers	2,187	1,877	108	243	216	340	467	314
	(2,081)	(1,160)	(145)	(370)	(278)	(329)	(702)	(341)
Dissimilarity Index	0.29	0.19	0.41	0.57	0.27	0.34	0.36	0.33
Average Neighborhood Price (Thousands of 2004 Dollars)	329							
	(232)							
Num. of Observations	39,872							

Note: Each observation is a neighborhood-month from January 1990 to November 2004. Poor households have an income of less than $50,000 in 1990 dollars. Standard deviations are presented in parentheses.

3 Empirical Framework

In this section, we first present a dynamic model of neighborhood choice and propose novel instrumental variables to identify key parameters of the model (Section 3.1), then describe how the parameters of the model are estimated and used to study segregation dynamics via simulation (Section 3.2). We conclude with a detailed discussion of the interpretation of model parameters (Section 3.4). Importantly, this framework requires only a panel of socioeconomic compositions of neighborhoods; segregation dynamics can

– and should, as we argue below – be studied without controlling for other neighborhood amenities and characteristics. In Section 5, we discuss how additional information, in particular data on prices, can be used to study additional related questions.

3.1 A Dynamic Model of Neighborhood Choice

A city is divided into J neighborhoods. At the beginning of each period t, household i observes state variable $\boldsymbol{b_{it}}$ as well as where they are currently located, j^*_{it}. They use this information to form expectations of their value of residing in each neighborhood, and then choose where to reside in order to maximize their expected utility. Specifically, household i faces the dynamic optimization problem

$$\max_{j_{i\tau}\in\mathbb{J}} \mathbb{E}\left[\sum_{\tau=t}^{\mathcal{T}} \delta^{\tau-t}\cdot\mathbb{E}[u_{ij_{i\tau}\tau}|\boldsymbol{b_{i\tau}}]\,|j^*_{it},\boldsymbol{b_{it}}\right], \tag{1}$$

where $j_{i\tau}$ and $\boldsymbol{b_{i\tau}}$ are the choice and state variables of household i in period τ respectively, $u(\cdot)$ is their flow indirect utility function, $\mathcal{T}$ is their time horizon, and δ is their inter-temporal discount factor. (Hereafter, we refer to all vectors and matrices in bold type.) $\mathbb{J}$ is each household's choice set, which includes moving to one of the $j\in\{1,\ldots,J\}$ neighborhoods, remaining in their current house ($j=J+1$), or leaving the city entirely ($J=0$).[10]

$\mathbb{E}[u_{ij\tau}|\boldsymbol{b_{i\tau}}]$ corresponds to the flow-utility that household i *expects* to obtain (just prior to making their choice in τ) if they choose neighborhood j, with *state variable* $\boldsymbol{b_{i\tau}}$ denoting the information set used to form that expectation. Our notation differs slightly from the standard literature to make clear the role of expectations in neighborhood sorting. Although $\mathbb{E}[u_{ij\tau}|\boldsymbol{b_{i\tau}}]$ is typically written as $u(j,\boldsymbol{b_{i\tau}})$, we want to make explicit the fact that the flow-utilities are not observed before decisions are made.[11] Note that so far we have not imposed any restrictions on the structural primitives of the expectation of flow-utility since $\boldsymbol{b_{i\tau}}$ is allowed to be unobserved and may vary freely across i and τ.

We define the value function as $V(\boldsymbol{b_{it}}) = \max_{j\in\mathbb{J}} \mathbb{E}[v_{ijt}|\boldsymbol{b_{it}}]$, where v_{ijt} refers to the cumulative utility of household i for choosing neighborhood j in t. The expectation of

[10] Here we borrow notation for the choice set from Bayer et al. (2016); as in their paper, we only observe data on homeowners, so in our application, $j=0$ also includes the outside option of renting within the city.

[11] To make our notation otherwise comparable with the standard literature, we follow the same notation as Aguirregabiria and Mira (2010)'s well-known survey of the dynamic discrete choice literature.

this cumulative utility given $\boldsymbol{b_{it}}$ is written as

$$\mathbb{E}[v_{ijt}|\boldsymbol{b_{it}}] = \mathbb{E}[u_{ijt}|\boldsymbol{b_{it}}] + \int \delta \cdot V\left(\boldsymbol{b_{it+1}}\right) dF_b\left(\boldsymbol{b_{it+1}}\,|j, \boldsymbol{b_{it}}\right). \quad (2)$$

$F_b\left(\boldsymbol{b_{it+1}}\,|j, \boldsymbol{b_{it}}\right)$ is the expected distribution of the state variable in $t+1$ conditional on the choice and the state variable from t. Next, we make the standard assumptions in the dynamic discrete choice literature (e.g., Aguirregabiria and Mira (2010)):

Assumption 1. (Additive Separability, Logit Error, Conditional Independence)

1. *$\mathbb{E}\left[u_{ijt}|\boldsymbol{b_{it}}\right] = \mathbb{E}\left[u_{ijt}|\boldsymbol{x_{it}}\right] + \epsilon_{ijt}$ where ϵ_{ijt} is the jth element of $\boldsymbol{\epsilon_{it}}$.*

2. *ϵ_{ijt} is i.i.d. extreme value type I.*

3. *$F_x\left(\boldsymbol{x_{it+1}}\,|j, \boldsymbol{x_{it}}, \boldsymbol{\epsilon_{it}}\right) = F_x\left(\boldsymbol{x_{it+1}}\,|j, \boldsymbol{x_{it}}\right)$ where $F_x(\cdot)$ is the cumulative density function of x.*

Assumption 1 implies

$$\mathbb{E}[v_{ijt}|\boldsymbol{b_{it}}] = \underbrace{\mathbb{E}[u_{ijt}|\boldsymbol{x_{it}}] + \int \delta \cdot \overline{V}\left(\boldsymbol{x_{it+1}}\right) f_{\boldsymbol{x}}\left(\boldsymbol{x_{it+1}}\,|j, \boldsymbol{x_{it}}\right)}_{\mathbb{E}[v_{ijt}|\boldsymbol{x_{it}}]} + \epsilon_{ijt}, \quad (3)$$

where $\overline{V}\left(\cdot\right)$ is the *integrated value function.*[12]

At the beginning of period t, households observe the state variable $\boldsymbol{b_{it}}$, form expectations and then choose (1) whether or not to move, and upon deciding to move they choose (2) an option in $\mathbb{J} = \{0, \ldots, J\}$. We classify households into G different demographic groups on the basis of their race and income, and impose the following additional restriction:

Assumption 2. *Let $\mathbf{1}(\cdot)$ be the indicator function and $\boldsymbol{x_{it}} = (j_{it-1}, \boldsymbol{x_{gt}})$. Then*

$$\mathbb{E}[v_{ijt}|\boldsymbol{x_{it}}] = \mathbf{1}\left(j \in \{0, \ldots, J\}\right) \cdot \left(\mathbb{E}\left[v_{gjt}|\boldsymbol{x_{gt}}\right] - \phi_g\right) + \mathbf{1}\left(j = J+1\right) \cdot \mathbb{E}\left[v_{gjt}|\boldsymbol{x_{gt}}\right] \quad (4)$$

Assumption 2 states that different households of the same group g are allowed to differ from each other only via their previous choice (j_{it-1}) beyond their idiosyncratic

[12] This function, defined as $\overline{V}\left(\boldsymbol{x_{it}}\right) = \int V\left(\boldsymbol{x_{it}}, \boldsymbol{\epsilon_{it}}\right) dG_{\boldsymbol{\epsilon}}(\boldsymbol{\epsilon_{it}})$, is the unique solution to the integrated Bellman equation $\overline{V}(\boldsymbol{x_{it}}) = \int \max_{j\in\mathbb{J}} \Big(\mathbb{E}[u_{ijt}|\boldsymbol{x_{it}}] + \epsilon_{ijt} + \delta \cdot \sum_{\boldsymbol{x_{it+1}}} \overline{V}\left(\boldsymbol{x_{it+1}}\right) f_x\left(\boldsymbol{x_{it+1}}\,|j, \boldsymbol{x_{it}}\right)\Big) dG_\epsilon(\boldsymbol{\epsilon_{it}})$, where $G_\epsilon(\boldsymbol{\epsilon_{it}})$ is the extreme value type I cumulative density function.

error $\boldsymbol{\epsilon_{it}}$.[13] It also implicitly states that moving costs do not vary within group or depend on the neighborhood of origin (j_{it-1}) and destination (j_{it}). These assumptions follow from Bayer et al. (2016) and are needed due to data limitations since we only observe variables at the group-neighborhood-month level.[14]

The term $\mathbb{E}\left[v_{gjt}|\boldsymbol{x_{gt}}\right]$ is the moving-cost-free component of $\mathbb{E}[v_{ijt}|\boldsymbol{x_{it}}]$ for households of group g. Defining $v^e_{gjt} = \mathbb{E}\left[v_{gjt}|\boldsymbol{x_{gt}}\right]$, $\boldsymbol{s^e_{jt}} = \mathbb{E}[\boldsymbol{s_{jt}}|\boldsymbol{x_{gt}}]$ and $\xi^e_{gjt} = \mathbb{E}[\xi_{gjt}|\boldsymbol{x_{gt}}]$, we decompose:

$$v^e_{gjt} = \boldsymbol{\beta'_g}\boldsymbol{s^e_{jt}} + \xi^e_{gjt}, \tag{5}$$

where $\boldsymbol{s_{jt}}$ is a vector of observed demographic shares representing the socioeconomic composition of each neighborhood and period, and ξ_{gjt} encompasses all other neighborhood amenities that are valued by households. It is important to note that equation (5) is written from the perspective of the household, *not* the econometrician, as the three variables reflect the expectations of households when they make their choice. Even if $\boldsymbol{x_{gt}}$ is not observed, v^e_{gjt} can be identified from choice data using standard approaches.[15] However, note that both variables on the right-hand-side are unobserved to the econometrician. In particular, while $\boldsymbol{s_{jt}}$ is observed, $\boldsymbol{s^e_{jt}}$ is unobserved.

Next, we impose assumptions that allow us to identify $\boldsymbol{\beta_g}$, the effect of the expected socioeconomic composition on the expected cumulative value of the neighborhood. We rewrite equation (5) based on observed quantities as

$$v^e_{gjt} = \boldsymbol{\beta'_g}\boldsymbol{s_{jt}} + \xi_{gjt}, \tag{6}$$

where $\xi_{gjt} = \xi^e_{gjt} + \boldsymbol{\beta'_g}(\boldsymbol{s^e_{jt}} - \boldsymbol{s_{jt}})$. This is analogous to equation (5) from the perspective of the econometrician since v^e_{gjt} can be identified (see below) and $\boldsymbol{s_{jt}}$ is observed. Throughout the remainder of the paper, we refer to ξ_{gjt} as simply "amenities", although

[13] Note that j_{it-1}, which is a component of $\boldsymbol{x_{it}}$, implicitly appears in equation (4): the last term of equation (4) is equivalent to $\mathbf{1}\left(j = j_{it-1}\right) \cdot \mathbb{E}\left[v_{gj_{it-1}t}|\boldsymbol{x_{gt}}\right]$.

[14] One can allow for heterogeneous parameters across individuals of the same group by observables such as wealth or the neighborhood of origin (e.g., Bayer et al. (2016)), but doing so turned out to be infeasible in our context because there are not enough households of certain types inside the same group. Because the observed distribution of residential choices varies greatly by socioeconomic group, some of the heterogeneity in moving costs by neighborhood of origin may be incorporated into our estimates of ϕ_g.

[15] This point is obvious, since $\boldsymbol{x_{gt}}$, whatever it may be, is the information set that the econometrician assumes was used to make choices as observed in the data. This point is typically left implicit in discrete choice models since it is often not made explicit that the left-hand-side of the equation involves an expectation conditional on a generic information set.

in reality it encompasses both amenities other than socioeconomic composition (ξ^e_{gjt}) as well as any forecast error of $\boldsymbol{s_{jt}}$ using information set $\boldsymbol{x_{gjt}}$ (e.g., from differences in the information sets of different types of households).

Instrumental Variables

In order to identify $\boldsymbol{\beta_g}$, we impose a restriction on the error term ξ_{gjt}.

Assumption 3. *IV Validity. For all g and t, there exists some $T > 0$ such that*

$$\mathbb{E}\left[\xi_{gjt} \left| \boldsymbol{v^e_{jt-T}}, \boldsymbol{s_{jt-(T+1)}}\right.\right] = \mathbb{E}\left[\xi_{gjt} \left| \boldsymbol{v^e_{jt-T}}\right.\right] \tag{7}$$

Assumption 3 states that sorting in t on the basis of amenities (ξ_{gjt}) does not use information that was used in the distant past (in the form of $\boldsymbol{s_{jt-(T+1)}}$) unless it is already embedded in more recent valuations ($\boldsymbol{v^e_{jt-T}} = (v^e_{1jt-T}, ..., v^e_{Gjt-T})$). Note that this assumption still allows households to use information from the past to form their expectation of ξ_{jt}. It simply requires that any information from the distant past ($t-(T+1)$ or before) that informs sorting on the basis of amenities in t must have been used by households that sorted more recently, in $t-T$.

Assumption 3 suggests a straightforward strategy to estimate $\boldsymbol{\beta}$ by instrumental variables: we can use $\boldsymbol{s_{jt-(T+1)}}$ as an instrument for $\boldsymbol{s_{jt}}$ in equation 5, provided that we control for $\boldsymbol{v^e_{jt-T}} = (v^e_{1jt-T}, ..., v^e_{Gjt-T})$. This strategy can only be implemented if the instrument is relevant, or more formally:

Assumption 4. *IV Relevance. For all g and t, and for some value of T that satisfies Assumption 3,* $\mathbb{E}\left[\boldsymbol{s_{jt}} \left| \boldsymbol{v^e_{jt-T}}, \boldsymbol{s_{jt-(T+1)}}\right.\right] \neq \mathbb{E}\left[\boldsymbol{s_{jt}} \left| \boldsymbol{v^e_{jt-T}}\right.\right]$.

Assumption 4 states that $\boldsymbol{s_{jt-(T+1)}}$ and $\boldsymbol{s_{jt}}$ are correlated to each other even conditional on $\boldsymbol{v^e_{jt-T}}$. This is the case because incomplete information and moving costs imply that some households will likely be mismatched to their neighborhood at any given point in time. To see this, note that because households lack perfect foresight, some households residing in a neighborhood as of $t-(T+1)$ would have sorted there in the past because of information that turned out to be orthogonal to $\boldsymbol{v^e_{jt-T}}$ (i.e., the information that led those households to sort to that neighborhood turned out to be irrrelevant to later inflows in $t-T$). These households now reside in a less-than-ideal neighborhood in t (they are mismatched), yet many of them remain there because moving is costly, thus they still contribute to the socioeconomic composition s_{jt}.

We formalize this intuition with the choice model above. Consider a generic period in the distant past, $\tau \le t-(T+1)$. Let $\boldsymbol{\omega_{gj\tau}} = \boldsymbol{x_{g\tau}} - \mathbb{E}\left[\boldsymbol{x_{g\tau}} \middle| \boldsymbol{v^e_{jt-T}}\right]$. Note that $\boldsymbol{v^e_{jt-T}}$ is influenced by information from periods *after* τ, so $\boldsymbol{\omega_{gj\tau}}$ corresponds to the portion of the information in τ that was later found out to be wrong given the information that was available to households in the subsequent period $(t-T)$. Indeed, some households who sorted in τ may now live in less-than-ideal neighborhoods precisely because of $\boldsymbol{\omega_{gj\tau}}$. These households chose their neighborhood using information that turned out to be irrelevant for future inflows (e.g., because of a possible state of the world that was unrealized). Some of these households will not have "fixed" their *ex post* mistake by t since their level of dissatisfaction does not exceed the cost of moving. They are precisely the mismatched households described above. Because the demographics of these mismatched households contribute to $\boldsymbol{s_{jt}}$, some erroneous information that they used ($\boldsymbol{\omega_{gj\tau}}$) contributes to $\boldsymbol{s_{jt}}$. However, since $\boldsymbol{\omega_{gj\tau}}$ turned out to be irrelevant to future inflows, it is not used to sort on the basis of ξ_{gjt} by Assumption 3. Note that even though we do not observe $\boldsymbol{x_{g\tau}}$ (and hence $\omega_{gj\tau}$), we can still isolate the variation in $\boldsymbol{s_{jt}}$ stemming from $\omega_{gj\tau}$ for all $\tau \le t-(T+1)$ and for all g by using the component of $s_{t-(T+1)}$ that is orthogonal to $v^e_{1jt-T}, ..., v^e_{Gjt-T}$ as an IV.

In a scenario with complete information, $\boldsymbol{\omega_{gj\tau}}$ would not exist: households would perfectly anticipate the evolution of neighborhoods, and hence no information would ever turn out to be wrong *ex post*. As a result the IVs would be irrelevant. However, incomplete information alone is insufficient to guarantee relevance. Without moving costs, households would immediately and freely re-optimize in the face of information that turned out to be wrong *ex post*, so $\boldsymbol{\omega_{gj\tau}}$ would not persist for more than one period. It is precisely these two frictions – incomplete information and moving costs – that jointly imply the IVs are relevant. Relevance of the IV is of course a testable assumption, and we show in Section 4 that it is satisfied.

Remark 1. The exclusion restriction on $\boldsymbol{\xi}$ (Assumption 3) reflects the idea that households in t do not use past information (from before $t-T$) in a more sophisticated manner than households in $t-T$. This is a restriction on the *relative* level of sophistication, not on the *absolute* level, so it is consistent with many formulations of expectations ranging from the narrowly myopic households of Schelling (1969) to highly sophisticated households. For instance, consider households with rational expectations who use their information set in the best way possible (their forecast errors are orthogonal to their information set). Let decisionmakers in t form their expectations with information from the last τ periods, and decisionmakers in $t-T$ form their expectation with information from the last τ' periods. Then Assumption 3 is compatible with any values of τ and τ' provided that $\tau \leq \tau' + T$. In particular, households in t are allowed to be somewhat more sophisticated than those in $t-T$ (i.e., $\tau > \tau'$ is allowed).[16]

Remark 2. Standard dynamic discrete choice approaches often parametrically specify the transition probability $f_{\boldsymbol{x}}\left(\boldsymbol{x_{it+1}}|j, \boldsymbol{x_{it}}\right)$ from equation (3) and assume $\boldsymbol{x}$ is observed or estimable by the econometrician.[17] We want to avoid such assumptions in our context. Because $\boldsymbol{x_{it}} = (j_{it-1}, \boldsymbol{x_{gt}})$, and $\boldsymbol{x_{gt}}$ determines the expected compositions of neighborhoods, such assumptions would restrict how neighborhood segregation would evolve over time. Because observed past choices reflect past information sets, we instead relate the information set in $\boldsymbol{x_{gt}}$ with past information sets in $\{\boldsymbol{x_{g't-T}}\ \forall g'\}$, allowing us to state the validity assumptions without explicitly stating what must be included or excluded in $\boldsymbol{x_{gt}}$. By not connecting $\boldsymbol{x_{gt}}$ to what econometricians observe implies, for instance, that the time horizon ($\mathcal{T}$) and the inter-temporal discount factor (δ) may vary across groups, and neither needs to be observed or identified. No further restriction on information sets is required beyond Assumptions 1, 2, 3.

3.2 Estimation

Our estimation strategy unfolds in two stages: we first estimate v_{gjt} and ϕ_g for all g, j and t (stage 1) and then we estimate $\boldsymbol{\beta_g}$ for all g (stage 2). Before describing our strategy in more detail, it is useful to describe exactly what data is required. First, we require population counts of each group in each neighborhood in each period, which we

[16] Although at first our identification strategy might look similar to strategies used in the production function literature, such as the "proxy variable" literature (e.g., Olley and Pakes (1996)) and the "dynamic panel" literature (e.g., Arellano and Bond (1991)), there are important differences. In our setup, there is a distinct asymmetry between the outcome variable (v^e_{gjt}) and the main explanatory variable ($\boldsymbol{s_{jt}}$): while v^e_{gjt} reflects the decisions of those who are choosing a new neighborhood in t, $\boldsymbol{s_{jt}}$ reflects the decisions of many other households as well (e.g., households who made past choices), which may have been mediated by moving costs and different information sets. We exploit this asymmetry to build an identification strategy that relates the information used by households of one group in t with the information used by all past decision makers. Thus, we need only restrict the information used by households in t *relative* to the information used by households in the past. In contrast, identification in the production function literature exploits *absolute* restrictions in the information sets of decision makers (e.g., firms). See Ackerberg (2020) for an illuminating discussion of the identifying assumption made in that literature. Our IV approach is also very different from the shift-share IV approach (e.g., Bartik (1991)). Although both IVs use past shares, ours assumes exogeneity of them only conditional on the valuations in period $t-T$, whereas shift-share IVs assume exogeneity of them unconditionally as discussed by Goldsmith-Pinkham, Sorkin and Swift (2020).

[17] See, Aguirregabiria and Mira (2010) for a survey of the literature.

denote as N_{gjt}. From this, we can derive the socioeconomic composition of residents, the $G-1\times 1$ vector $\boldsymbol{s_{jt}}$ whose gth element represents the share of group g:

$$s_{gjt}=\frac{N_{gjt}}{\sum_{g'}N_{g'jt}} \tag{8}$$

In addition, we require information on the total number of inflows of each group into each of the J neighborhoods, which we denote as I_{gjt}. Note that I_{gjt} cannot be expressed solely in terms of N_{gjt} and N_{gjt-1}, since it also embeds new information about the number of households who chose to remain in the same house from $t-1$ to t.

Stage 1: Estimation of v^e_{gjt} and ϕ_g

This stage follows closely from Bayer et al. (2016). First, we use the choices of only those who moved in period t to estimate the cumulative utilities v^e_{gjt}. Having decided to move, household i solves the following optimization problem:

$$\max_{j\in\{0,\ldots,J\}} v^e_{g_ijt}-\phi_{g_i}+\epsilon_{ijt} \tag{9}$$

Following Assumption 1, the choice-specific probabilities are

$$\begin{aligned}P\left(j_{it}=j\,|\,j\notin\{J+1\}\,,j_{it-1}\right)&=\frac{\exp\left(v^e_{g_ijt}-\phi_g\right)}{\sum_{j'=0}^{J}\exp\left(v^e_{g_ij't}-\phi_g\right)}\\&=\frac{\exp\left(v^e_{g_ijt}\right)}{\sum_{j'=0}^{J}\exp\left(v^e_{g_ij't}\right)}\end{aligned} \tag{10}$$

Because moving costs are assumed to not vary by the neighborhood of origin or destination (Assumption 2), they cancel out. Following Berry (1994), we estimate $\hat{v}^e_{gjt}$ for $j\in\{0,\ldots J\}$ as

$$\hat{v}^e_{gjt}=\log\left(I_{gjt}\right)-\log\left(I_{g0t}\right). \tag{11}$$

Next, we consider the choice of whether or not to stay in the same home to identify the moving cost parameter ϕ_g. For household i who resided in j last period, the probability of choosing option $J+1$ (not moving) is

$$
\begin{aligned}
P\left(j_{it}=J+1 \mid j_{it-1}=j\right) &= P\left(v^e_{g_ijt}+\epsilon_{iJ+1t} > v^e_{g_ij't}-\phi_{g_i}+\epsilon_{ij't}\ ,\ \forall j' | \ j_{it-1}=j\right) \\
&= \frac{\exp\left(v^e_{g_ijt}\right)}{\sum_{j'=0}^{J}\exp\left(v^e_{g_ij't}-\phi_{g_i}\right)+\exp\left(v^e_{g_ijt}\right)}
\end{aligned} \tag{12}
$$

where the first line must hold for all $j'=0,\ldots,J$, and the second line follows from the logit formula (Assumption 1.2). The data analog to $P\left(j_{it}=J+1 \mid j_{it-1}=j\right)$ is simply $\frac{N_{g_ijt}-I_{g_ijt}}{N_{g_ijt-1}}$, or the proportion of group g_i households residing in neighborhood j in $t-1$ who decided to stay in the same home in the following period. Hence, equation (12) yields the J empirical moment conditions

$$
h_j\left(\phi_g;\boldsymbol{\hat{v}^e_{gt}}\right)=\frac{N_{gjt}-I_{gjt}}{N_{gjt-1}}-\frac{\exp\left(\hat{v}^e_{gjt}\right)}{\sum_{j'=0}^{J}\exp\left(\hat{v}^e_{gj't}-\phi_g\right)+\exp\left(\hat{v}^e_{gjt}\right)} \tag{13}
$$

By plugging in our estimates of $\hat{v}^e_{gjt}$ from equation (11) into equation (12), we can estimate ϕ_g by GMM using the moment conditions in (13). In Appendix C, we report results from Monte Carlo simulations based on the model from Section 3.1 that show that $\boldsymbol{\hat{\phi}}$ is a consistent estimator of $\boldsymbol{\phi}$.

Stage 2: Estimation of $\boldsymbol{\beta_g}$

We rewrite equation (6) based on the observed quantities as

$$
\hat{v}^e_{gjt}=\boldsymbol{\beta'_g s_{jt}}+\boldsymbol{\gamma'_g\hat{v}^e_{jt-T}}+\underbrace{\xi_{gjt}+\hat{v}^e_{gjt}-v^e_{gjt}-\boldsymbol{\gamma'_g\hat{v}^e_{jt-T}}}_{\text{error}_{gjt}}, \tag{14}
$$

We estimate $\boldsymbol{\beta_g}$ in the equation above via Two Stage Least Squares using $\boldsymbol{s_{jt-(T+1)}}$ as an IV for $\boldsymbol{s_{jt}}$, controlling for $\boldsymbol{\hat{v}^e_{jt-T}}$. Based on Assumption 1, $\hat{v}^e_{gjt}$ converges to v^e_{gjt}, and $\boldsymbol{\hat{v}^e_{jt-T}}$ converges to $\boldsymbol{v^e_{jt-T}}$. Moreover, based on Assumption 3, $\boldsymbol{s_{jt-(T+1)}}$ is uncorrelated to ξ_{gjt}, conditional on $\boldsymbol{v^e_{jt-T}}$. In Appendix C, we report results from Monte Carlo simulations based on the model from Section 3.1 that show that $\boldsymbol{\hat{\beta}}$ is a consistent estimator of $\boldsymbol{\beta}$.

3.3 Simulation of Segregation Dynamics

We simulate the dynamics of segregation by considering different counterfactual initial values of the elements of equations (4) and (5). Specifically, we simulate how the demographic composition of each neighborhood will evolve given a counterfactual initial state $\tilde{\boldsymbol{s}}^0 = (\tilde{\boldsymbol{s}}_{\boldsymbol{1}}^{\boldsymbol{0}}, ..., \tilde{\boldsymbol{s}}_{\boldsymbol{J}}^{\boldsymbol{0}})$, matrix of initial counterfactual values of amenities $\tilde{\boldsymbol{\xi}}^{\boldsymbol{0}}$, whose (g, j)th element is $\tilde{\xi}_{gj}^0$, vector of initial counterfactual distribution of households $\tilde{\boldsymbol{N}}^{\boldsymbol{0}}$, whose (g, j)th element is $\tilde{N}_{gj}^0$, vector of counterfactual response parameters $\tilde{\boldsymbol{\beta}}$, and vector of counterfactual moving costs $\tilde{\boldsymbol{\phi}}$. Each counterfactual is defined by a vector $\left(\tilde{\boldsymbol{s}}^{\boldsymbol{0}}, \tilde{\boldsymbol{\xi}}^{\boldsymbol{0}}, \tilde{\boldsymbol{N}}^{\boldsymbol{0}}, \tilde{\boldsymbol{\beta}}, \tilde{\boldsymbol{\phi}}\right)$ where tildes are used to denote that each of these are counterfactual objects that we can manipulate in order to analyze the effects of different forces on segregation.

Below we define the recursive function $\boldsymbol{s}\left(\cdot; \tilde{\boldsymbol{\xi}}^{\boldsymbol{0}}, \tilde{\boldsymbol{N}}^{\boldsymbol{0}}, \tilde{\boldsymbol{\beta}}, \tilde{\boldsymbol{\phi}}\right)$, such that $\boldsymbol{s}\left(\tilde{\boldsymbol{s}}^{\tau}; \tilde{\boldsymbol{\xi}}^{\boldsymbol{0}}, \tilde{\boldsymbol{N}}^{\boldsymbol{0}}, \tilde{\boldsymbol{\beta}}, \tilde{\boldsymbol{\phi}}\right) = \tilde{\boldsymbol{s}}^{\boldsymbol{\tau+1}}$ for each $\tau \geq 0$, which characterizes the evolution of segregation. To simplify the exposition, we write $\boldsymbol{s}(\tilde{\boldsymbol{s}}^{\tau})$ with the understanding that the vector $\left(\tilde{\boldsymbol{\xi}}^{\boldsymbol{0}}, \tilde{\boldsymbol{N}}^{\boldsymbol{0}}, \tilde{\boldsymbol{\beta}}, \tilde{\boldsymbol{\phi}}\right)$ is held constant. For $\tau \geq 0$, The $(g, j)^{\text{th}}$ element of the function $\boldsymbol{s}\left(\tilde{\boldsymbol{s}}^{\boldsymbol{\tau}}\right)$, $\tilde{s}_{gj}^{\tau+1}$, is calculated as

$$\tilde{s}_{gj}^{\tau+1}\left(\tilde{\boldsymbol{s}}^{\boldsymbol{\tau}}\right) = \frac{N_{gj}\left(\tilde{\boldsymbol{s}}^{\boldsymbol{\tau}}\right)}{\sum\limits_{g'} N_{g'j}\left(\tilde{\boldsymbol{s}}^{\boldsymbol{\tau}}\right)} \tag{15}$$

where

$$\begin{aligned} N_{gj}\left(\tilde{\boldsymbol{s}}^{\boldsymbol{\tau}}\right) = \quad & N_{gj}\left(\tilde{\boldsymbol{s}}^{\boldsymbol{\tau-1}}\right) \times \left(\frac{\exp\left(v_{gj}^e(\tilde{s}^\tau)\right)}{\sum_{j'=0}^J \exp\left(v_{gj'}^e(\tilde{s}^\tau) - \tilde{\phi}_g\right) + \exp\left(v_{gj}^e(\tilde{s}^\tau)\right)}\right) + \\ & \textstyle\sum_{k=1}^J N_{gk}\left(\tilde{\boldsymbol{s}}^{\boldsymbol{\tau-1}}\right) \times \left(\frac{\exp\left(v_{gj}^e(\tilde{s}^\tau) - \tilde{\phi}_g\right)}{\sum_{j'=0}^J \exp\left(v_{gj'}^e(\tilde{s}^\tau) - \tilde{\phi}_g\right) + \exp\left(v_{gk}^e(\tilde{s}^\tau)\right)}\right) \end{aligned} \tag{16}$$

with $v_{g0}^e\left(\tilde{\boldsymbol{s}}^{\boldsymbol{\tau}}\right)$ normalized to zero, and

$$v_{gj}^e\left(\tilde{\boldsymbol{s}}^{\boldsymbol{\tau}}\right) = \tilde{\boldsymbol{\beta}}_{\boldsymbol{g}}' \tilde{\boldsymbol{s}}_{\boldsymbol{j}}^{\boldsymbol{\tau}} + \tilde{\xi}_{gj}^0. \tag{17}$$

Equation (15), which is analogous to (8), simply defines the group shares. Equation (16) is based on the logit formula (e.g., see equation (12)). The first term on the right-hand side corresponds to the simulated number of households who resided in neighborhood j in the previous period and remained in their house , incurring no moving costs. The second term represents the simulated number of households who resided in neighborhood k in the previous period and then moved to neighborhood j next period (households with $k = j$ moved houses within neighborhood j). By implementing this simulation simultaneously for all neighborhoods, we incorporate all endogenous feedback that spills over from one neighborhood to another. Throughout the simulation, we keep the total number of households of each group across all neighborhoods in the Bay Area constant at the level implied by $\tilde{\boldsymbol{N}}^{\mathbf{0}}$ to ensures that our results do not reflect aggregate demographic changes in the SF Bay area, which are outside of our model.[18]

By repeatedly evaluating $\boldsymbol{s}(\cdot)$ starting from $\tilde{\boldsymbol{s}}^{\mathbf{0}}$, we can use equations (15), (16) and (17) to construct a *simulated trajectory* $\mathbb{T}(\tilde{\boldsymbol{s}}^{\mathbf{0}}) = \{\tilde{\boldsymbol{s}}^{\mathbf{0}}, \tilde{\boldsymbol{s}}^{1}, \tilde{\boldsymbol{s}}^{2}...\}$. We define a *steady state* as follows:

Definition 1. State $\boldsymbol{s}^{\star}$ is a *steady state* if $\mathbb{T}(\tilde{\boldsymbol{s}}^{\mathbf{0}})$ converges to $\boldsymbol{s}^{\star}$ for some $\tilde{\boldsymbol{s}}^{\mathbf{0}}$.[19]

The steady state corresponds to how neighborhoods would look in the long-run under a given counterfactual, so it can be useful when considering the long-run effects of counterfactual policies on segregation.[20] Moreover, identifying trajectories matters in practice if frictions imply slow convergence toward steady-state. We could in principle identify all steady states by conducting a grid search of all possible counterfactuals $\left(\tilde{\boldsymbol{s}}^{\mathbf{0}}, \tilde{\boldsymbol{\xi}}^{\mathbf{0}}, \tilde{\boldsymbol{N}}^{\mathbf{0}}, \tilde{\boldsymbol{\beta}}, \tilde{\boldsymbol{\phi}}\right)$ and computing simulated trajectories $\mathbb{T}(\tilde{\boldsymbol{s}}^{\mathbf{0}})$ for each of them, but the complexity of such an undertaking is formidable. Even just conducting a grid search on $\tilde{\boldsymbol{s}}$ and fixing all other initial conditions is computationally complex beyond the scope of this paper, because of the high dimensionality of $\tilde{\boldsymbol{s}}$. For tractability, we restrict our analysis to a specific set of informative counterfactuals.

[18]This is in practice achieved as follows. First, we ignore the outside option of leaving the Bay Area homeowners market by omitting $\exp\left(v^{e}_{g0} - \phi_j\right)$ from the denominators of the probabilities in equation (16). This ensures that these probabilities are conditional on remaining within the Bay Area. Second, we do not consider households who lived outside the Bay Area and moved in, as this decision is outside of our choice model.

[19]This notion of "steady state" in this paper has been sometimes referred to as "equilibrium" in the theoretical literature on the dynamics of segregation (Schelling (1969, 1971); Becker and Murphy (2000)). We view "steady state" as a more appropriate term because neighborhoods are understood to be always in Perfect Bayesian Equilibrium in our setup.

[20]One might be concerned that the linear specifications of equation (17) unduly constrains the simulated trajectories and the steady state that we obtain. In previous work (Caetano and Maheshri (2017)), we show that a simple linear specification allows for substantial flexibility in these trajectories even in a far more restrictive context (e.g., with only 2 groups, with $\tilde{s}_{j't}$ restricted to equal $s_{j't}$ for all $j' \neq j$, and with myopic households: $\mathbb{E}[\boldsymbol{s_t}|\boldsymbol{x_{gt}}] := \boldsymbol{s^e_t} = \boldsymbol{s_{t-1}}$).

Choices of $\left(\tilde{\boldsymbol{s}}^{\mathbf{0}}, \tilde{\boldsymbol{\xi}}^{\mathbf{0}}, \tilde{\boldsymbol{N}}^{\mathbf{0}}, \tilde{\boldsymbol{\beta}}, \tilde{\boldsymbol{\phi}}\right)$

As noted above, we must fix initial conditions and parameters before performing any simulation. We discuss these choices here. Let $\mathcal{T}$ correspond to November 2004, the final period of our sample.

1. Initial socioeconomic composition ($\tilde{\boldsymbol{s}}^{\mathbf{0}}$) - Our baseline choice for the initial socioeconomic composition is the observed $\boldsymbol{s}_{\mathcal{T}}$. We also consider a counterfactual in which households of each group are reallocated across neighborhoods so that the socioeconomic compositions of all neighborhoods are identical. This sheds light on the dynamic effects of a policy that fully integrated the Bay Area.

2. Initial amenities ($\tilde{\boldsymbol{\xi}}^{\mathbf{0}}$) - Our baseline choices for the initial unobserved expected value of amenities are their estimated values in the final period of our sample, i.e., $\tilde{\boldsymbol{\xi}}^{\mathbf{0}} = \hat{\boldsymbol{\xi}}_{\mathcal{T}}$, whose (g,j)th element is $\hat{\xi}^0_{gj\mathcal{T}} = \hat{v}^e_{gj\mathcal{T}} - \hat{\boldsymbol{\beta}}'_{\boldsymbol{g}} \boldsymbol{s}_{\boldsymbol{j}\mathcal{T}}$. This ensures that our simulation corresponds to one in which there are no exogenous shocks to amenities after November 2004. Still, the future unobserved expected value of amenities is allowed to endogenously respond to changes in socioeconomic compositions via $\boldsymbol{\beta}$, see the discussion below about the interpretation of $\boldsymbol{\beta}$.

3. Initial distribution of households ($\tilde{\boldsymbol{N}}^{\mathbf{0}}$) - Our baseline choice is the observed $\boldsymbol{N}_{\mathcal{T}}$.

4. Response parameters ($\tilde{\boldsymbol{\beta}}$) - Our baseline choice for the values of the response parameters is their estimated values $\hat{\boldsymbol{\beta}}$. We consider counterfactuals where specific elements of $\tilde{\boldsymbol{\beta}}$ are set to zero (corresponding to scenarios where households are "race-blind", "income-blind", or both).

5. Moving Costs ($\tilde{\boldsymbol{\phi}}$) - Our baseline choice for the moving cost parameters are their estimated values $\hat{\boldsymbol{\phi}}$. We also consider a counterfactual in which there is a one-time amnesty where $\tilde{\boldsymbol{\phi}} = \mathbf{0}$ for a single period, and one in which moving costs are permanently set to zero.

Remark 3. There is an important connection between assumptions on expectations and the endogenous feedback loop. Note that $\boldsymbol{s}\left(\boldsymbol{s}^e_{\boldsymbol{t}}\right) = \boldsymbol{s}_{\boldsymbol{t}}$ by construction, since observed choices of each group in t are made when $\tilde{\boldsymbol{s}} = \mathbb{E}[\boldsymbol{s}_{\boldsymbol{t}} | \boldsymbol{x}_{\boldsymbol{gt}}] = \boldsymbol{s}^e_{\boldsymbol{t}}$. Thus, if we assumed that households perfectly forecast the compositions (i.e., $\boldsymbol{s}^e_{\boldsymbol{t}} = \boldsymbol{s}_{\boldsymbol{t}}$ for all g), then that would imply $\boldsymbol{s}\left(\boldsymbol{s}_{\boldsymbol{t}}\right) = \boldsymbol{s}_{\boldsymbol{t}}$; that is, it would imply that data are always observed in steady state and that no feedback loop exists. More generally, the trajectory of convergence towards the steady state is likely affected by expectations, so it is crucial that we avoid strong assumptions on the formation of households' expectations if we want to study the dynamics of segregation.

3.4 Interpretation of Model Parameters

Interpretation of $\boldsymbol{\beta}$ The coefficient matrix $\boldsymbol{\beta}$ captures the various responses of households of different groups to their expectations of the socioeconomic compositions of neighborhoods. It does not reveal whether the response is mediated through a change in the flow utility or the continuation value associated with a neighborhood choice. To see this, consider equation (3), and define u^e_{gjt} (the flow utility) and CV^e_{gjt} (the continuation value of households if they choose neighborhood j in t) as the averages of $\mathbb{E}\left[u_{ijt}|\boldsymbol{x_{it}}\right]$ and $\int \delta \cdot \overline{V}\left(\boldsymbol{x_{it+1}}\right) f_{\boldsymbol{x}}\left(\boldsymbol{x_{it+1}}\left|j, \boldsymbol{x_{it}}\right.\right)$ across all households of group g respectively. Then we can write $v^e_{gjt} = u^e_{gjt} + CV^e_{gjt}$. For each g and g', $\beta_{g,g'} = \frac{\partial v^e_{gjt}}{\partial s_{g'jt}} = \frac{\partial u^e_{gjt}}{\partial s_{g'jt}} + \frac{\partial CV^e_{gjt}}{\partial s_{g'jt}}$, i.e., it represents the total marginal effect of an increase in the expected g' share on the group g valuation of that neighborhood, which incorporates the effect on both the flow utility and the continuation value.

We focus on this "reduced-form" effect because it allows us to study segregation dynamics without the additional assumptions required for this decomposition (see Remark 2). As Manski (2004) argues, choice data alone is insufficient to separately identify expectations and preferences. For instance, suppose a neighborhood is expected to increase its poor share, and we observe rich households responding to it by reducing their demand for that neighborhood. From choice data alone, we could not conclude that they responded to prejudice against poor households (a preference) as opposed to a signal that the neighborhood would become less desirable to them in the future for some other reason (an expectation), or both. While this would prevent us from identifying, say, households' willingness to pay to avoid residing close to poor neighbors, it would not restrict us from analyzing how households sort into or out of a neighborhood in response to an increase in the poor share since this is fundamentally related to households' choices and not their preferences *per se*. By imposing the assumptions that would allow us to identify households' marginal willingness to pay for their neighbors' characteristics without actually observing data on households' expectations, we would necessarily restrict how households' form their expectations. This is unwise in our setting, as this would in turn restrict the simulated trajectories that we identify (Remark 3).

$\boldsymbol{\beta}$ includes any type of discriminatory sorting on the basis of the socioeconomic composition of neighbors, including pure socioeconomic animus (or affinity) and statistical discrimination. It is useful to elaborate on what may constitute statistical discrimination in the context of neighborhood sorting. In the example above, suppose rich households inferred from the expected increase in poor share in t that the quality of the neighborhood school will decline in the future. A response to that expectation would qualify as statistical discrimination.[21] Thus, $\boldsymbol{\beta}$ includes not only sorting on the basis of changes in socioeconomic compositions *per se* ($\boldsymbol{s_{jt}}$), but also sorting on the basis of expected future endogenous changes in other amenities (e.g., ξ_{jt+1}) due to changes in socioeconomic compositions.

Moreover, $\boldsymbol{\beta}$ may also reflect supply-driven discrimination. For instance, suppose we found that Black households responded positively to an increase in the Black share. This would be possible even if Blacks exhibited no demand-driven discrimination, whether taste-based or statistical. Indeed, the same pattern could alternatively be explained by Black households simply facing obstacles to residing in neighborhoods without Blacks because of discrimination on the part of, say, the mortgage market (e.g., Ladd (1998)) or real estate agents (e.g., Ondrich, Ross and Yinger (2003)). Using the language of Christensen and Timmins (2019), in this example supply-driven discrimination would "steer" Blacks toward Black neighborhoods, which would lead us to find that Blacks respond positively to Black share even if there was no demand-driven discrimination on their part. If we had information about how the choice set of certain groups are more restricted because of such supply-driven discrimination, we could in principle separately identify such effect. Because we do not, we follow the standard approach of assuming all groups have the same choice set.

Thus, $\boldsymbol{\beta}$ reflects the overall *ability* of households to discriminate, i.e. to sort on the basis of the expected socioeconomic composition of the neighborhood for whatever reason. This ability is affected by both demand and supply considerations, and frictions (other than moving costs, which we explicitly model separately) may restrict or enhance this ability, so they show up in $\boldsymbol{\beta}$ as well. We are not aware of any empirical paper that separately identifies these different types of neighborhood discrimination.

[21] In a world of complete information, households would not use the neighbors' attributes to predict other amenities in the future, as they would be able to know their values directly. Incomplete information leads them to use such information. See Fang and Moro (2011) for a survey of models of statistical discrimination.

Interpretation of $\boldsymbol{\phi}$ The moving cost parameter, $\boldsymbol{\phi}$, refers to the *instantaneous* moving cost of changing homes. That is, although $\boldsymbol{\phi}$ enters separately from $\boldsymbol{\beta}$ in the model, this only applies to moves in period t. The expected costs of future moves (in response to unforeseen changes in neighborhood characteristics) are not included in $\boldsymbol{\phi}$, but they do impact CV_{gjt}; hence, by the argument above, they are loaded into $\boldsymbol{\beta}$. As a result, $\boldsymbol{\beta}$ may also contain components related to the interaction between the anticipated possibility of forecast errors and expected future moving costs. Indeed, households may recognize that they are unable to perfectly predict the future socioeconomic compositions of neighborhoods and any associated effects on other neighborhood amenities, and this may necessitate a future (costly) move.

4 Results

Our empirical analysis covers eight socioeconomic groups – all combinations of four races and two income groups – each of whom are allowed to respond heterogeneously to unobserved amenities as well as to five endogenous amenities – the shares of Blacks, Hispanics, and Asians (relative to Whites), the share of the poor (relative to the rich), and average neighborhood prices.[22]

4.1 Estimation Results

In Table 2, we present estimates of the responses to the socioeconomic compositions of neighborhoods ($\boldsymbol{\beta_g}$) along with the moving costs (ϕ_g) for households of each group. The endogenous amenities $\boldsymbol{s_{jt}}$ are instrumented by $\boldsymbol{s_{jt-13}}$ in equation (14).[23]

Since White (poor) share is the omitted race (income) amenity, the responses $\beta_{g,g'}$ are interpreted as the response of group g to a marginal increase in $s_{g'jt}$ relative to a marginal increase in the share of White (rich) neighbors. We find that households of each group respond positively to neighbors of the same race and rich households

[22]We lack sufficient data to precisely estimate $\boldsymbol{\beta}$ if we allowed each of the eight groups to respond to race and income in an unrestricted, non-separable way (i.e., 8x7 instead of 8x4 estimates of $\beta_{g,g'}$). We also lack sufficient data to precisely estimate $\boldsymbol{\pi}$ if we allowed each of the eight groups to respond to prices heterogeneously. Instead, we can only estimate π_g separately for rich households and for poor households (i.e., 2x1 instead of 8x1 estimates of π_g) With more data, a more flexible specification could be estimated.

[23]For robustness, we also estimated a specification where we included a cubic B-spline of $v_{g'jt-12}$ for all g' with four knots as controls in equation (14) to ensure that we were appropriately controlling for households' valuations of neighborhoods in period $t = T$. Our results were effectively unchanged.

respond positively to rich neighbors. Hispanics respond most positively to neighbors of their own race, followed by Asians and Blacks. Own race responses are stronger for poor households than rich households. Interestingly, not all responses are reciprocated: e.g., rich Hispanics respond negatively to Blacks, and poor Hispanics respond positively to Blacks, but Blacks of both income groups show little response to Hispanics. Altogether, these heterogeneous responses may give rise to complex dynamics.

Table 2: Responses to the Socioeconomic Compositions of Neighborhoods

	White		Black		Hispanic		Asian	
	Rich	Poor	Rich	Poor	Rich	Poor	Rich	Poor
Responses to:								
Black Share	-9.47	-7.41	9.41	10.71	-1.08	3.51	-3.67	-1.30
	(0.35)	(0.43)	(0.41)	(0.42)	(0.36)	(0.39)	(0.39)	(0.36)
Hispanic Share	-15.02	-5.93	-0.32	0.24	25.78	28.19	-1.13	4.94
	(0.55)	(0.69)	(0.51)	(0.52)	(0.59)	(0.63)	(0.50)	(0.58)
Asian Share	-4.50	-10.10	-0.87	-2.34	-1.84	-4.02	18.47	21.40
	(0.34)	(0.44)	(0.38)	(0.35)	(0.40)	(0.38)	(0.51)	(0.52)
Poor Share	-4.77	4.74	-5.11	2.01	-8.50	-0.20	-12.15	0.47
	(0.36)	(0.43)	(0.29)	(0.28)	(0.34)	(0.35)	(0.40)	(0.39)
Moving Costs	28.57	28.70	27.44	27.60	28.04	28.16	28.06	27.64
	(0.01)	(0.02)	(0.03)	(0.03)	(0.02)	(0.02)	(0.02)	(0.01)
R^2	0.79							
Num. of Observations	147,840							

Notes: The first eight columns of responses are 2SLS estimates of $\boldsymbol{\beta}_g$ from equation (14). We use $s_{g'jt-13}$ for all g' as instrumental variables. White is the omitted racial share and rich is the omitted income share. Moving costs are estimated by GMM (see equation 13). All standard errors clustered by group-month. The p-values for both the Cragg-Donald and the Kleinbergen-Paap weak identification tests are less than 0.001, which imply strong first stages.

Own race responses range from roughly one third of moving costs to the entirety of moving costs, while cross race responses tend to be smaller. [24] This allows for the possibility that substantial amenity mismatch may accumulate since many households may be locked into a neighborhood that is no longer their most preferred neighborhood. Although our estimates of moving costs are generally statistically different from each other, they are similar in magnitude across all socioeconomic groups (the maximum variation in these moving costs is less than 10% of the estimates).

In the Appendix, we present raw OLS estimates of $\boldsymbol{\beta}$ (Table 4). Our OLS estimates of $\boldsymbol{\beta}$ are much larger in magnitude than our IV estimates since there are many confounding reasons why similar households would choose similar neighborhoods (e.g., they tend to value other amenities more similarly), all of which would bias the OLS estimates upward in magnitude. The OLS bias is most pronounced for the within-group parameter estimates, as expected. We also report estimates of $\boldsymbol{\beta}$ for different values of T (the period corresponding to our IVs) in Figures 12 and 13. Larger values of T weaken Assumption 3 resulting in an IV that is more likely to be valid. We find that estimates of every element of $\boldsymbol{\beta}$ changes very little for $T = 13, \ldots, 36$.

4.2 Counterfactual Analysis

4.2.1 Baseline

We simulate the evolution of the socioeconomic compositions of neighborhoods setting $\tilde{\boldsymbol{\beta}}$ and $\tilde{\boldsymbol{\phi}}$ equal to their estimated values, and $\tilde{\boldsymbol{s}}$, and $\tilde{\boldsymbol{N}}$ at their observed November 2004 values. This set of initial conditions and parameters corresponds to our "Baseline" simulation, which can be interpreted as the simulated dynamic trajectory of segregation from November 2004 onward in the absence of future external shocks. In Figure 2, we present a graph of the number of neighborhoods that experience at least 1, 2, 5 or 10 simulated moves that change their socioeconomic composition. If, for instance, a rich White homeowner simply left a neighborhood, that would count as one change (one outflow). If instead they were replaced by another rich White homeowner, that

[24] As discussed in Kennan and Walker (2011), household-level moving costs in such discrete choice frameworks can be interpreted as also including ϵ (defined in Assumption 1), so they may vary substantially across households. This can explain why some households would move even with such large gaps between $\boldsymbol{\beta}$ and $\boldsymbol{\phi}$. Thus, moving costs *conditional on moving* are far less prohibitive than the moving cost estimates shown in Table 2. As a robustness check, we allowed for moving costs to vary by both group and year, but we found little heterogeneity over time.

would count as zero changes. If they were replaced by a homeowner of a different race or income level, that would count as two changes (one outflow plus one inflow). We describe neighborhoods experiencing such changes to their socioeconomic compositions as "in flux."

Initially, and for several decades to follow, nearly all neighborhoods are in flux. From this, we conclude that the Bay Area is not observed to be in steady state.[25] Despite substantial moving costs, the amenities of the neighborhoods where households are observed to reside are sufficiently unattractive to some households that most neighborhoods experience turnover. Over time, changes in the socioeconomic compositions of these neighborhoods feedback and also spill over to other neighborhoods, which in turn changes their relative attractiveness to homeowners of all socioeconomic groups. This process is slow and non-monotonic: although it takes 90-100 years for the Bay Area to approximate steady state[26] there are brief episodes of greater churn after 10 years and 33 years when gradual changes in a small number of neighborhoods spill over sufficiently such that many neighborhoods are in flux all at once.

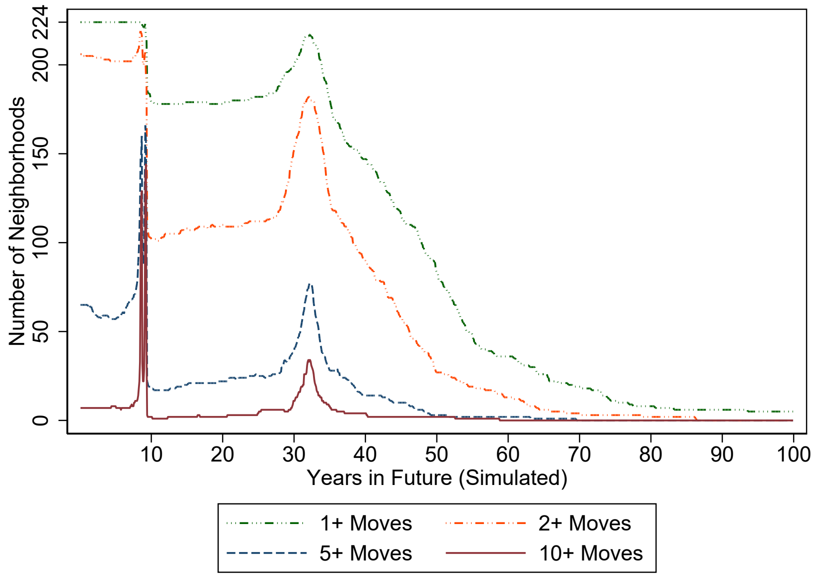
The outcome of this pattern of sorting is a change in both the levels of segregation and average neighborhood prices in the Bay Area. We present these results in Figure 3. The top panel corresponds to our baseline simulation, and the remaining nine panels correspond to different counterfactuals that we discuss in more detail in the remainder of this section .In each panel, we present simulated changes in the dissimilarity index for each race (pooling income groups) and for each income group (pooling races) across all Bay Area neighborhoods. In every simulation, the change for each group is presented as an arrow with two arrowheads. The first arrowhead corresponds to the medium-run (10 years into the future) change in the dissimilarity index, and the second arrowhead corresponds to the long-run (154 years into the future, after steady state is reached).

As shown in the top panel labeled "Baseline", all races experience modest increases in segregation, though this can be a very slow process. White households experience the smallest increase in segregation in both absolute and relative terms, though it occurs fairly quickly. In the medium-run, White households are 21% more segregated, but

[25]This also implies that neighborhoods are not observed at "tipping points" since they correspond to an unstable steady state. Hence, small deviations in our simulation due to, say, estimation error, should leave our long-run conclusions effectively unchanged, which we confirmed empirically.

[26]It takes 154 years for all Bay Area neighborhoods to experience no moves (see Appendix Figure 14).

Figure 2: Number of Neighborhoods In Flux (Simulated)

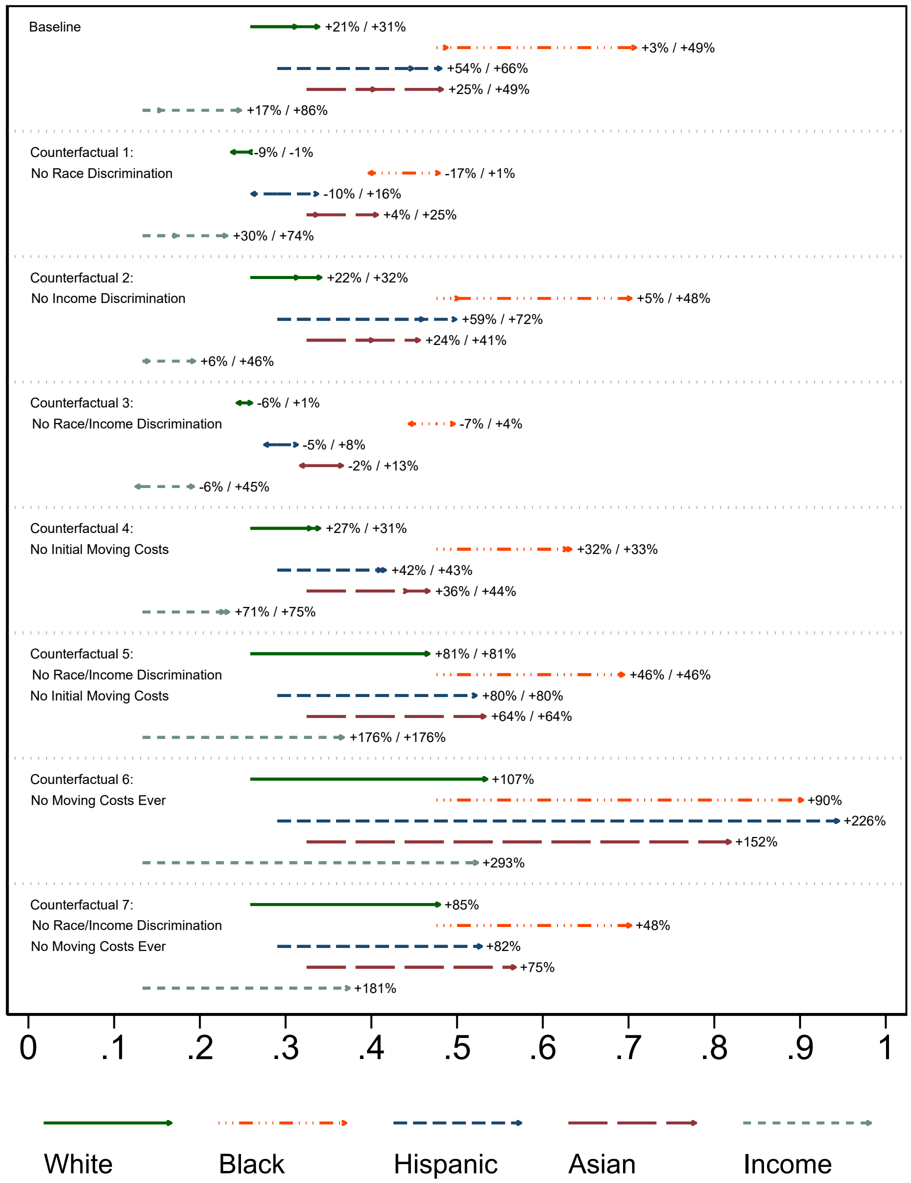


Notes: Figure shows the number of neighborhoods with at least one, two, five or ten moves that change their socioeconomic composition (out of a total of 224 neighborhoods). Simulation begins in November 2004.

in the long run they are only 31% more segregated. Black households start off more segregated than all other races and remain so throughout the simulation. However, Black segregation increases by only 3% in the medium run, though it increases by 49% in the long run.[27] Hispanic homeowners experience the largest and most rapid absolute and relative (54% in the medium-run, 66% in the long-run) increases in segregation, followed by Asians, who experience a 25% increase in segregation in the medium-run, and a 49% increase in the long-run. Although income segregation increases substantially in the long-run, this process unfolds rather slowly (17% in the medium-run, 86% in the long-run).

[27]We speculate that Black segregation increases very little in the medium-run because moving costs are a more strongly binding constraint to Black households. Indeed, we find that the ratio of mismatch (defined as the difference between the utility households receive from their observed neighborhoods with the utility they receive from their ideal neighborhood) to moving costs is highest for Black households over the sample period. This implies that Black households are least able to mitigate mismatch in practice.

Figure 3: Medium- and Long-Run Changes in Race and Income Segregation (Simulated)



Notes: The arrows represent the changes in simulated Dissimilarity Indices for households of each race and income from November 2004 in the absence of exogenous shocks. The first arrowhead corresponds to a "medium-run" that is 10 years into the future, and the second arrow corresponds to a "long-run" that is over 154 years into the future after neighborhoods have reached steady state. Numbers correspond to the relative change in dissimilarity. A Black dissimilarity index of, say, 0.60, means that 60% of Black homeowners would have to be relocated in order to generate an equal distribution of Blacks across all Bay Area neighborhoods. Details of each counterfactual are presented in Section 4.2.

4.2.2 The Roles of Discriminatory Responses: Race and Income

Our estimates of $\hat{\boldsymbol{\beta}}$ reveal systematic discriminatory responses of homeowners of all socioeconomic groups. To isolate their roles in explaining the patterns of segregation dynamics presented in Figures 2-3, we consider a series of counterfactuals in which households are either "race-blind", i.e., unresponsive to the racial composition of their neighbors ($\tilde{\beta}_{g,g'} = 0$ for all g and $g' \in \{\text{Black}, \text{Hispanic}, \text{Asian}\}$), "income-blind", i.e., unresponsive to the income composition of their neighbors ($\tilde{\beta}_{g,\text{poor}} = 0$ for all g), or both race- and income-blind ($\tilde{\boldsymbol{\beta}} = \mathbf{0}$). As shown in Figure 4, discriminatory racial responses are responsible for the non-monotonic process of convergence; in their absence, there are no episodes of intense churn.

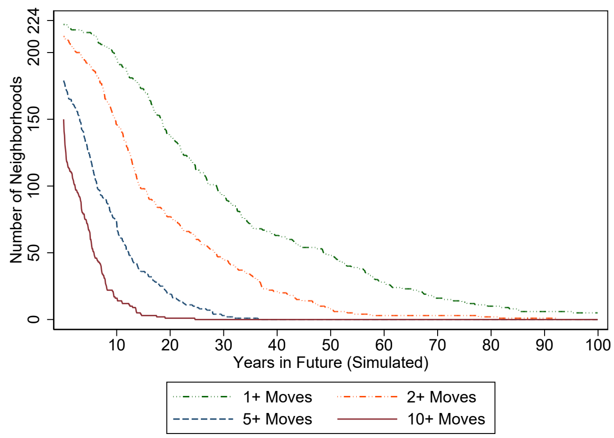
We present the simulated increase in segregation under each of these counterfactuals in Figure 3 (panels labeled Counterfactual 1-3). By comparison with the baseline results it is clear that removing racial discrimination has a profound impact on segregation in both the medium and long run (Counterfactual 1). Indeed, segregation is expected to decrease slightly in the medium run for all races except for Asians in the absence of racial discrimination, and any long run increases are modest at best. Our findings of decreases in segregation in the medium run is evidence that households and neighborhood amenities are mismatched – what is keeping households of a given race in a particular neighborhood is the presence of same-race neighbors; once they no longer value the race of their neighbors, they are much more likely to move to a different neighborhood with a very different racial composition. On the other hand, income segregation has little effect on segregation, either in the presence of racial discrimination (Counterfactual 2 vs. baseline) or in the absence of discrimination (Counterfactual 3 vs. Counterfactual 1).[28]

Remark 4. In Counterfactual 3, we find a gradual dynamic adjustment even in the absence of discrimination (when there is no scope for Schelling-style segregation dynamics). This arises solely due to moving costs. When moving is costly, stayers and

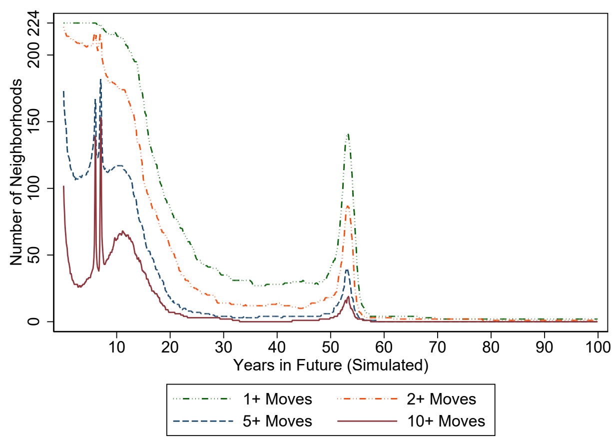
[28]We refrain from considering counterfactual changes in $\boldsymbol{\xi}$ because we are unable to estimate the causal relationship between a specific amenity included in $\boldsymbol{\xi}$ and residential decisions, which would be an actual policy parameter. (In Section 5, we discuss why estimating such a relationship, even if feasible, might be inadvisable depending on the objectives of the researcher.) In principle, large variation in $\boldsymbol{\xi}$ might lead to a completely different trajectory due to the possibility of multiple steady states. However, small variation in $\boldsymbol{\xi}$ should affect our long-run conclusions only to the extent that it perturbed the location of the steady state by slightly altering the trajectory since we found that the data were not observed near a tipping point (Footnote 17). We verified this empirically.

Figure 4: Number of Neighborhoods In Flux - No Discrimination (Simulated)

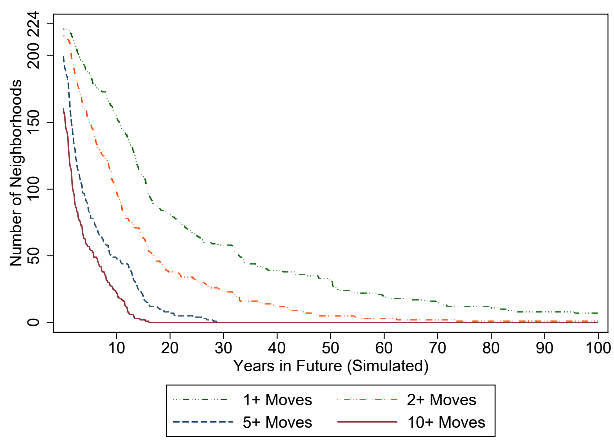
(a) No Racial Discrimination



(b) No Income Discrimination



(c) No Racial or Income Discrimination



Notes: Each panel shows the number of neighborhoods with at least one, two, five or ten moves that change their socioeconomic composition (out of a total of 224 neighborhoods) under a different counterfactual. Simulation begins in November 2004.

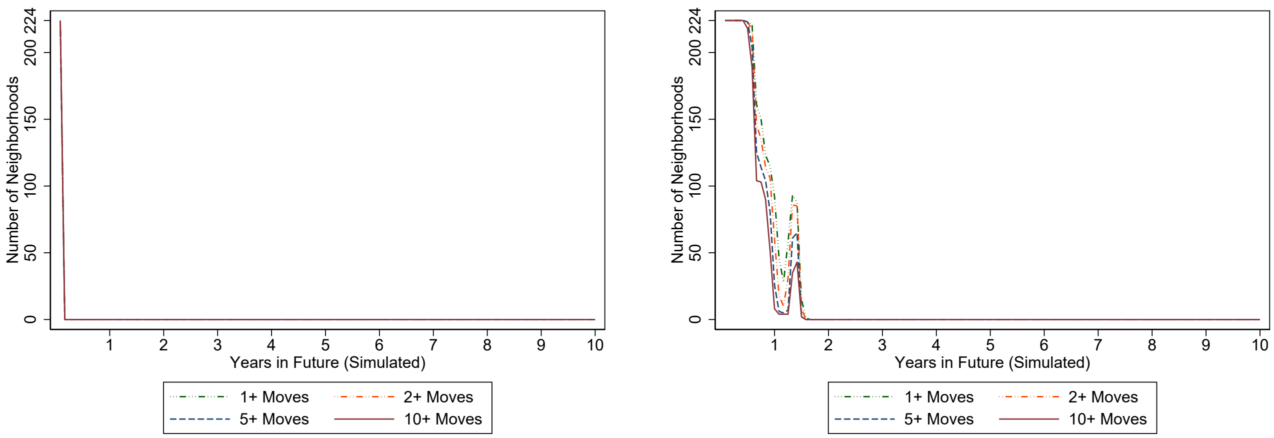
movers do not face symmetric decisions, as the former group does not bear moving costs while the latter group does. The probabilities that each of these types of households choose a particular neighborhood remains fixed in the absence of discrimination, however their proportions in a given neighborhood change because some households are mismatched at the beginning of the simulation. This leads to the dynamic adjustment shown in panel 3 of Figure (4) and the medium- and long-run effects on segregation shown in Figure (3). Note that these effects (shown in Counterfactual 3) are quite small relative to the effects of discrimination on segregation. When moving is not costly, stayers and movers face symmetric decisions, and dynamic adjustment is immediate (Counterfactual 7).

4.3 The Role of Moving Costs

The gradual declines of Figures 2 and 4 suggest that moving costs play an important role in the dynamics of segregation. To explore this further, we consider a counterfactual in which all homeowners enjoy a one-time moving-cost amnesty at the beginning of the simulation ($\tilde{\boldsymbol{\phi}}$ is set to zero for the first iteration of the simulation and then set to $\hat{\boldsymbol{\phi}}$ thereafter). As shown in the first panel of Figure 5, the Bay Area converges to a steady state instantaneously. In the first period, the lack of moving costs allows households to eliminate their mismatch (per their *ex ante* expectation in t). However, this does not imply that there is no mismatch in t or in further periods, as forecast errors may lead households to reside in neighborhoods that turn out to be suboptimal. Nevertheless, this mismatch is quite small relative to moving costs, which are restored in future periods.

Figure 5: Number of Neighborhoods In Flux - No Moving Costs (Simulated)

(a) No Initial Moving Costs (b) No Moving Costs Ever



Notes: Each panel shows the number of neighborhoods with at least one, two, five or ten moves that change their socioeconomic composition (out of a total of 224 neighborhoods) under a different counterfactual. Simulation begins in November 2004.

In order to illustrate this point, consider a different counterfactual where we eliminate moving costs both today and in every future period by permanently setting $\tilde{\boldsymbol{\phi}} = \mathbf{0}$. We do not consider this counterfactual to be particularly sensible because households would have to be repeatedly surprised by the future elimination of moving costs every time it occurs.[29] However, this exercise is valuable as it allows us to gauge the role of incomplete information, which is hidden in our context but would perhaps play a more important role in a context with much lower moving costs or much higher discrimination than the one we encounter in our sample. The dynamics of this second counterfactual are shown in the second panel of Figure 5. While convergence is still much faster than in the baseline case with moving costs (note that the x-axis corresponds only to a 10 year period as opposed to a 100 year period), it is not instantaneous, owing to the fact that forecast errors would still trigger further moves (which would remain costless in this scenario), leading to the feedback loop discussed in Schelling (1969).

[29]This "surprise" must occur because $\boldsymbol{\beta}$ also contains the marginal effects of $\boldsymbol{s_{jt}}$ on the continuation values and was estimated using data generated in a world with expectations of non-zero future moving costs (see Section 3.4). Note that we would encounter a related issue if we instead decided to separately identify the flow utility and the continuation value components of $\boldsymbol{\beta}$. Doing so would require us to assume households are forward looking and anticipate future moving costs in a very specific way, but if this assumption was invalid, it would lead to misestimation of the simulated trajectories. For instance, in practice households may discount the future differently depending on their socioeconomic group.

We explore the interaction between moving costs and segregation in counterfactuals 4-7 of Figure 3. Counterfactual 4 (which maintains baseline discriminatory responses) leads to higher segregation levels across the board relative to baseline. When we shut off discriminatory responses in counterfactual 5, the resulting increases in segregation are much less pronounced.

A comparison of counterfactuals 4 and 5 reveals that discrimination would play a quantitatively similar role in a world without (initial) moving costs, with perhaps the exception of income segregation, which would occur much more rapidly without moving costs (though it would still converge to a similar level in the long run). On the contrary, in counterfactuals 6 and 7 when moving costs are permanently eliminated, segregation would increase dramatically in counterfactual 6 though much less under in the absence of discriminatory responses (counterfactual 7), which highlights the role of the endogenous feedback loop. This suggests that any permanent reductions in moving costs could cause rapidly increasing segregation. Thus, we conclude that frictions, especially moving costs, mitigate the role of discrimination on segregation.

4.4 The Role of the Initial Allocation of Households

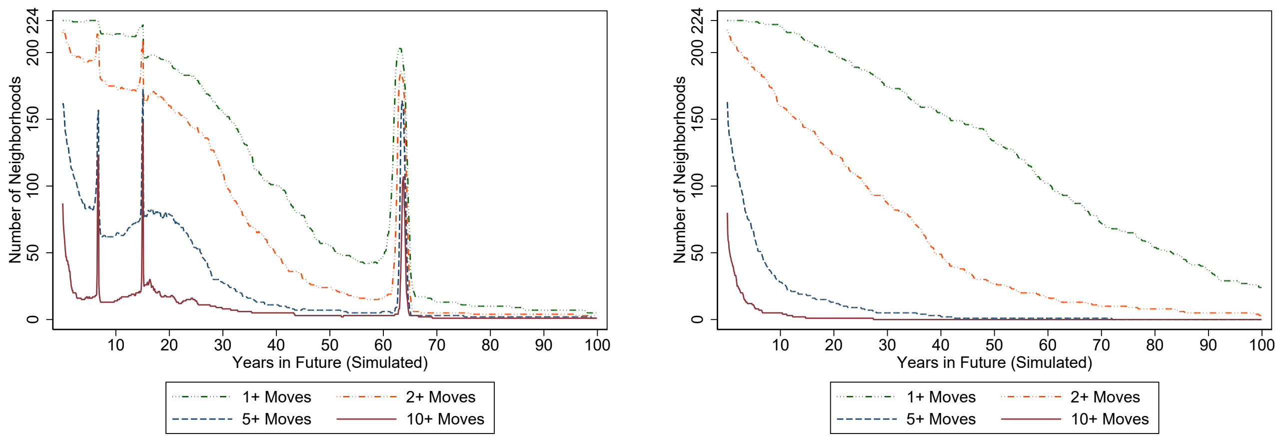
We now consider a hypothetical policy in which households are re-allocated so that all neighborhoods have the exact same initial socioeconomic compositions (but, importantly, other neighborhood amenities are unchanged). The first panel of Figure 6 plots the number of neighborhoods in flux after the full integration policy. As compared with the benchmark in Figure 2, this re-arrangement of households takes only slightly longer to reach steady state, thought there are additional episodes of churn farther out into the future. This could reflect the fact that such a policy leads to misalignment that takes longer to undo because of moving costs. Eliminating discrimination, as in the second panel of Figure 6, ensures convergence is monotonic as before, thought it also seems to slow down convergence as expected.

We explore the relationship between initial socioeconomic compositions and segregation dynamics in Figure 7 under six counterfactuals. When starting in a fully integrated Bay Area, the segregation level is zero by construction. Counterfactual 0' shows that this integration policy would reduce segregation in the medium-run relative to the baseline counterfactual, but it would not lead to large relative reductions in

Figure 6: Number of Neighborhoods In Flux - Full Integration (Simulated)

(a) Baseline Responses and Moving Costs

(b) No Race or Income Discrimination



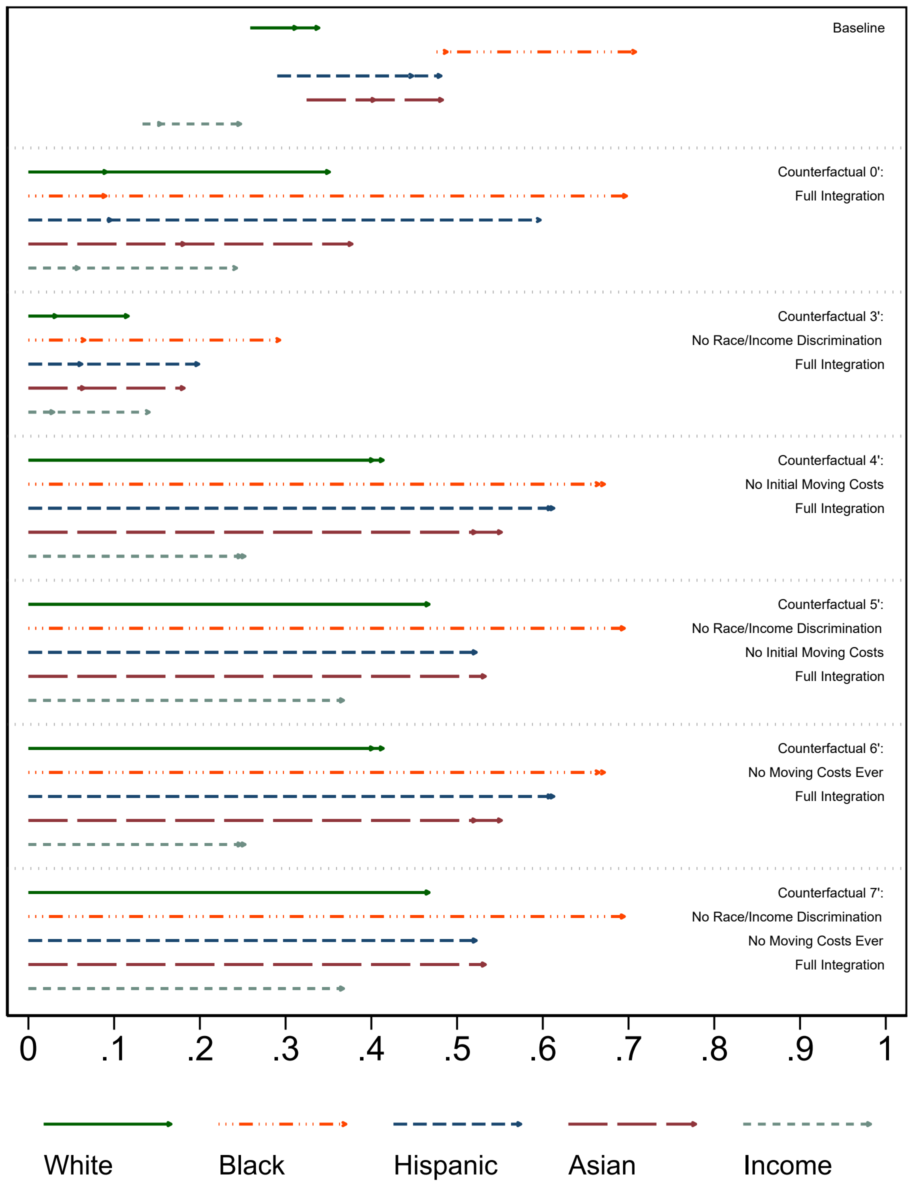
Notes: Each panel shows the number of neighborhoods with at least one, two, five or ten moves that change their socioeconomic composition (out of a total of 224 neighborhoods) under a different counterfactual. Simulation begins in November 2004.

segregation in the long-run (and would even increase Hispanic segregation). However, when we eliminate discrimination in counterfactual 3', segregation of all types would reduce dramatically. This suggests that discrimination offsets the desegregating benefits of policies that reallocate households across neighborhoods.[30]

In counterfactuals 4'-7', moving costs are eliminated either in a one time amnesty or permanently. This rapidly intensifies the levels of segregation in the medium-run relative to counterfactual 0', but changes the long-run levels of segregation very little. The stark contrast with our results from Section (5) that did not start from a fully integrated state is evidence of multiple steady states, since initial conditions seem to matter in the long-run.

[30]This is consistent with the empirical literature on demographic change associated with school segregation, which Morrill (1989) summarizes as "Over the longer term, the effect of mandatory busing is to ... [foster] polarization between a poorer, minority-dominated central city and richer, white-dominated suburbs."

Figure 7: Steady State Changes in Race and Income Segregation - Full Integration (Simulated)



Notes: The arrows represent the changes in simulated Dissimilarity Indices for households of each race and income from November 2004 in the absence of exogenous shocks. The first arrowhead corresponds to a "medium-run" that is 10 years into the future, and the second arrow corresponds to a "long-run" that is over 166 years into the future after neighborhoods have reached steady state. Numbers correspond to the relative change in dissimilarity. A Black dissimilarity index of, say, 0.60, means that 60% of Black homeowners would have to be relocated in order to generate an equal distribution of Blacks across all Bay Area neighborhoods. Details of each counterfactual are presented in Section 4.4.

5 Incorporating Additional Information on Prices and Amenities

In this section, we extend the framework to incorporate data on neighborhood prices. This allows us to conduct two additional exercises. First, we closely replicate our previous results, which supports our claim that data on choices alone is sufficient to analyze important aspects of segregation. Second, we present the results of additional counterfactual simulations that reveal the role of prices in explaining segregation dynamics and the endogenous effects of segregation on home price appreciation. By incorporating price into the model, the number of parameters that need to be identified – and hence the number of identifying assumptions – increases. Thus, a parsimonious model is preferred unless the question of interest involves the role of prices on segregation or vice versa.

5.1 Empirical Framework with Price Data

5.1.1 Set Up

In this section, we consider the unique role of prices as determinants of segregation. We modify our specification of expected utility in equation (5) as

$$v_{gjt}^{e} = \boldsymbol{\theta_g'} \boldsymbol{s_{jt}^{e}} + \pi_g P_{jt}^{e} + \xi_{gjt}^{e} \tag{18}$$

where $P_{jt}^{e} = \mathbb{E}\left[P_{jt}|\boldsymbol{x_{gt}}\right]$ is the expected neighborhood price, and ξ_{gjt}^{e} now encompasses the expected value of the neighborhood due to all other neighborhood amenities besides socioeconomic composition and price. The parameter vector $\boldsymbol{\theta_g}$ corresponds to group g responses to the expected socioeconomic composition of the neighborhood *net of price effects*, in contrast with $\boldsymbol{\beta_g}$, which corresponded to those responses inclusive of price effects.

Because the socioeconomic composition might be expected to be capitalized in prices, we need to include an additional equation:[31]

$$P_{jt}^{e} = \alpha + \boldsymbol{\rho}' \boldsymbol{s_{jt}^{e}} + \eta_{jt}^{e}. \tag{19}$$

[31] Technically, it is possible that the effect of $\boldsymbol{s_{jt}^{e}}$ on P_{jt}^{e} changes also with g, since P_{jt}^{e} is the expectation of group g with respect to P_{jt}. Because we do not have data on price expectations, we cannot explicitly consider this possibility. If $\boldsymbol{\rho_g}$ represents this potentially heterogeneous effect, then η_{jt}^{e} in equation (19) includes the term $(\boldsymbol{\rho_g} - \boldsymbol{\rho})' \boldsymbol{s_{jt}^{e}}$.

Figure 8: Causal Graph of Socioeconomic Composition, Price and Valuation

$$\begin{tikzcd} \boldsymbol{s^{e}_{g'jt}} \arrow[d, "\rho_{g'}"] \arrow[dr, "\theta_{g,g'}"] & \\ P^{e}_{jt} \arrow[r, "\alpha_g"] & v^{e}_{gjt} \end{tikzcd}$$

The following figure helps us connect these two equations, and how they relate to equation (5). We represent causal relationships as directed arrows, and for simplicity we focus on the causal effects of a single element of the vector $\boldsymbol{s^{e}_{jt}}$, the expected share of group g' $s^{e}_{g'jt}$.

Group g households' valuation of neighborhood j are affected by their expectations of the group g' share in that neighborhood both directly (encapsulated by $\theta_{g,g'}$) and indirectly through the expected price response (encapsulated by $\rho_{g'}$ and α_g). Our baseline model without price incorporates this entire causal effect into the single parameter $\beta_{g,g'}$. As such, we can represent the mathematical relationship between the parameters of the model with price data and the parameters of the baseline model as

$$\beta_{g,g'} = \theta_{g,g'} + \rho_{g'}\alpha_g \tag{20}$$

For intuition, it is useful to explicitly describe a hypothetical sequence of events that might occur within a single period t. Consider a change in $s^{e}_{g'jt}$ at the very beginning of period t. As households of group g change their expectation regarding the share of group g', they will anticipate their value for the neighborhood (v^{e}_{gjt}) will change in two ways: first, directly depending on whether they value $s_{g'jt}$ per se, which is reflected in $\theta_{g,g'}$. Second, indirectly via prices. Indeed, households will anticipate changes in the neighborhood price (via $\rho_{g'}$), which in turn should affect v^{e}_{gjt} via α_g. For each g, the coefficient $\beta_{g,g'}$ incorporates the effect of $\boldsymbol{s^{e}_{g'jt}}$ on v^{e}_{gjt} via these two channels. Note that we make no assumptions on the particulars of the process by which $s^{e}_{g'jt}$ causes v^{e}_{gjt}. We simply posit that however households expect this process to be, it can be decomposed into one price channel, $\rho_{g'} \cdot \alpha_g$ and a second residual channel $\theta_{g,g'}$. In particular, the mechanism that leads changes in their expected share of group g' ($s^{e}_{g'jt}$) to affect changes in prices (P_{jt}) is not specified; it may include expected adjustments in the supply-side or the demand-side, or both, and it may or may not include the expectation that the market will clear (i.e., actual demand equals actual supply).[32]

In our main analysis, we simply identified $\boldsymbol{\beta_g}$. This allowed us to characterize segregation without imposing additional assumptions required to disentangle the direct and indirect effects of $\boldsymbol{s_{jt}}$ on v^e_{gjt}. We intentionally selected the set of endogenous amenities parsimoniously by focusing on the two *primitive* endogenous dimensions along which households sort: race and income. However, that specification did not allow for us to study the potential role of prices in separately explaining segregation. To do so, we must identify $\theta_{g,g'}$, α_g and $\rho_{g'}$ for all g and g', which requires additional assumptions.

We re-write equation (18) from the perspective of the econometrician as

$$v^e_{gjt} = \boldsymbol{\theta'_g s_{jt}} + \pi_g P_{jt} + \xi_{gjt} \tag{21}$$

where $\xi_{gjt} = \xi^e_{gjt} + \boldsymbol{\theta'_g} \left(\boldsymbol{s^e_{jt}} - \boldsymbol{s_{jt}}\right) + \pi_g \left(P^e_{jt} - P_{jt}\right)$. Analogously, we re-write equation (19) from the perspective of the econometrician as

$$P_{jt} = \alpha + \boldsymbol{\rho' s_{jt}} + \eta_{jt}, \tag{22}$$

where $\eta_{jt} = \eta^e_{jt} + \boldsymbol{\rho'} \left(\boldsymbol{s^e_{jt}} - \boldsymbol{s_{jt}}\right)$.

5.1.2 Identification

Instead of Assumption 3, we make the following assumption to identify both $\boldsymbol{\theta_g}$ and π_g for all g.

Assumption 3′. *IV Validity (ξ). For all g and t, there exists some $T > 0$ such that*

$$\mathbb{E}\left[\xi_{gjt} \,\middle|\, \boldsymbol{v^e_{jt-T}}, \boldsymbol{s_{jt-(T+1)}}, P_{jt-(T+1)}\right] = \mathbb{E}\left[\xi_{gjt} \,\middle|\, \boldsymbol{v^e_{jt-T}}\right] \tag{23}$$

Assumption **3′** replaces Assumption 3 and states that the components of $s_{jt-(T+1)}$ and $P_{jt-(T+1)}$ that are orthogonal to $\boldsymbol{v^e_{jt-T}}$ are valid IVs for $\boldsymbol{s_{jt}}$ and P_{jt}. The argument for the plausibility of this validity assumption is analogous to the argument discussed in Section 3: information from the distant past ($\tau < t - T$) that turned out to be

[32]This hypothetical sequence of events is related to the sequence described by Bayer, McMillan and Rueben (2004*a*) about how steady state is assumed to be achieved in each period in their simulations. The key departure here is that we do not assume that these expectations are connected to what is observed in the data, so data may not be in steady state at any given period t.

irrelevant for the sorting decisions of the near past $(t-T)$ should not be relevant for sorting decisions in t. The argument for the relevance assumption of these IV is also analogous. In particular, note that P_{jt} is in part determined by households' willingness to *sell* their homes, which in turn depends on how mismatched to their neighborhood these households currently are. Since some of these mismatched households made their original sorting decision before $t-T$, P_{jt} should be correlated to information from the far past that turned out to be irrelevant more recently. For instance, the "unrealized" component of the price that a household purchased their home in $\tau < t-T$ (i.e., the component due to $\boldsymbol{\omega_{gj\tau}} = \boldsymbol{x_{g\tau}} - \mathbb{E}\left[\boldsymbol{x_{g\tau}} \middle| \boldsymbol{v^e_{jt-T}}\right]$, see Section 3) may still influence the household's willingness to sell in t.

In order to identify $\boldsymbol{\rho}$ in equation (22), we make the following assumption, which allows us to use the component of $\boldsymbol{s_{jt-(T+1)}}$ that is orthogonal to $\boldsymbol{v^e_{jt-T}}$ as an IV for $\boldsymbol{s_{jt}}$:

Assumption 6′. *IV Validity (η). For all t, there exists some $T>0$ such that*

$$\mathbb{E}\left[\eta_{jt} \middle| \boldsymbol{v^e_{jt-T}}, \boldsymbol{s_{jt-(T+1)}}\right] = \mathbb{E}\left[\eta_{jt} \middle| \boldsymbol{v^e_{jt-T}}\right]$$

Assumption **6′** is closely related to Assumption **3′**. Note that η_{jt} incorporates the price capitalization of "demand-side" unobservables (ξ_{gjt} in equation (21)) as well as supply-side ones (not explicitly modeled in this paper, but potentially included in η_{jt}). Hence, Assumption **6′** states that the price capitalization of other amenities in t does not rely on the information from the distant past ($\tau < t-T$) that was ignored by all households who sorted in the more recent past $(t-T)$. The relevance of $\boldsymbol{s_{jt-(T+1)}}$ as an IV for $\boldsymbol{s_{jt}}$ conditional on $\boldsymbol{v^e_{jt-T}}$ follows analogously as well, i.e. due to mismatched households.

Estimation of the model and simulation are analogous to the case without price data; full details are presented in Appendix B.

5.2 Results with Price Data

We present estimates of $\boldsymbol{\theta}$, $\boldsymbol{\pi}$, $\boldsymbol{\rho}$ and $\boldsymbol{\phi}$ in Table 3. The signs and relative magnitudes of our estimates of $\boldsymbol{\theta}$ are roughly similar to the estimates of $\boldsymbol{\beta}$ from our baseline model with the exception that rich households of all races respond less intensely to poor neighbors than before. This suggests that much of the response to poor households that was identified in our main analysis reflected differential responses to neighborhood prices as opposed to the income of neighbors. All households respond negatively to higher neighborhood prices, but poor White and Asian households are over eight and three times more price-sensitive than their rich counterparts respectively. Our estimates of $\boldsymbol{\rho}$ presented in the far right column imply that a 10 percentage point increase in the expected Black share of a neighborhood, all else constant, leads to a reduction in average price of \$12,300. This effect is over four times as large for the same increase in the expected Hispanic or poor shares of a neighborhood and roughly the same size for the same increase in the expected Asian share of a neighborhood. Altogether, we cannot reject that $\hat{\beta}_{g,g'} = \hat{\theta}_{g,g'} + \hat{\boldsymbol{\rho}}'_{\boldsymbol{g}} \hat{\boldsymbol{\alpha}}_{\boldsymbol{g}}$ for all g, g', which suggests that our estimates of $\hat{\boldsymbol{\beta}}$ were not biased by confounders correlated to price.

Table 3: Responses to the Socioeconomic Compositions and Prices of Neighborhoods

	White		Black		Hispanic		Asian		
	Rich	Poor	Rich	Poor	Rich	Poor	Rich	Poor	Prices
Responses to:									
Black Share	-9.82	-9.17	8.89	10.17	-1.84	2.84	-3.67	-2.27	-0.123
	(0.38)	(0.47)	(0.41)	(0.40)	(0.35)	(0.35)	(0.36)	(0.36)	(0.010)
Hispanic Share	-16.08	-13.78	-1.97	-1.85	22.05	25.05	2.13	0.67	-0.563
	(0.55)	(0.72)	(0.53)	(0.52)	(0.61)	(0.61)	(0.64)	(0.64)	(0.012)
Asian Share	-4.49	-11.66	-1.34	-2.80	-2.42	-4.60	18.24	20.64	-0.094
	(0.34)	(0.42)	(0.36)	(0.33)	(0.44)	(0.44)	(0.55)	(0.55)	(0.013)
Poor Share	-5.81	-3.44	-7.24	0.55	-12.69	-3.02	-13.45	-4.49	-0.584
	(0.47)	(0.50)	(0.36)	(0.32)	(0.39)	(0.39)	(0.45)	(0.45)	(0.013)
Response to Price (Thousands)	-1.36	-12.72	-3.45	-1.39	-6.45	-5.03	-2.04	-7.12	
	(0.38)	(0.90)	(0.64)	(0.46)	(0.07)	(0.08)	(0.70)	(0.74)	
Moving Costs	28.57	28.70	27.44	27.60	28.04	28.16	28.06	27.64	
	(0.01)	(0.02)	(0.03)	(0.03)	(0.02)	(0.02)	(0.02)	(0.01)	
R^2					0.79				0.43
Num. of Observations					147,840				36,960

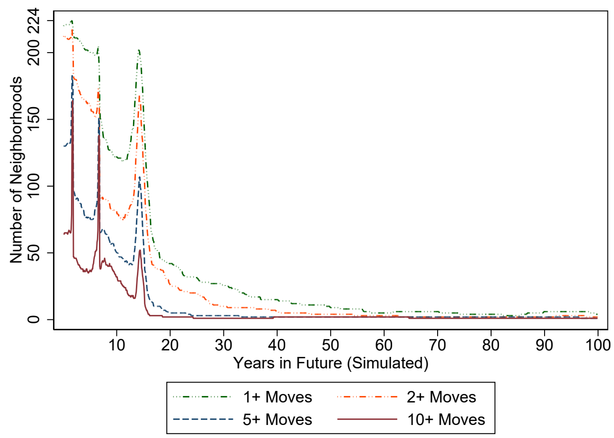
Notes: π_g is allowed to vary only by income group. We use $s_{g'jt-13}$ for all g' and P_{jt-13} as instrumental variables. White is the omitted racial share and rich is the omitted income share. Price as a dependent variable is denominated in millions of dollars. Moving costs are estimated by GMM (see equation 13). All standard errors clustered by group-month. The p-values for both the Cragg-Donald and the Kleinbergen-Paap weak identification tests for both regressions are less than 0.001, which imply strong first stages.

To isolate the role of endogenous price responses on segregation dynamics presented in Figures 2-3, we consider two counterfactuals. In each, households all have symmetric price responses ($\tilde{\pi}_g$ is equal to a population weighted sum of $\hat{\pi}_{\text{rich}}$ and $\hat{\pi}_{\text{poor}}$ for all groups g). In the second counterfactual, households are additionally unresponsive to the socioeconomic compositions of neighborhoods ($\tilde{\boldsymbol{\theta}} = \mathbf{0}$). A comparison of Figure 9 with Figure 2 and the third panel of Figure 4 shows that price responses have little effect on segregation dynamics.

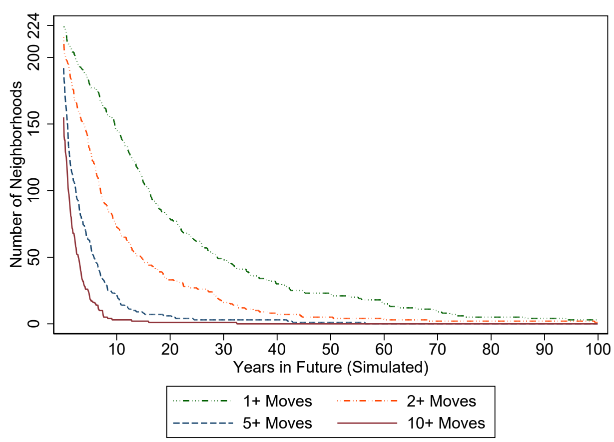
We present the simulated increase in segregation under all previously considered counterfactuals (baseline and 1-7) along with each of these new counterfactuals (8-9) in Figure 10. Our results in counterfactuals 1-7 are effectively unchanged in the models with and without price data. Comparing counterfactual 8 with the baseline counterfactual, we see that differential price responses have a minuscule effect on segregation dynamics in both the medium-run and the long-run. The similarity of the effects under counterfactuals 8 and 9 (and counterfactuals 3 and 9) also reveal that differential price responses and discrimination do not interact to impact segregation dynamics in a meaningful way.

Figure 9: Number of Neighborhoods In Flux - Alternative Price Responses (Simulated)

(a) Symmetric Price Responses

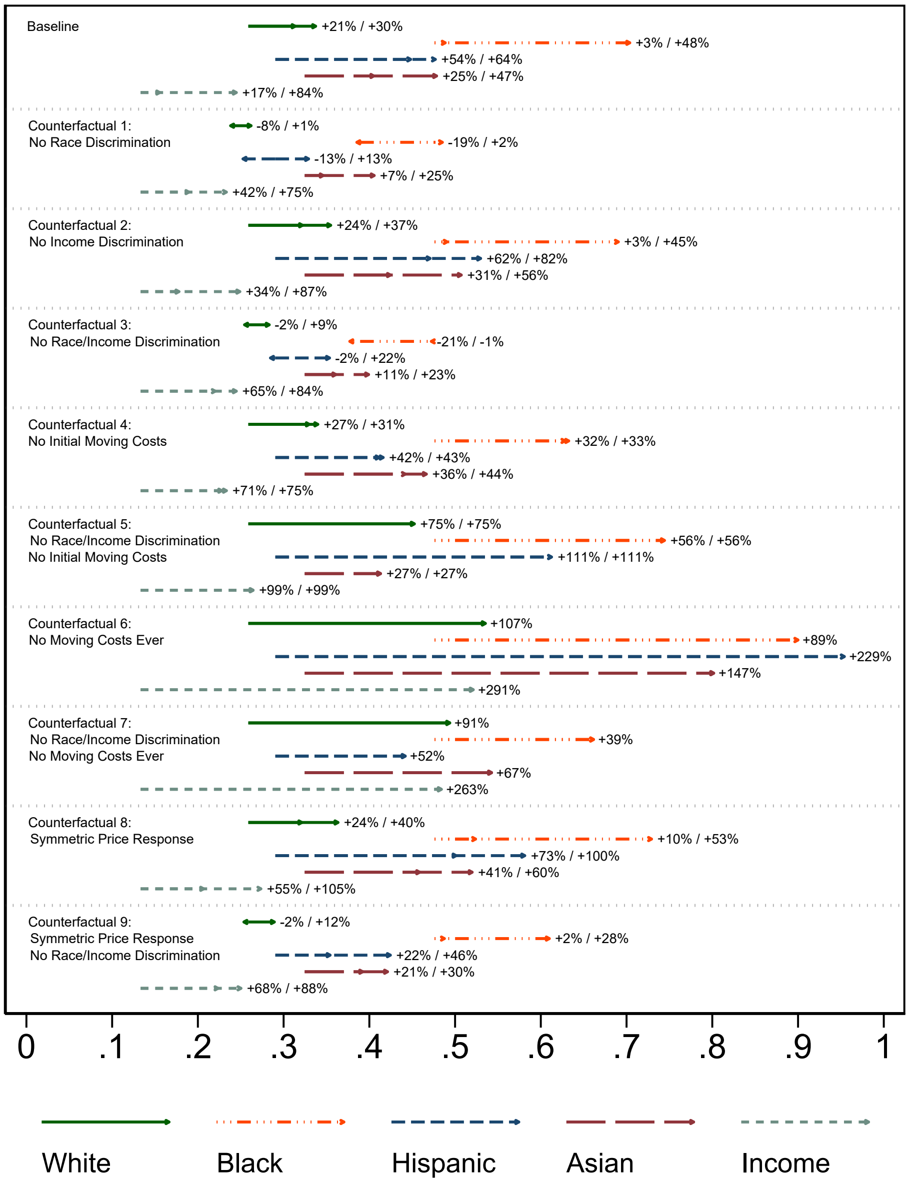


(b) Symmetric Price Responses, No Racial or Income Discrimination



Notes: Each panel shows the number of neighborhoods with at least one, two, five or ten moves that change their socioeconomic composition (out of a total of 224 neighborhoods) under a different counterfactual. Simulation begins in November 2004.

Figure 10: Medium- and Long-Run Changes in Average Neighborhood Prices (Simulated)



Notes: The arrows represent the changes in simulated average neighborhood prices at various quantiles from November 2004 in the absence of exogenous shocks. The first arrowhead corresponds to a "medium-run" that is 10 years into the future, and the second arrow corresponds to a "long-run" that is over 166 years into the future after neighborhoods have reached steady state. All prices are in constant 2004 dollars. Details of each counterfactual are presented in Section 4.2.

Although incorporating price data explicitly does not affect our analysis of segregation dynamics, it does allow us to explore how neighborhood prices will evolve under these counterfactual scenarios. We present these results in Figure 11. As shown in the baseline counterfactual, the socioeconomic sorting of households across neighborhoods has a negligible impact on prices in the medium-run. In the long-run, socioeconomic sorting does cause more substantial price appreciation in the most expensive neighborhoods, but only on the order of 10% over the course of a century and a half. This pales in comparison with the large price appreciation that we observe in the Bay Area, which suggests that ξ is the most relevant factor to study housing price dynamics. When moving costs are eliminated, we find larger price appreciation between the 25th and 75th percentiles, on the order of 10%.

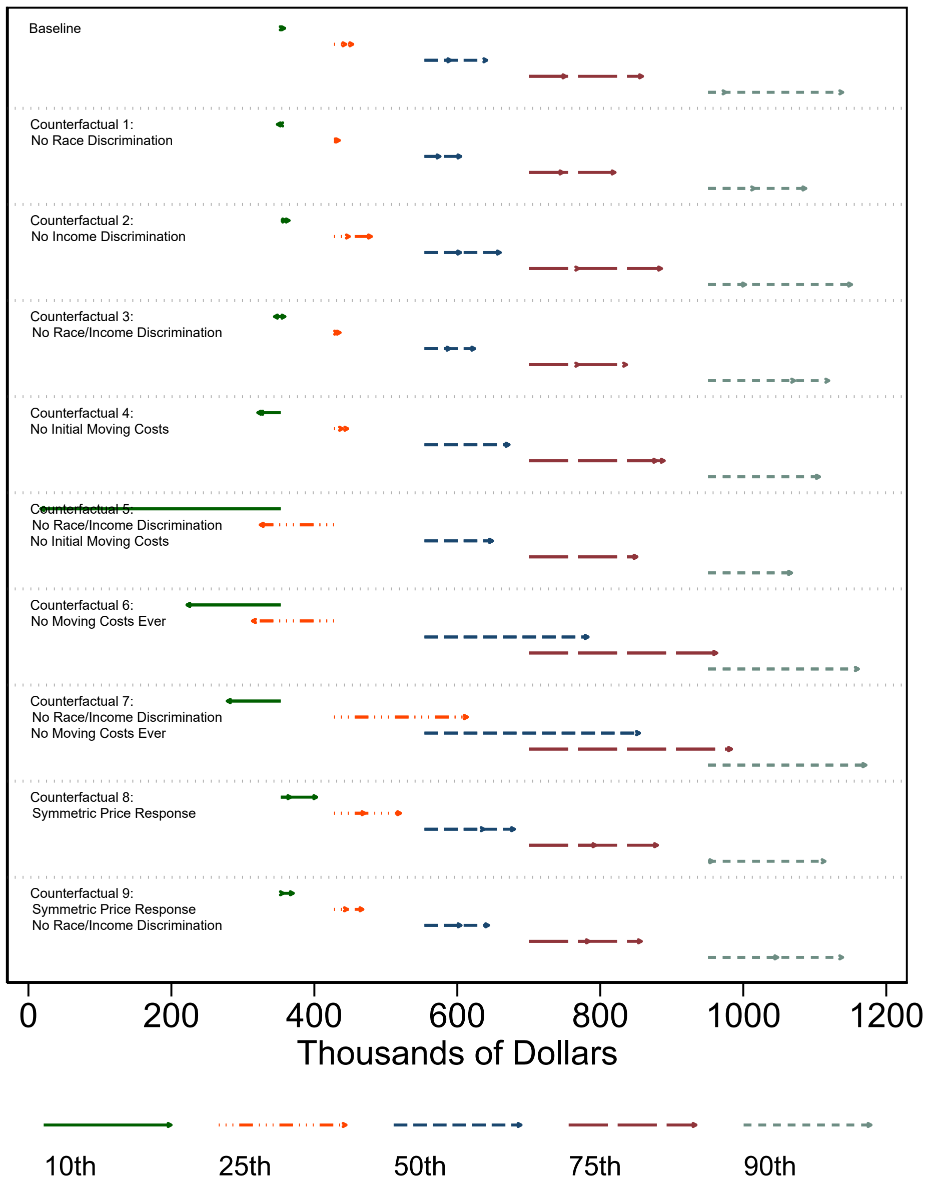
6 Conclusion

Neighborhoods constantly evolve: their amenities are not static and their residents are in flux. Theoretical models of segregation tend to attribute this evolution to endogenous changes in neighborhood residents arising from discrimination, while disaggregated models of residential choice tend to attribute this evolution to exogenous changes in other amenities. In this paper, we develop an empirical framework that synthesizes these two approaches and provides new perspectives on how the aggregate phenomenon of segregation arises from the accumulation of disaggregate residential choices.

We use this framework to study the determinants of race and income segregation in the San Francisco Bay Area from 1990 to 2004. By delineating the interconnected roles of socioeconomic discrimination, other neighborhood amenities, incomplete information, moving costs, initial allocations of households across neighborhoods, and heterogeneity in price-sensitivity, we explore the underlying forces that drive segregation through counterfactual analyses. We find that while racial discrimination leads to segregation, frictions (primarily moving costs and incomplete information) prevent much desired sorting from occurring, which weakens the feedback loop generated by endogenous discriminatory sorting. Other amenities also contribute to segregation, though to a lesser extent.

An important caveat in our analysis is that we do not observe the socioeconomic composition of renters over time. This may be less damaging to our conclusions if the aspects of the expected composition of neighborhoods that are most relevant to sorting

Figure 11: Medium- and Long-Run Changes in Average Neighborhood Prices - Full Integration (Simulated)



Notes: The arrows represent the changes in simulated average neighborhood prices at various quantiles from November 2004 in the absence of exogenous shocks. The first arrowhead corresponds to a "medium-run" that is 10 years into the future, and the second arrow corresponds to a "long-run" that is over 166 years into the future after neighborhoods have reached steady state. All prices are in constant 2004 dollars. Details of each counterfactual are presented in Section 4.4.

decisions are the ones proxied by the actual composition of homeowners (e.g., different allocations of local public goods spending depending on the socioeconomic composition of local taxpayers). However, this may be a concern in neighborhoods with lower rates of homeownership if the aspects of the expected composition of neighborhoods that are most relevant to sorting decisions are the compositions of the people that *use* public goods and, at the same time, landlords' socioeconomic status is a poor predictor of tenants' socioeconomic status. In any case, because renters face relatively lower moving costs than homeowners, we would expect to find patterns of segregation somewhere in between our baseline findings and our counterfactual findings without moving costs. Future research with access to better data is needed to address these issues.

Richer data would also provide opportunities to study socioeconomic segregation at finer levels; for instance, we could consider more income groups or we could disaggregate Hispanics by race, etc. Ultimately, we view our framework as a platform for the empirical analysis of determinants of segregation that can be directly adapted to various contexts. The use of this framework to study sorting along different demographic dimensions (e.g., race, income, partisanship, education) in different settings (e.g., neighborhoods, schools, virtual communities, physical venues) could prove valuable in revealing the importance of different cleavages in our society.

References

Aaronson, Daniel, Daniel A Hartley and Bhashkar Mazumder. 2020. "The Effects of the 1930s HOLC'Redlining'Maps.".

Ackerberg, Daniel A. 2020. "Timing Assumptions and Efficiency: Empirical Evidence in a Production Function Context." *Working Paper* .

Aguirregabiria, Victor and Pedro Mira. 2010. "Dynamic discrete choice structural models: A survey." *Journal of Econometrics* 156(1):38–67.

Angrist, J.D. and J.S. Pischke. 2009. *Mostly harmless econometrics: An empiricist's companion*. Princeton Univ Pr.

Arellano, Manuel and Stephen Bond. 1991. "Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations." *The review of economic studies* 58(2):277–297.

Banzhaf, H. Spencer and Randall P. Walsh. 2013. "Segregation and Tiebout sorting: The link between place-based investments and neighborhood tipping." *Journal of Urban Economics* 74(0):83 – 98.

Bartik, Timothy J. 1991. "Who benefits from state and local economic development policies?".

Bayer, P. and C. Timmins. 2005. "On the equilibrium properties of locational sorting models." *Journal of Urban Economics* 57(3):462–477.

Bayer, P. and C. Timmins. 2007. "Estimating Equilibrium Models Of Sorting Across Locations." *The Economic Journal* 117(518):353–374.

Bayer, P., F. Ferreira and R. McMillan. 2007. "A Unified Framework for Measuring Preferences for Schools and Neighborhoods." *Journal of Political Economy* 115(4):588–638.

Bayer, P., N. Keohane and C. Timmins. 2009. "Migration and hedonic valuation: The case of air quality." *Journal of Environmental Economics and Management* 58(1):1–14.

Bayer, P., R. McMillan and K. Rueben. 2004*a*. An equilibrium model of sorting in an urban housing market. Technical report National Bureau of Economic Research.

Bayer, Patrick and Robert McMillan. 2012. "Tiebout sorting and neighborhood stratification." *Journal of Public Economics* 96(11-12):1129–1143.

Bayer, Patrick, Robert McMillan, Alvin Murphy and Christopher Timmins. 2016. "A Dynamic Model of Demand for Houses and Neighborhoods." *Econometrica* 84(3):893–942.

Bayer, Patrick, Robert McMillan and Kim S. Rueben. 2004*b*. "What drives racial segregation? New evidence using Census microdata." *Journal of Urban Economics* 56(3):514 – 535.

Becker, G.S. and K.M. Murphy. 2000. *Social Economics: Market Behavior in a Social Environment*. Harvard University Press.

Berry, S., J. Levinsohn and A. Pakes. 1995. "Automobile Prices in Market Equilibrium." *Econometrica* 63(4):841–890.

Berry, S.T. 1994. "Estimating Discrete-Choice Models of Product Differentiation." *The RAND Journal of Economics* pp. 242–262.

Caetano, Gregorio. 2019. "Neighborhood sorting and the value of public school quality." *Journal of Urban Economics* 114:103193.

Caetano, Gregorio and Vikram Maheshri. 2017. "School segregation and the identification of tipping behavior." *Journal of Public Economics* 148:115–135.

Caetano, Gregorio and Vikram Maheshri. 2023. "Explaining Recent Trends in US School Segregation." *Journal of Labor Economics* 41(1):175–203.

Card, D., A. Mas and J. Rothstein. 2008a. "Tipping and the Dynamics of Segregation." *Quarterly Journal of Economics* 123(1):177–218.

Christensen, P. and C. Timmins. 2019. "Sorting or Steering: Experimental Evidence on the Economic Effects of Housing Discrimination." *NBER Working Paper* .

Cutler, David M, Edward L Glaeser and Jacob L Vigdor. 1999. "The rise and decline of the American ghetto." *Journal of political economy* 107(3):455–506.

Epple, D. and H. Sieg. 1999. "Estimating Equilibrium Models of Local Jurisdictions." *Journal of Political Economy* 107(4).

Epple, D., T. Romer and H. Sieg. 2001. "Interjurisdictional sorting and majority rule: An empirical analysis." *Econometrica* 69(6):1437–1465.

Epple, D., T. Romer and H. Sieg. 2003. "Interjurisdictional sorting and majority rule: An empirical analysis." *Econometrica* 69(6):1437–1465.

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.

Fang, Hanming and Andrea Moro. 2011. "Theories of statistical discrimination and affirmative action: A survey." *Handbook of social economics* 1:133–200.

Ferreira, Fernando. 2010. "You can take it with you: Proposition 13 tax benefits, residential mobility, and willingness to pay for housing amenities." *Journal of Public Economics* 94(9):661–673.

Goldsmith-Pinkham, Paul, Isaac Sorkin and Henry Swift. 2020. "Bartik instruments: What, when, why, and how." *American Economic Review* 110(8):2586–2624.

Ihlanfeldt, Keith R and Benjamin Scafidi. 2002. "Black self-segregation as a cause of housing segregation: Evidence from the multi-city study of urban inequality." *Journal of Urban Economics* 51(2):366–390.

Kennan, John and James R Walker. 2011. "The effect of expected income on individual migration decisions." *Econometrica* 79(1):211–251.

Kiel, Katherine A and Jeffrey E Zabel. 1996. "House price differentials in US cities: Household and neighborhood racial effects." *Journal of housing economics* 5(2):143–165.

Kuminoff, Nicolai V, V Kerry Smith and Christopher Timmins. 2013. "The new economics of equilibrium sorting and policy evaluation using housing markets." *Journal of Economic Literature* 51(4):1007–1062.

Ladd, Helen F. 1998. "Evidence on discrimination in mortgage lending." *Journal of Economic Perspectives* 12(2):41–62.

Li, Nicholas Y. 2023. "Racial Sorting, Restricted Choices, and the Origins of Residential Segregation in U.S. Cities." *mimeo.* .

Manski, Charles F. 2004. "Measuring expectations." *Econometrica* 72(5):1329–1376.

Massey, Douglas and Nancy A Denton. 1993. *American apartheid: Segregation and the making of the underclass*. Harvard university press.

McGuire, Martin. 1974. "Group segregation and optimal jurisdictions." *Journal of Political Economy* 82(1):112–132.

Morrill, Richard L. 1989. "School busing and demographic change." *Urban Geography* 10(4):336–354.

Myrdal, G. 1944. "An American dilemma.".

Olley, Steven and Ariel Pakes. 1996. "The dynamics of productivity in the telecomunications equipment industry." *Econometrica* 64:1263–97.

Ondrich, Jan, Stephen Ross and John Yinger. 2003. "Now you see it, now you don't: why do real estate agents withhold available houses from black customers?" *Review of Economics and Statistics* 85(4):854–873.

Ouazad, Amine and Romain Rancière. 2016. "Credit standards and segregation." *Review of Economics and Statistics* 98(5):880–896.

Ringo, Daniel. 2013. "Residential Clustering by Race and Income: What Drives the Sorting?" *Mimeo* .

Rosen, Sherwin. 1974. "Hedonic prices and implicit markets: product differentiation in pure competition." *Journal of political economy* 82(1):34–55.

Schelling, Thomas C. 1969. "Models of Segregation." *The American Economic Review* 59(2):488–493.

Schelling, Thomas C. 1971. "Dynamic Models of Segregation." *Journal of Mathematical Sociology* 1:143–186.

Shertzer, Allison, Tate Twinam and Randall P Walsh. 2021. "Zoning and segregation in urban economic history." *Regional Science and Urban Economics* p. 103652.

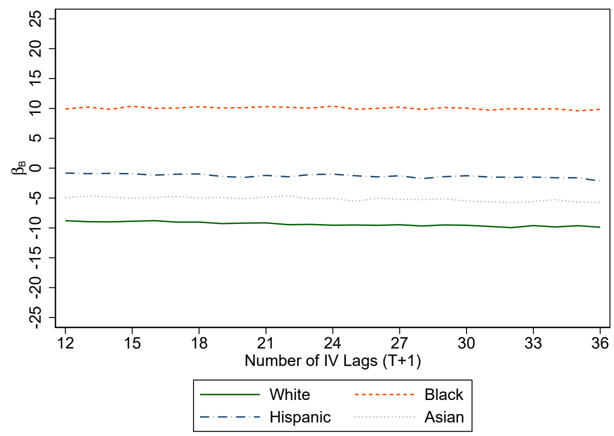
A Online Appendix: Tables and Figures

Table 4: OLS Estimates of Responses to the Race and Income Compositions of Neighborhoods ($\boldsymbol{\beta}$)

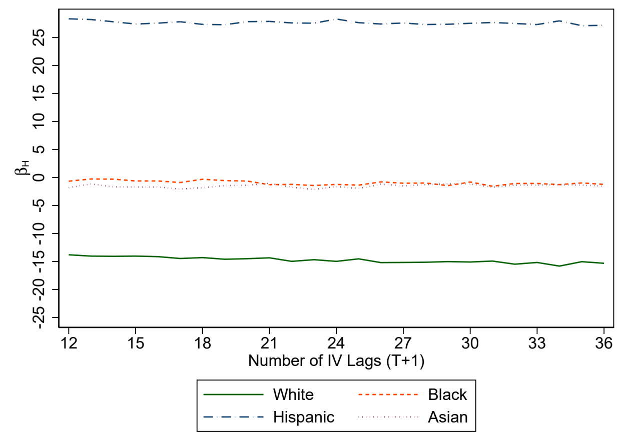
	White		Black		Hispanic		Asian	
	Rich	Poor	Rich	Poor	Rich	Poor	Rich	Poor
Responses to:								
Black Share	-12.25	-7.32	13.47	16.35	-0.81	8.49	-4.81	1.02
	(0.35)	(0.52)	(0.43)	(0.43)	(0.33)	(0.35)	(0.37)	(0.34)
Hispanic Share	-17.69	2.66	8.18	8.86	35.92	43.64	-1.76	16.25
	(0.51)	(0.53)	(0.47)	(0.38)	(0.67)	(0.53)	(0.46)	(0.48)
Asian Share	-5.63	-9.99	-0.44	-2.64	-0.21	-3.54	24.24	26.63
	(0.30)	(0.38)	(0.34)	(0.28)	(0.35)	(0.37)	(0.56)	(0.61)
Poor Share	-7.01	1.28	-7.52	1.11	-12.77	-2.60	-16.17	-2.91
	(0.35)	(0.39)	(0.30)	(0.27)	(0.30)	(0.35)	(0.37)	(0.34)
R^2	0.38							
Num. of Observations	147,840							

Notes: This specification includes only group-month fixed effects as controls. White is the omitted racial share and rich is the omitted income share. All standard errors clustered by group-month.

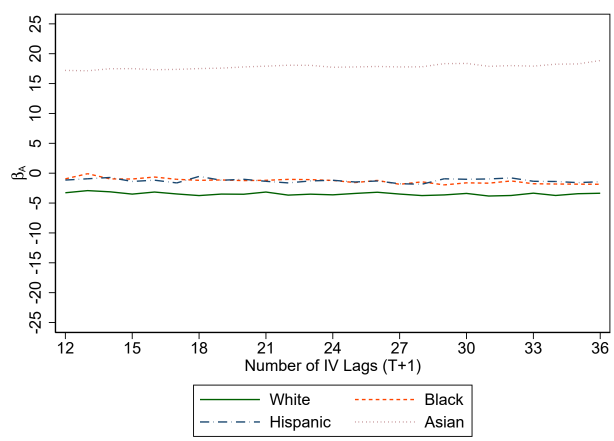
Figure 12: Responses of Rich Households of Different Races to Race and Income Compositions for Different Values of T



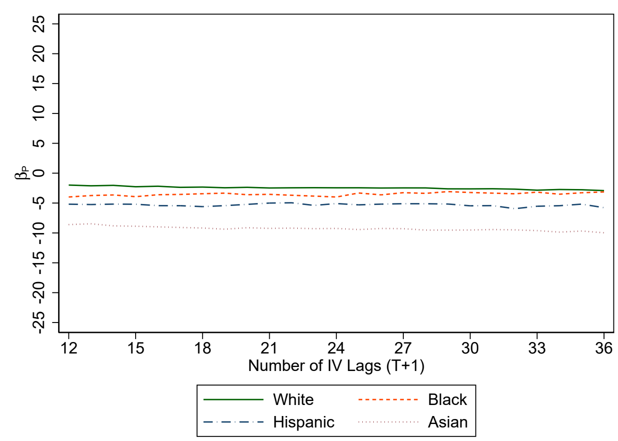
(a) Responses to Black Share



(b) Responses to Hispanic Share



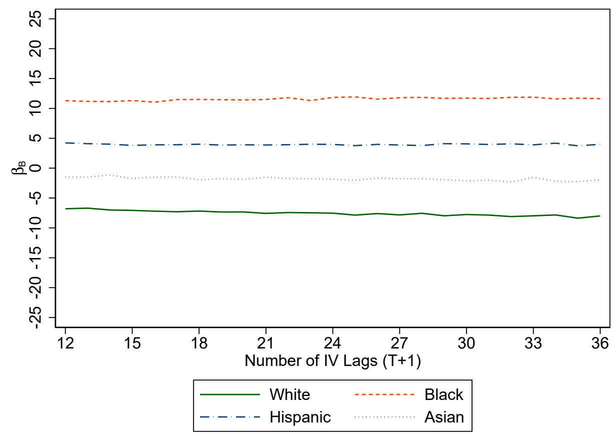
(c) Responses to Asian Share



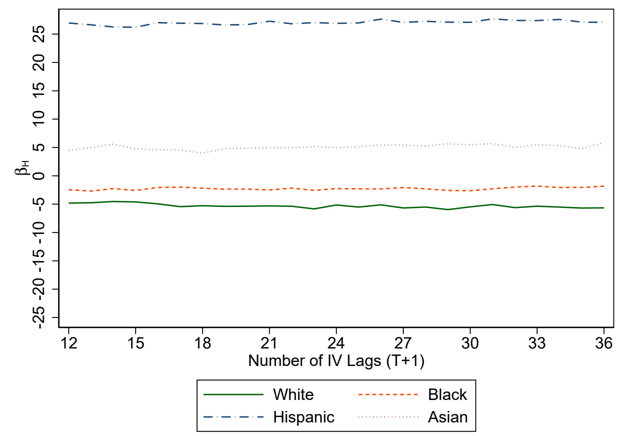
(d) Responses to Poor Share

Notes: Each panel shows $\hat{\beta}_{g,g'}$ for all g' for different values of $T+1$, the lag of the Instrumental Variable, $\boldsymbol{s_{jt-(T+1)}}$. In all panels of this figure g represents poor Whites, Blacks, Hispanics and Asians. We control for $\boldsymbol{v_{jt-T}}$ in all specifications. For comparison, we keep the sample constant in all cases, so the first 37 months of the sample are not used in the estimation of $\boldsymbol{\beta}$ in any specification.

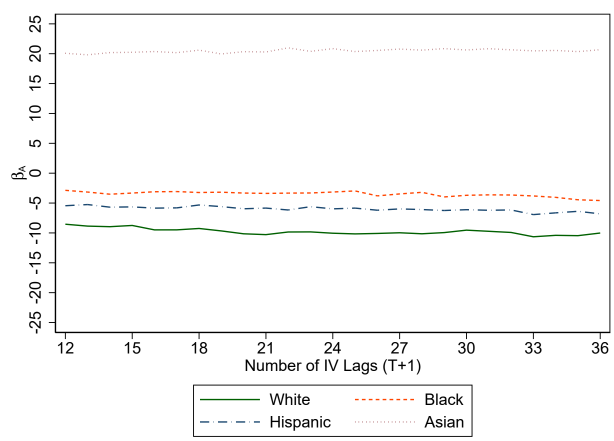
Figure 13: Responses of Poor Households of Different Races to Race and Income Compositions for Different Values of T



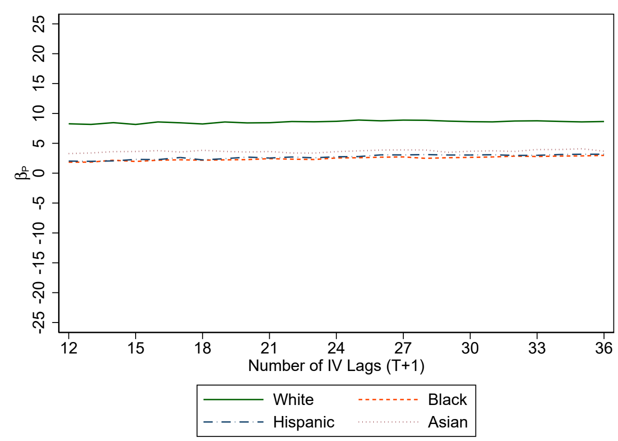
(a) Responses to Black share



(b) Responses to Hispanic Share



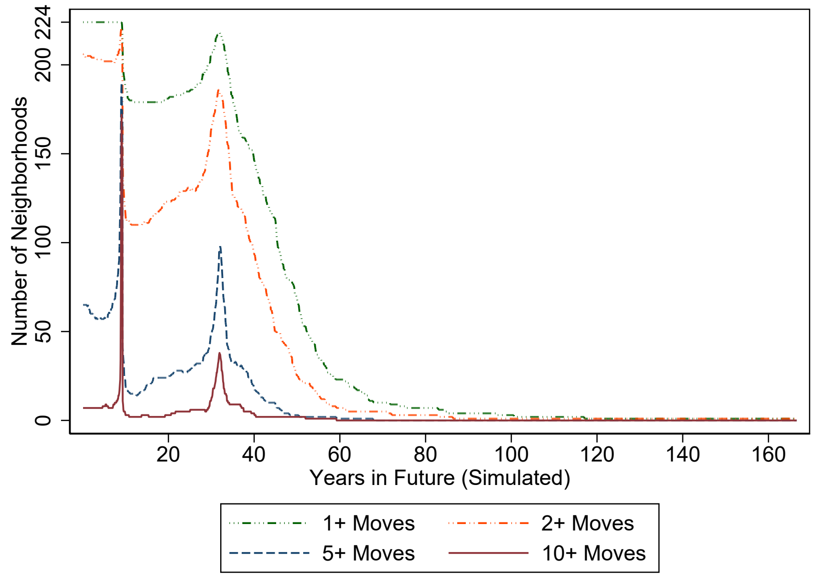
(c) Responses to Asian Share



(d) Responses to Poor Share

Notes: Each panel shows $\hat{\beta}_{g,g'}$ for all g' for different values of $T+1$, the lag of the Instrumental Variable, $\boldsymbol{s_{jt-(T+1)}}$. In all panels of this figure g represents poor Whites, Blacks, Hispanics and Asians. We control for $\boldsymbol{v_{jt-T}}$ in all specifications. For comparison, we keep the sample constant in all cases, so the first 37 months of the sample are not used in the estimation of $\boldsymbol{\beta}$ in any specification.

Figure 14: Number of Neighborhoods In Flux (Simulated)



Notes: Figure shows the number of neighborhoods with at least one, two, five or ten moves that change their socioeconomic composition (out of a total of 224 neighborhoods). Simulation begins in November 2004 and continues for 200 years.

B Estimation and Simulation With Price Data

B.1 Estimation

Stage 1: Estimation of v^e_{gjt} and ϕ_g

This stage follows exactly as in the baseline case without price data presented in Section 3.2.

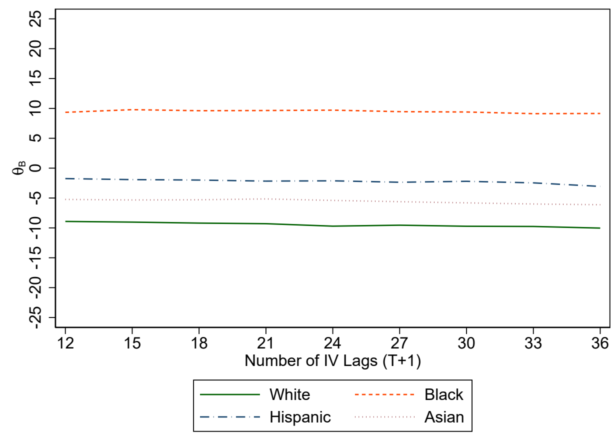
Stage 2: Estimation of $\boldsymbol{\beta_g}$, π_g and ρ_g

We re-write equation (21) based on observed quantities as

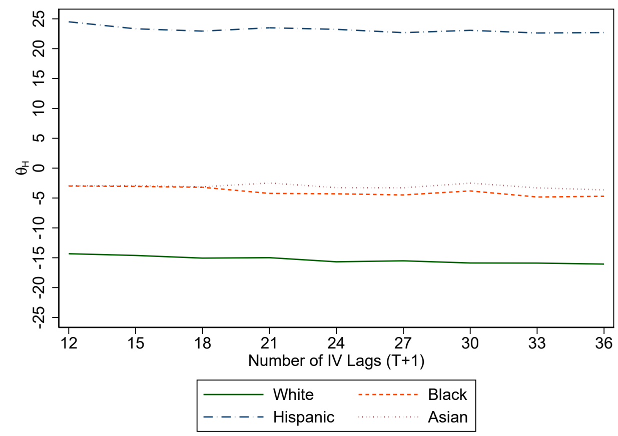
$$\hat{v}^e_{gjt} = \boldsymbol{\beta_g'} \boldsymbol{s_{jt}} + \pi_g P_{jt} + \boldsymbol{\gamma_g'} \boldsymbol{\hat{v}^e_{jt-T}} + \underbrace{\xi_{gjt} + \hat{v}^e_{gjt} - v^e_{gjt} - \boldsymbol{\gamma_g'} \boldsymbol{\hat{v}^e_{jt-T}}}_{\text{error}_{gjt}}, \tag{24}$$

We estimate $\boldsymbol{\beta_g}$ in the equation above via Two Stage Least Squares using $\boldsymbol{s_{jt-(T+1)}}$ and $P_{jt-(T+1)}$ as an IV for $\boldsymbol{s_{jt}}$, controlling for $\boldsymbol{\hat{v}^e_{jt-T}}$. Based on Assumption 1, $\hat{v}^e_{gjt}$ converges to v^e_{gjt}, and $\boldsymbol{\hat{v}^e_{jt-T}}$ converges to $\boldsymbol{v^e_{jt-T}}$, and Assumption **3′** implies $\boldsymbol{s_{jt-(T+1)}}$ and $P_{jt-(T+1)}$ are both uncorrelated to ξ_{gjt}, conditional on $\boldsymbol{v^e_{jt-T}}$.

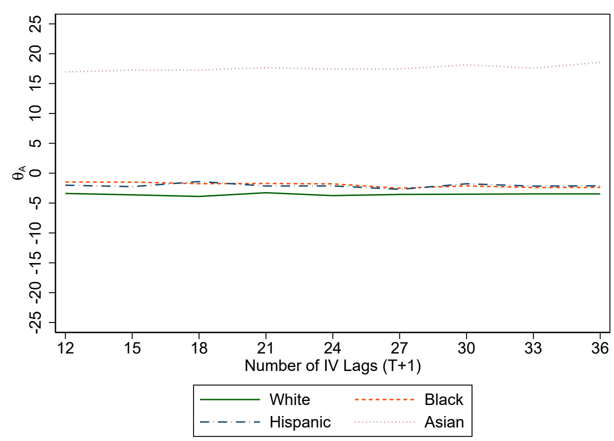
Figure 15: Responses of Rich Households of Different Races to Race and Income Compositions ($\theta_{g,g'}$) for Different Values of T: Price Explicitly Included



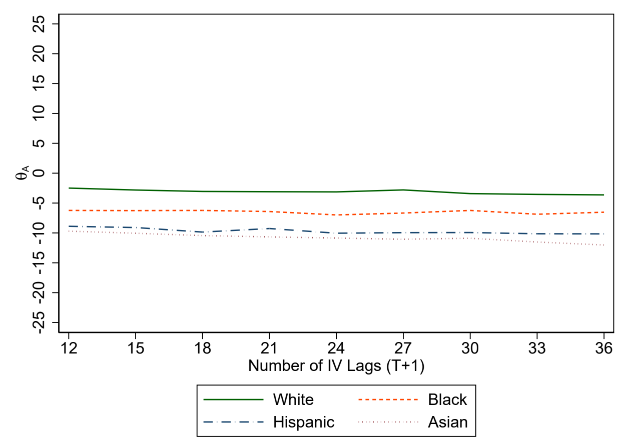
(a) Responses to Black Share



(b) Responses to Hispanic Share



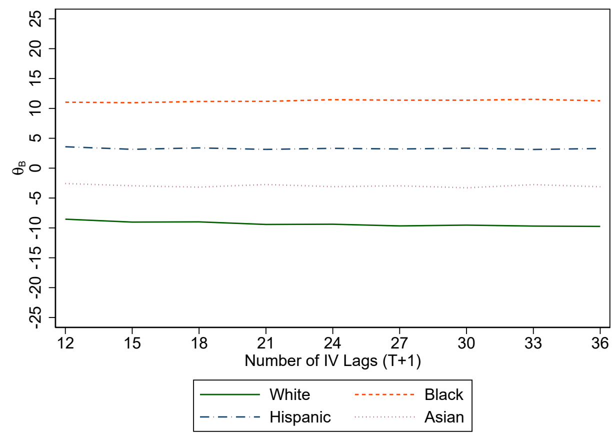
(c) Responses to Asian Share



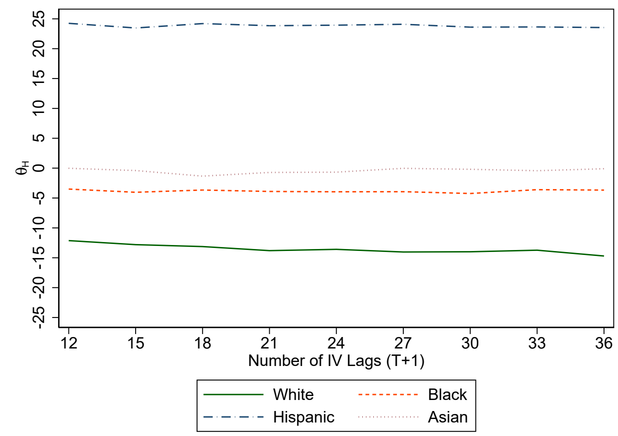
(d) Responses to Poor Share

Notes: Each panel shows $\hat{\theta}_{g,g'}$ for all g' for different values of $T+1$, the lag of the Instrumental Variable, $\boldsymbol{s_{jt-(T+1)}}$. In all panels of this figure g represents poor Whites, Blacks, Hispanics and Asians. We control for $\boldsymbol{v_{jt-T}}$ in all specifications. For comparison, we keep the sample constant in all cases, so the first 37 months of the sample are not used in the estimation of $\boldsymbol{\theta}$ in any specification.

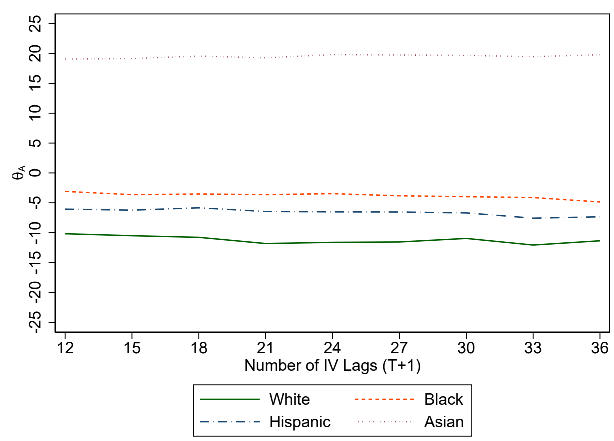
Figure 16: Responses of Poor Households of Different Races to Race and Income Compositions ($\theta_{g,g'}$) for Different Values of T: Price Explicitly Included



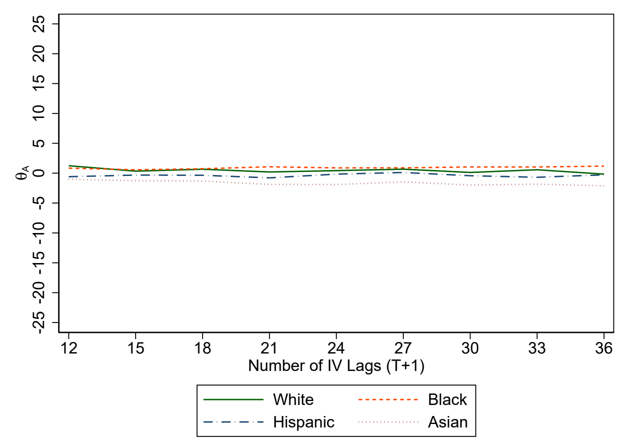
(a) Responses to Black share



(b) Responses to Hispanic Share



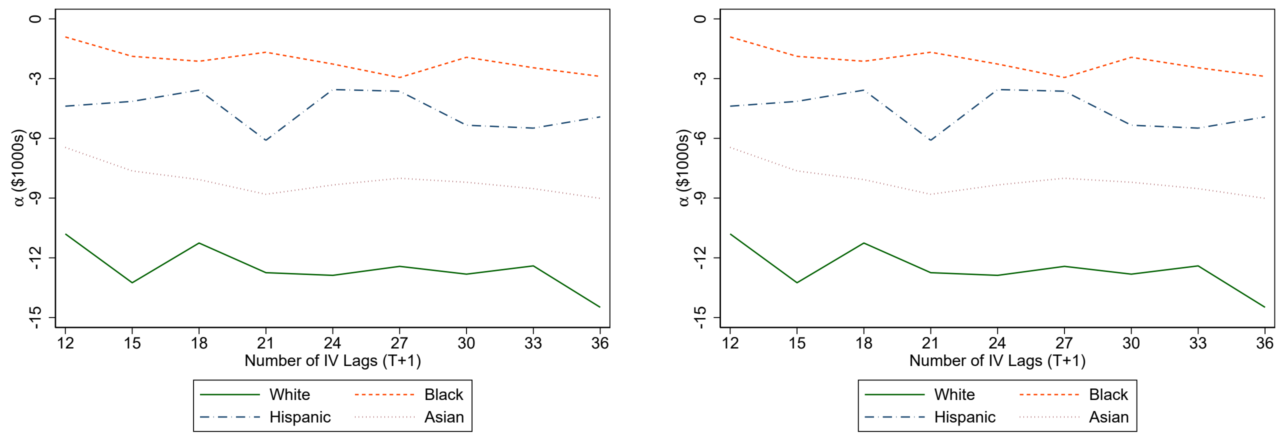
(c) Responses to Asian Share



(d) Responses to Poor Share

Notes: Each panel shows $\hat{\theta}_{g,g'}$ for all g' for different values of $T+1$, the lag of the Instrumental Variable, $\boldsymbol{s_{jt-(T+1)}}$. In all panels of this figure g represents poor Whites, Blacks, Hispanics and Asians. We control for $\boldsymbol{v_{jt-T}}$ in all specifications. For comparison, we keep the sample constant in all cases, so the first 37 months of the sample are not used in the estimation of $\boldsymbol{\theta}$ in any specification.

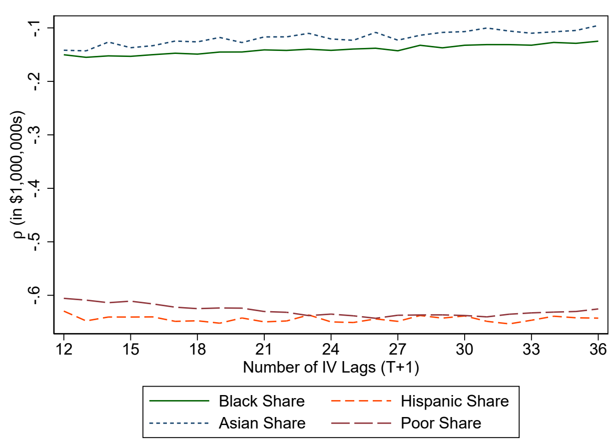
Figure 17: Responses of Households to Prices (α_g) for Different Values of T: Price Explicitly Included



(a) Rich Households (b) Poor Households

Notes: The plot shows $\hat{\alpha}_g$ for all g for different values of T, the lag of the Instrumental Variables, $\boldsymbol{s_{jt-T}}$ and P_{jt-T}. We set $T' = 12$ in all cases. For comparison, we keep the sample constant in all cases, so the first 37 months of the sample are not used in the estimation of $\boldsymbol{\alpha}$ for all values of T.

Figure 18: Implicit Price of Race and Income Compositions ($\rho_{g'}$) for Different Values of T



Notes: Each panel shows $\hat{\rho}_{g'}$ for all g' for different values of $T+1$, the lag of the Instrumental Variable, $\boldsymbol{s_{jt-(T+1)}}$. In all panels of this figure g represents poor Whites, Blacks, Hispanics and Asians. We control for $\boldsymbol{v_{jt-T}}$ in all specifications. For comparison, we keep the sample constant in all cases, so the first 37 months of the sample are not used in the estimation of $\boldsymbol{\rho}$ in any specification.

Similarly, re-writing equation (22) to include control variable $\boldsymbol{v^e_{jt-T}}$, we obtain

$$P_{jt} = \alpha + \boldsymbol{\rho}'\boldsymbol{s_{jt}} + \boldsymbol{\lambda}'\boldsymbol{\hat{v}^e_{jt-T}} + \underbrace{\eta_{jt} - \boldsymbol{\lambda}'\boldsymbol{\hat{v}^e_{jt-T}}}_{\text{error}_{jt}}. \tag{25}$$

We estimate $\boldsymbol{\rho}$ in equation (22) via Two Stage Least Squares using $\boldsymbol{s_{jt-(T+1)}}$ as an IV for $\boldsymbol{s_{jt}}$, controlling for $\boldsymbol{\hat{v}^e_{jt-T}}$. Based on Assumption 1, $\boldsymbol{\hat{v}^e_{jt-T}}$ converges to $\boldsymbol{v^e_{jt-T}}$, and Assumption **6′** implies $\boldsymbol{s_{jt-(T+1)}}$ is uncorrelated to η_{jt} conditional on $\boldsymbol{v^e_{jt-T}}$.

B.2 Simulation

Given counterfactual values of the parameter vector $(\boldsymbol{\tilde{\phi}}, \boldsymbol{\tilde{\beta}}, \boldsymbol{\tilde{\pi}}, \boldsymbol{\tilde{\rho}})$ and counterfactual initial conditions $(\boldsymbol{\tilde{s}^0}, \boldsymbol{\tilde{\xi}^0}, \boldsymbol{\tilde{\eta}^0}, \boldsymbol{\tilde{N}^0})$, we obtain $\boldsymbol{s^{\tau+1}} = \boldsymbol{s}(\boldsymbol{\tilde{s}^\tau})$ for each $\tau \geq 0$ using equations (15) and (16), as in the simulation without prices. However, instead of using equation (17), we update $v^e_{gj}(\boldsymbol{\tilde{s}^\tau})$ according to

$$v^e_{gj}\left(\boldsymbol{\tilde{s}^\tau}\right) = \boldsymbol{\tilde{\beta}'_g}\boldsymbol{\tilde{s}^\tau_j} + \tilde{\pi}_g P_j\left(\boldsymbol{\tilde{s}^\tau}\right) + \tilde{\xi}^0{}_{gj}, \tag{26}$$

where

$$P_j\left(\boldsymbol{\tilde{s}^\tau}\right) = \boldsymbol{\tilde{\rho}}'\boldsymbol{\tilde{s}^\tau_j} + \tilde{\eta}^0_j. \tag{27}$$

B.3 Interpretation of New Model Parameters

Interpretation of $\boldsymbol{\pi}$ $\boldsymbol{\pi}$ incorporates two components: $\pi_g = \frac{\partial v^e_{gjt}}{\partial P_{jt}} = \frac{\partial u^e_{gjt}}{\partial P_{jt}} + \frac{\partial CV^e_{gjt}}{\partial P_{jt}}$. The flow utility component should be negative since households prefer to pay lower prices for their house, all else constant. However, the sign of the continuation value component is theoretically ambiguous since it depends on whether the price of a neighborhood today (relative to the price of the actual house they buy) signals disproportionate expected future appreciation relative to an otherwise comparable neighborhood, which would have consequences for homeowners' expected wealth.[33]

[33] The buying and selling of a house may impact household wealth. Despite its undeniable importance when studying the behavior of homeowners, we do not explicitly model the effects of moving on wealth, and we do not allow for household heterogeneity by wealth either. In our context, doing so would substantially increase the number of groups of households that we would need to consider and would render our analysis infeasible since there are not enough households of each race and income level to study their decisions by wealth levels. Note, however, that wealth is partially incorporated in our analysis since our parameters are allowed to vary by group, and these groups may have different wealth on average.

Interpretation of $\boldsymbol{\rho}$ The price response parameter, $\boldsymbol{\rho}$, incorporates both supply-side and demand-side considerations, though we are agnostic as to their particulars. $\boldsymbol{\rho}$ should simply be understood as the best linear approximation of the (potentially non-linear) causal expectation function (CEF) of $\boldsymbol{s_{jt}}$ on P_{jt} (Angrist and Pischke (2009)). We do not invoke the assumptions that allow this equation to be interpreted as a hedonic price equation from which tastes or costs can be recovered (Rosen (1974)), nor do we invoke the assumptions that allow $\boldsymbol{\rho}$ to be interpreted as a marginal willingness to pay parameter. For instance, $\boldsymbol{\rho}$ would equal zero if household demand was unresponsive to changes in $\boldsymbol{s_{jt}}$, or if the supply of housing was perfectly elastic. We would be unable to distinguish between these two scenarios given an estimate of $\hat{\boldsymbol{\rho}} = \mathbf{0}$, but this would not affect our analysis of segregation dynamics.

C Monte Carlo Simulations

C.1 Set up

We consider the exact model from Section 3 with $G = 2$ instead of $G = 8$ for simplicity and clarity. We denote the two groups as A and B. For each sample, $N_{A0} = 1,000$, $N_{B0} = 9,000$, $J = 224$ and $T = 179$ (the values of J and T correspond to the values in our data). Each sample is created as follows:

1. Draw ξ_{gj0} from $\xi_{gj0} \sim \mathcal{N}(0, 4)$ for each g and $j \geq 1$, and normalize $\xi_{g00} = 0$.

2. For $j \geq 1$, calculate $s_{j0} = \frac{N_{Aj0}}{N_{Aj0}+N_{Bj0}}$, where N_{gj0} is obtained by

$$N_{gj0} = N_{g0} \cdot \frac{\exp(\xi_{gj0})}{\sum_{k=0}^{J} \exp(\xi_{gk0})}.$$

   i.,e., households sort only on the basis of ξ in period 0. Steps 1 and 2 create the initial conditions, which are used to simulate the sample for $t \geq 1$. Data for $t = 0$ is not included as part of the simulated sample.

3. For $t \geq 1$ and $j \geq 1$ and draw η_{gjt} from $\eta_{gjt} \sim \mathcal{N}(0, 1)$.

4. For $t \geq 1$ and $j \geq 1$, obtain ξ_{gjt} from the following equation: $\xi_{gjt} = 0.8\xi_{gjt-1} + \eta_{gjt} - 0.8\eta_{gjt-1}$.

5. For $t \geq 1$, obtain $s_{jt} = \frac{N_{Ajt}}{N_{Ajt}+N_{Bjt}}$, where N_{gjt} is obtained by

$$N_{gjt} = \quad N_{gjt-1} \times \left(\frac{\exp\left(v^e_{gjt}\right)}{\sum_{j'=0}^{J} \exp\left(v^e_{gj't} - \phi_g\right) + \exp\left(v^e_{gjt}\right)} \right) + \\ + \sum_{k=1}^{J} N_{gkt-1} \times \left(\frac{\exp\left(v^e_{gjt} - \phi_g\right)}{\sum_{j'=0}^{J} \exp\left(v^e_{gj't} - \phi_g\right) + \exp\left(v^e_{gkt}\right)} \right)$$

6. with $v^e_{g0t} = 0$ and $v^e_{gjt} = \beta_g \cdot s_{jt} + \xi_{gjt}$ for $j \geq 1$. We set $\phi_A = 20$, $\phi_B = 15$, $\beta_A = 3$ and $\beta_B = -3$.[34]

These steps create a sample of N_{gjt} for $g = A, B$, $j = 1, ..., J$, and $t = 1, ..., T$. This is one sample. We repeat these steps to create a total of 1,000 samples.

C.2 Monte Carlo Results

Table 5 shows the Monte Carlo results. For each different estimator, we show both the bias and the standard deviation, which are the two inputs in the Mean Square Error calculation ($MSE = Bias^2 + SD^2$).

The first row shows the results for the moving cost estimator of ϕ_g. It performs well, with small biases, which is to be expected given that the Monte Carlo assumes the dynamic model from Section 3.1.

The remaining rows of the table show the results for the discriminatory response estimators of β_g. For context, we start with the naive OLS estimator, which is biased for both β_A and β_B, as expected. The other rows show the performance of the 2SLS estimator of β_g using $s_{jt-(T+1)}$ as an IV for s_{jt}, controlling for $\boldsymbol{v^e_{jt-T}}$. We consider different values of T to show how the results tend to be fairly stable as T grows. Irrespective of the lag, the IV estimators of β_A and β_B tend to perform well.

[34] We obtained nearly identical results under an alternative Monte Carlo simulation where we keep the total number of households of each group constant over time by instead calculating $s_{jt} = \frac{\tilde{N}_{Ajt}}{\tilde{N}_{Ajt}+\tilde{N}_{Bjt}}$, with $\tilde{N}_{gjt} = N_{g0} \frac{N_{gjt}}{\sum_{k=1}^{J} N_{gkt}}$. Allowing for changes over time in the total number of households of a given group across all J neighborhoods does not change the results, as expected, since such variation is absorbed by group-month fixed effects.

Table 5: Monte Carlo Results: Estimates of ϕ_g and β_g

	Linear Controls			
	Group A		Group B	
	Bias	SD	Bias	SD
ϕ_g: $\phi_A = 20$, $\phi_B = 15$	0.0002	0.0001	0.0010	0.0002
β_g: $\beta_A = 3$, $\beta_B = -3$				
OLS: $v^e_{gjt} = s_{jt}\beta_g + \xi_{gjt}$	0.1900	0.0593	-0.2103	0.0630
IV: $s_{jt-(T+1)}$ as instrument for s_{jt} inv$^e_{gjt}$ = $s_{jt}\beta_g + \boldsymbol{\gamma}' \cdot \boldsymbol{v^e_{jt-T}} + \text{error}_{gjt}$				
s_{jt-13} as IV, $\boldsymbol{v^e_{jt-12}}$ as controls	-0.0039	0.0292	0.0045	0.0290
s_{jt-19} as IV, $\boldsymbol{v^e_{jt-18}}$ as controls	0.0040	0.0294	-0.0024	0.0285
s_{jt-25} as IV, $\boldsymbol{v^e_{jt-24}}$ as controls	0.0058	0.0297	-0.0046	0.0288
s_{jt-31} as IV, $\boldsymbol{v^e_{jt-30}}$ as controls	0.0072	0.0311	-0.0040	0.0311
s_{jt-37} as IV, $\boldsymbol{v^e_{jt-36}}$ as controls	0.0080	0.0319	-0.0043	0.0320

Notes: Each specification includes group-month fixed effects as controls. Group B is the omitted group share.