Segregation by Choice? Separating Preferences from Restrictions in Housing Markets*

Milena Almagro — Chicago Booth & NBER

Aradhya Sood — University of Toronto

March 21, 2025

PRELIMINARY. NOT FOR DISTRIBUTION.

Abstract

How much of the observed segregation in neighborhoods is attributed to restrictions in the households' choice sets, such as those imposed by discrimination versus household preferences? To answer this question, we build and estimate a two-sided housing matching model that endogeneizes the differences in choice sets across demographic groups using a novel dataset that links households and developments to the historic street grid of the 1940 Minneapolis metro area. Using an instrumental variable approach, we first document that having *de jure* discriminatory restrictions in neighborhoods reduces the likelihood of non-White and White Eastern European immigrant households by 9.3 and 11.9 percentage points, respectively. Next, we demonstrate—both theoretically and through Monte Carlo simulations—that traditional discrete choice models of neighborhood demand with mis-specified choice sets produce biased estimates of preferences for neighborhood characteristics. We estimate our two-sided housing matching model, which allows for the joint estimation of different choice sets and preference parameters. Our results suggest that different demographic groups faced substantially different choice sets: on average, minority households are 26 percentage points less likely to have a neighborhood in their choice set compared to non-minority households even after controlling for *de jure* forms of discrimination. Moreover, ignoring these differences in choice sets over-estimates the role that *de jure* discriminatory tools play in shaping the location choices of minority households and, consequently, on segregation.

*Authors are listed in alphabetical order. All co-authors contributed equally. We thank the Mapping Prejudice team, and in particular Michael Corey, Kirsten Delegard, Kevin Ehrman-Solberg, and Penny Petersen, for sharing their covenants data and insights as well as the Urban Transition Historical GIS Project and John R. Logan for sharing their 1940 street grid for Minneapolis. We thank Dan Aaronson, Abi Adams-Prassl, Victor Aguirregabiria, Sara Bagagli, Leah Boustan, Giovanni Compiani, Mike Dinerstein, James Feigenbaum, Rick Hornbeck, Samuel Myers Jr., Steven L. Ross, Paulo Somaini and seminar participants at the AREUEA National Meetings, Federal Reserve Banks of Chicago and Minneapolis, Junior Spatial Economics Conference at Sciences Po, Russell Sage Foundation Emerging Scholars Conference, Southern Economic Association, and Urban Economics Association for their helpful comments. This research is supported by the Russell Sage Foundation and Gates Foundation. Ben Couillard, Nanxi Fu, Jason Hong, Sarah Ji, Xiaofeng Liu, Jonathan Potter, Sriram Tolety, and Elias van Emmerick provided excellent research assistance. Emails: milena.almagro@chicagobooth.edu and aradhya.sood@rotman.utoronto.ca.

1. Introduction

Segregation across space is observed everywhere, yet the underlying causes driving it remain unclear. Understanding the causes behind this segregation is central to a range of economic and policy questions, as segregation affects children's learning outcomes and their upward mobility (Collins and Margo, 2000; Durlauf, 2004; Chetty et al., 2016; Chyn and Katz, 2021). Although there is a long literature studying segregation, a central challenge remains distinguishing restrictions in the neighborhood choice sets for certain households—such as those arising from supply-side discrimination—from demand-side preferences of households. Much of the literature attributes racial and ethnic segregation to preferences to live among others of their own race or ethnicity (Bayer et al., 2007, 2022). However, recent research has found that discrimination, whether explicit or implicit, can also play a role in the observed sorting of households across neighborhoods (Aaronson et al., 2021; Sood and Ehrman-Solberg, 2023; Christensen and Timmins, 2022, 2023). In this paper, we ask: how much do restrictions in the choice set of neighborhoods versus preferences of households contribute towards segregation and shape equilibrium outcomes?

To understand the drivers behind observed segregation patterns, we construct a novel dataset that links households from the 1940 individual-count US census data and housing developments in the 1940 Minneapolis metro area to the street grid of that time. To our knowledge, this is the first dataset compiled at the development-household level in a historical context. Additionally, we can also link lot-level data on the *de jure* instrument of housing discrimination in the form of racial covenants. These covenants, added to property deeds, explicitly prohibited certain racial and ethnic minorities from buying or renting homes and served as the primary *de jure* tool of housing discrimination in urban and suburban neighborhoods in the early to mid-20th century in the US.

This setting helps us in several ways. First, because the Minneapolis metro area was rapidly expanding during this time, the supply-side decisions to admit households into a neighborhood were made at the development level rather than on an individual house seller level, creating a better empirical setting for us to study restrictions on the supply side

of the housing market. Second, covenants in the housing market served a dual purpose in the housing market: non-minority households not impacted by racial covenants might have preferred living in areas with such restrictions, while simultaneously, these covenants limited and shaped the residential options available to minority households. This dual role highlights the importance of distinguishing between preferences and restrictions when examining how choice-set restrictions arising from either discrimination or neighborhood design policies influence segregation and spatial inequality. This is relevant not only in a historical context but also in relation to current debates about zoning regulations.

We disentangle the role of the drivers behind segregation both theoretically and empirically. First, we present reduced-form results indicating that racial covenants had a first-order effect on the sorting of households across space. Using an instrumental variable approach to account for the non-random placement of covenants over space, we find that living in a fully covenanted neighborhood decreased the probability of non-White households by 9.3 percentage points. Additionally, we show that the impact of racial covenants extended beyond racial minorities to include ethnic minorities, including White Eastern European immigrant families. Specifically, having a fully covenanted neighborhood reduced the White Eastern European immigrant households by 11.9 percentage points. These results highlight that *de jure* discrimination affected the segregation patterns of households in 1940. However, these reduced-form results cannot separate between the importance of preference for covenants by non-minority households and the function of covenants as restrictions in the neighborhood choice set of minority households.

Given the dual role that restrictions like racial covenants can play in sorting, we theoretically show the significance of accounting for differences in choice sets in residential choice models. Using a simplified discrete-choice model, we analyze how ignoring that certain demographic groups may only have access to a restricted choice set, as is standard in most residential choice models, can bias the estimation of preference parameters. Our theoretical findings indicate that the bias resulting from the mis-specified model is correlated with how the characteristics in the *latent choice set* for the restricted groups compare to the unrestricted full set of neighborhoods. Specifically, if the average characteristics in the restricted latent choice set are, on average, smaller (or larger) than the average across

all neighborhoods, then the estimate of the mis-specified model will have a downward (or upward) bias. Intuitively, if the practitioner mistakenly assumes that households facing choice set restrictions can choose from all available options, but in reality, they select from options with lower (or higher) values of a given characteristic, the resulting estimates would wrongly imply that the restricted demographic group sorts more negatively (positively) on this characteristic than under their true preferences. We further quantify this bias using Monte Carlo simulations.[1] Our results indicate that the difference between true and estimated parameters in a mis-specified model widens as the share of unavailable neighborhoods for the restricted demographic group increases.

We propose a two-sided matching housing market model where we model supply-side restrictions as differences in the neighborhood choice sets across different demographic groups following the literature on selective admissions and latent choice sets (Gandhi, 2019; Abaluck and Adams-Prassl, 2021; Agarwal and Somaini, 2022). Our two-sided matching housing market model of selective admissions not only allows us to include sources of observable discrimination, such as racial covenants, but also implicit restrictions that limit the choice sets available to minority groups, which are often difficult to measure. Concretely, on the supply side, we model real-estate developers who have preferences regarding the households they "accept" into their development according to an admission rule that depends on the household's race or ethnicity and captures the likelihood of a neighborhood being included in different demographic groups' choice sets. These supply-side preferences, in turn, generate differences in the household's latent choice sets, where these choice sets are explicitly linked to race or ethnicity. On the demand side, heterogeneous households from different demographic groups choose their most preferred development, partly based on their preferences over the presence of covenants in the development and the demographic composition of the development within their actual choice set.

We estimate our model parameters using a two-step approach following Berry et al. (1993) and Bayer et al. (2007) and extend their procedure to allow for the estimation of latent choice sets. In the first step, we recover the supply-side admission rule parameters and

[1] Crawford et al. (2021) perform a similar simulation exercise.

the household's preferences for development characteristics that are heterogeneous across demographic groups. In the second step, we recover the average preference parameters for the development characteristics. We estimate the model separately for demographic groups: *minority* and *non-minority*. We use our reduced-form results to classify households into these two groups.

The separate identification of household preference parameters and developer's admission-rule parameters in the first step of the estimation closely follows the arguments in Abaluck and Adams-Prassl (2021), which relies on the observation that the symmetry of the Slutsky equation fails when certain options are not in the agent's choice sets. Specifically, they demonstrate that having data on market shares and demand elasticities is sufficient to individually identify demand parameters from the parameters that govern the formation of the latent choice sets in the parametric model. The identification of the heterogeneous preference parameters, which is estimated in this first step, closely follows the approach of Bayer et al. (2007) and relies on individual (household-level) data and the differences in the final neighborhood match between minority and non-minority households.

In the second step, we face several endogeneity concerns. First, house prices and demographic composition within a development are equilibrium outcomes and, thus, can be functions of unobservable demand shocks. To address these endogeneity concerns, we construct two instruments: one based on steep land slopes that affect construction costs and, therefore, house prices (Saiz, 2010); and another using a shift-share migration instrument to instrument for the share of minorities within a development (Card, 2001; Boustan, 2010; Tabellini, 2020). Second, adding covenants is a decision of the developer and may also be influenced by unobservable location characteristics. To address this endogeneity concern, we use an urbanization frontier index from our reduced-form analysis as an instrument for the presence of covenants, which captures the availability of land for development in a radial outward expansion in 1940. Our relevance condition primarily relies on the correlation between the radial outward expansion of the city into virgin or agricultural land (Coeurdacier et al., 2022), conditional on the distance to the city center, and the growing popularity of covenants over time. Since racial covenants were primarily added to newly built housing, developments closer to the urbanized frontier of the

metropolitan area were more likely to have racial covenants in 1940 than those closer to the city center.

Preliminary results indicate that minority households faced substantial restrictions in choice sets: they were 26 percentage points less likely to have the option of living in the average development even after controlling for observable factors of discrimination, namely racial covenants. On the demand side, we also find that minority households are more sensitive to house prices compared with non-minority households and, thus, are more price elastic. Notably, when we estimate a model that ignores differences in neighborhood choice sets across demographic groups, we find that the willingness-to-pay for racial covenants is 60% larger compared to our two-sided matching model, overstating the role that household preferences play in the observed segregation.

Literature review: Our paper contributes to the literature on neighborhood segregation and the role of policy instruments in shaping the spatial distribution of households in several ways.

First, we contribute to the residential choice literature by considering both supply and demand side channels through which observed segregation can materialize. Typically, the existing literature only considers how preferences for co-patrons may shape within-city segregation (Bayer et al., 2007; Caetano and Maheshri, 2021; Almagro et al., 2024) due to the lack of data on the supply side. Two exceptions are Christensen and Timmins (2022) and Christensen and Timmins (2023), who measure discrimination by surveying landlords. We depart from their work in several ways. Firstly, we model and estimate the supply-side admission rules where the probability that a household is "accepted" in a neighborhood is a function of the household's demographics, which could arise from many implicit and explicit forms of discrimination beyond landlord discrimination. Secondly, we quantify the importance of preferences versus choice set restrictions on several equilibrium outcomes, such as house prices, segregation, and spatial inequality, allowing us to disentangle the role of demand and supply side channels behind discrimination.

Second, our paper contributes to a large literature measuring the role of how policy design contributes to segregation (Myers Jr, 1995; Rothstein, 2017; Trounstine, 2018; Fishback et al., 2020). We differ from previous studies in several ways. First, while most of the

literature measures localized treatment effects, often using regression discontinuity-based methods (Aaronson et al., 2021; Shertzer et al., 2016). Our study looks at the impact on the entire metro area and allows us to measure the aggregate importance of different channels that contribute to segregation. In addition, our novel data and model not only quantify the relative role of *de jure* discrimination via racial covenants, but also the role of other forms of discrimination that are implicitly captured by differences in admission rules across demographic groups along with household preferences.

Finally, to estimate differences in choice sets in the two-sided matching model, our methodology closely relates to the recent advances in the estimation of preference parameters with unobserved choice set heterogeneity (Gandhi, 2019; Agarwal and Somaini, 2022).[2] As opposed to previous papers that consider unobserved and heterogeneous choice sets, our paper is the first to estimate latent choice sets in the context of a housing market. Importantly, understanding how restrictions on choice sets are formed is key to understand how much of the observed segregation is explained by those restrictions as opposed to preferences on sorting and housing prices and understand their impact on spatial inequality and segregation.

2. Historical Background and Data

2.1 Historical Background

Racial covenants restricted racial and ethnic minorities from moving into covenanted neighborhoods during the first half of the twentieth century (see Appendix Figures D.1-D.3 for examples of racial covenants in property deeds). These covenants were private legal agreements between real estate developers and homeowners, making them a *de jure* form of segregation (Ellen and Steil, 2019). As an instrument of neighborhood design, racial covenants solved a collective action problem and prevented future changes in the neighborhood's demographic composition toward minority households (Tiebout, 1956; Schelling, 1969). In US metro areas that were being largely urbanized during the first half of the twentieth century, including the Minneapolis metro area, almost all racial covenants were added by developers at the stage of subdivision on new developments. Developers

[2]See Crawford et al. (2021) for an extensive survey on this topic.

would advertise racial covenants in their developments (see Appendix Figure D.4 for an example of such advertisements). The enforcement of racial covenants relied on "injured parties," usually neighbors, bringing lawsuits against those violating racial covenants as well as steering away of minorities by real estate agents and developers from covenanted neighborhoods.

Four key historical events established racial covenants as the primary tool for neighborhood design in urban and suburban areas of the US during the first half of the 20th century. First, in 1917, the *Buchanan v. Warley* US Supreme Court ruling prohibited municipalities from creating racially segregated neighborhoods through zoning regulations. As a result, the use of *private* racial covenants became more appealing (Lowe, 1948). Second, the prevalence of racial covenants increased when the National Association of Real Estate Boards amended its charter in 1924 to promote their use (Jones-Correa, 2000). Third, when the Supreme Court reaffirmed the legal enforceability of racial covenants in the 1926 *Corrigan v. Buckley* ruling, leading to their widespread adoption across the country (Crooks, 1948). Finally, the federal government solidified the use of racial covenants when, in its 1938 underwriting manual, the Federal Housing Administration (FHA) favored mortgages linked with "subdivision regulations and suitable restrictive covenants" (FHA, 1938). However, in 1948, the US Supreme Court declared racial covenants unenforceable in the *Shelley v. Kraemer* ruling, effectively ending the use of racial covenants to design segregated neighborhoods.

In addition to racial covenants, uncoordinated actions by developers, real estate agents, and neighbors likely contributed to steering households away from certain neighborhoods. There were also instances of collective group action, such as the threat of violence (Albright et al., 2021) and local government policies—like the placement of industrial zoning (Shertzer et al., 2016)—that influenced some developments to be welcoming while others were unwelcoming to racial and ethnic minorities in 1940. Our two sided matching model allows us to include sources of observable discrimination, such as racial covenants, but also these implicit restrictions that limited the choice sets available to minority groups.

2.2 Data

We combine data on racial covenants, historic developments, the historical street grid for Minneapolis metro area, and the 1940 US individual count decennial census. Linking all of these data sources together allow us to map racial covenants and households to the developments and geography of 1940. We focus on 1940 for two reasons. First, 1940 is when the use of racial covenants was at its peak. Second, 1940 is one of the two census years that contains data on house rents and value as well on the previous locations of households, which allow us to include household's moving costs in our residential choice which we eventually use as a demand-only shifter in the estimation of our two-sided matching model. More detailed information about how we construct our dataset can be found in Appendix A.

Data on Racial Covenants: The data on racial covenants comes from the Mapping Prejudice Project, which provides a complete census of all covenanted deeds and development maps in Hennepin County, Minnesota, where the Minneapolis metro area is located. In addition to the location of 24,131 racial covenants, Mapping Prejudice also provides information on the specific text of the racial and ethnic restrictions and the date of covenant execution.[3]

Historic Boundaries and Street Grids: We geocode and digitize enumeration district boundaries, which are similar to modern census tracts, for Hennepin County from 1900 to 1940. Unfortunately, linking households, developments and covenants to the present-day street day delivers less than 45% of successful matches as the historic street grid has changed in the past 80 years due to highway construction and new development over time. To overcome this problem, we digitize and geocode the 1940 street grids for Minneapolis and five neighboring suburban cities—St. Louis Park, Edina, Richfield, Hopkins, and Robbinsdale—using City Atlases and the Sanborn insurance maps that existed by 1940. Together with Minneapolis, these six cities form a contiguous part of the Minneapolis metro area and accounted for roughly 92% of the 1940 population of the

[3]Mapping Prejudice analyzed roughly three million pages of sales deeds and subdivision maps in Hennepin County, covering the period 1900-1960. Five people double-checked the existence, language, and date of the covenant execution for each deed and map, limiting the scope of error in data collection.

Hennepin County.

Data on Developments: Using the present-day development-level data from Hennepin County's assessor office misses out on developments that were destroyed or changed between 1940 and today, which translates into 29.7% of households in the sample not being matched to a development. To overcome this missing-development problem, we collect data on historic developments which come from the historic Atlases of the cities of Minneapolis (1940), St. Louis Park (1939), Edina (1940), Richfield (1941), Hopkins (1954), and Robbinsdale (1941). These data provide to-scale maps of each development in these cities (see Figure D.5 for an example of such Atlas map). We digitize and geocode the historic atlases to construct the 1940s geo-located geography of each development and locate the lots within these developments using Hennepin County's assessor data. In what follows, we use these developments as our definition of neighborhoods. Importantly, our data also contain information on the year that a development was originally platted and on the year the construction of the house was completed for most houses in our sample.

Individual-Count US Census Data: We use the full-count 1940 US decennial census data for the six Minneapolis metro area cities in our sample. The data include information about individuals like race, country of birth (to determine ethnicity), occupation, ownership status as well as residence five years ago. Since endogamy is common among races and ethnicities during this time period, we assume that the racial and ethnic characteristics of the head of household is the representative of the household, which we also confirm using the individual count data.[4] We do our analysis at the household level. There are 164,653 households in the six Minneapolis metro area cities in 1940.

The census data also contain other household information like the address, group quarter status, housing value, and rental price. We only observe rents and housing values for renter and owner-occupied households, respectively. While we conduct analysis separately for additional analysis, for the baseline, we impute the rental value of owner-occupied housing using the methodology laid out in Katz et al. (2017). The imputed rental value of owner-occupied housing combined with rental price data gives us monthly rental value for 91.5% of households in our sample. The remaining households lived in group

[4] 99.7% of of households in our sample marry within their race or ethnicity.

quarters, for which we observe no housing price data.

Other Data: We also digitize and geocode a 1899 map of Hennepin County, which highlights the locations of wetlands and other semi-submerged areas so we can measure the underlying land quality in the lake geography of Hennepin County. In addition, we obtain data on elevation and slope from the Natural Resources Conservation Service, which completed a comprehensive survey of Hennepin County in 1953. We also digitize a 1923 map highlighting the location of heavy and light industries as well as commercial zoning areas in the city of Minneapolis. Last, we also digitize the 1913 streetcar map of the Minneapolis metro area.

2.3 Construction of Geo-referenced Household Level Dataset

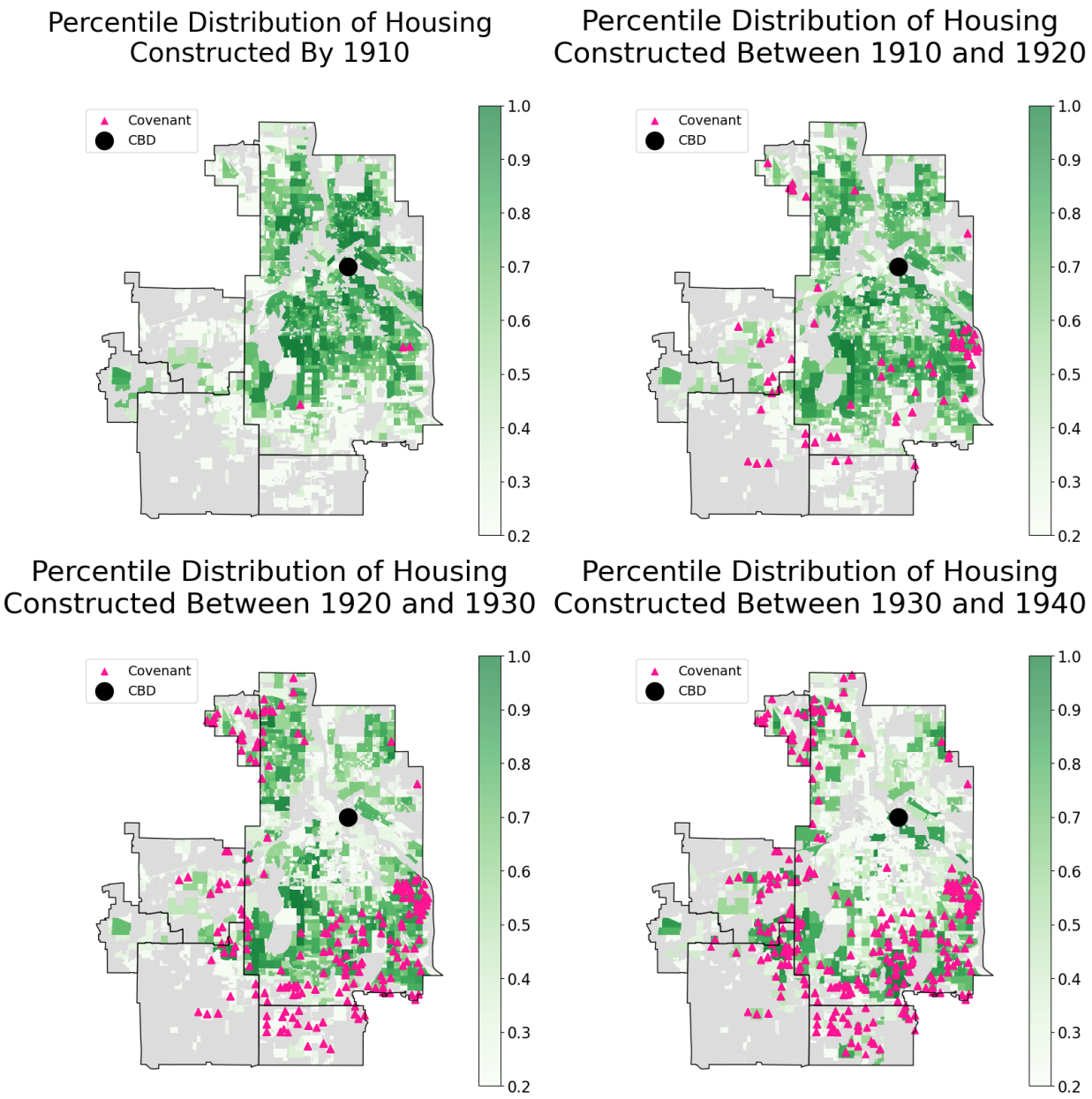
For our analysis, we need to link the households and racial covenants to geolocated developments in 1940.

We link 164,653 households to geolocated developments by matching the addresses from the census to the historical street grid map. In the 1940 census, we have complete addresses for 75.1% households. Unfortunately, the remaining 24.9% do not have a complete address in the census. For those households with an incomplete address, we impute missing street names and house numbers for 91% of missing addresses using the census enumeration ordering, as enumerators collected the census data on the same side of the street in order Logan and Parman (2017). Since this order is preserved in the census, we can impute missing house numbers and street names for most households. Adding those to the households with a full address, we can match 97.7% of the census households to a geolocated historic development using the historical street grid that we digitized.[5] Since we know the enumeration district for all households in the census, we place the remaining unmatched 2.3% of the households randomly across developments within the enumeration district.[6] Next, we link the information of racial covenants to the geolocated historic developments. Last, we also geolocate the presence of lakes and wetlands, the streetcar network, and the industrial and commercial zone areas, which allows us to measure the

[5]Recall that only 45% of the households can successfully be linked to an address using the present-day street grid.

[6]Later on in our empirical results, we confirm that our results are not driven by the random placement of these 2.3% of households.

Figure 1: Expansion of Housing Construction and Racial Covenants over Time



Note: This figure plots the distribution of housing construction in a given development in four periods–by 1910, 1911-20, 1921-30, 1931-40. Developments are color-coded by the percentile (between 0 and 1) based on the number of houses constructed in that period. The magenta triangles indicate addition of racial covenants in a development up to that time period. The city boundary of Minneapolis, St. Louis Park, Edina, Robbinsdale, Richfield, and Hopkins is marked in black. Grey areas represent developments with no housing construction by 1940. The black dot is the location of the central business district (CBD), which is where the St. Anthony Falls are located.

distance to these characteristics from a given development. See Appendix A for further details on linking across various datasets.

Figure 1 plots the distribution of new housing construction by development in four different periods–by 1910, 1911-20, 1921-30, 1931-40. Developments are color-coded in green by the percentile (between 0 and 1) based on the number of houses constructed in that period in that development. As the figure shows, the construction is expanding radially outwards over time from the central business district (CBD) indicated in the black dot. In addition, the figure plots the racial covenants up to that decade in magenta triangles. As the figure shows, the use of racial covenants continuously spread after 1910, the year of first racial covenant, and particularly after 1930.

3. Empirical Evidence of the Effects of Racial Covenants

In this section, we provide reduced-form evidence of the effects of racial covenants on the neighborhood ethnicity and racial composition in the 1940 Minneapolis metro area housing market. The reduced form results help us with the following three points. First, they demonstrate that the demographic groups impacted by racial covenants included all racial minorities as well as ethnic minorities. During the early to mid-20th century, this included White immigrant households from Eastern European countries. Second, these results help define whether the household belongs to the minority and non-minority groups. We use this classification to estimate structural parameters separately for each demographic group. Third, the findings motivate our structural analysis, as the reduced form results alone cannot distinguish between the preferences of non-minority households for covenants and the role of covenants as restrictions in the neighborhood choice set of minority households in creating the observed segregation.

3.1 Reduced-Form Model

To study the effects of racial covenants on the neighborhood ethnicity and racial composition, we estimate the following model, where i indicates a household, j indicates a development in 1940:

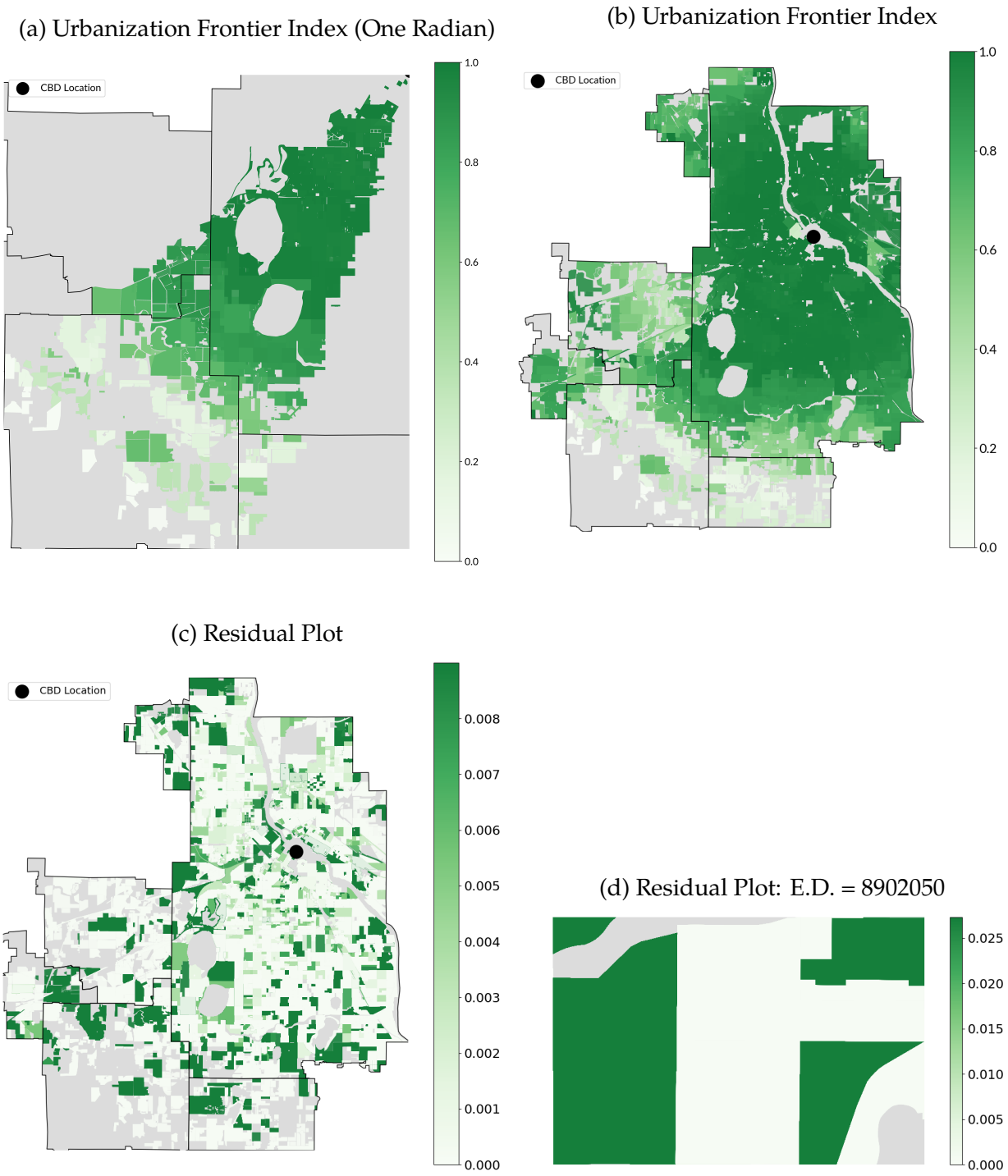
$$y_{ij} = \beta_0 + \beta_1 \text{Cov-Share}_{j(i)} + \beta_2 \text{Dist-CBD}_j + \beta_3 X_i + \beta_4 X_j + \eta_k(j) + \epsilon_{ij} \tag{1}$$

In this model, y is an indicator for a household's race and immigrant ethnicity in development j. Cov-Share$_{j(i)}$ is the share of the development j that had racial covenants. Dist-CBD$_j$ is the log of distance of the development j to the metro area's CBD. X_i is household's ownership status, an indicator for occupation, marital status, and the number of members and working members of the household. X_j is development level location quality variables including the percent of wetlands, log of distance to heavy industry zoned area, log of distance to commercial zoned area, log of distance to nearest lake, and log of distance to the streetcar. $\eta_k(j)$ indicates the enumeration district k fixed effects within which development j is located.

Given that covenants were a decision made by the developer at the time of platting their development, it is important to account for the non-random placement of racial covenants across space. Moreover, Sood and Ehrman-Solberg (2023) show that racial covenants were more likely to be added in lower-middle and middle-income areas and in neighborhoods that were more likely to face demographic transitions in the future. To control for this source of endogeneity, we include enumeration district fixed effects and use variation across developments within the same enumeration districts k to estimate our parameters.

Even within enumeration districts, the location of covenants could be correlated with time-varying unobservable characteristics. To address this concern, we instrument the share of covenanted development (Cov-Share$_{j(i)}$) with the share of built up area in the previous time period in developments that are neighboring to a given development j and are closer to the city center. We denote this share as f_j and call it *the urbanization frontier index*. Our relevance condition is based on the radial outward expansion of the city as seen in Figure 1. Thus, given that racial covenants gained popularity over time during the first half of the 20th century, as highlighted in Section 2, and were mostly added on new developments, our hypothesis suggests that developments closer to the urbanized frontier were more likely to have racial covenants than those away from the urbanization frontier. In addition, since racial covenants exhibit economies of scale as their discriminatory effects should increase in the number of covenants in the vicinity, developers could have had more incentives to add covenants in areas that had lower past levels of development and thus had more potential for these economies of scale to kick in.

Figure 2: Urbanization Frontier and Residuals



Note: Figure (a) plots the urbanization frontier index (between 0 and 1) for each development within a given radian of size 36°C. Figure (b) plots this same index for the developments in the entire metro area. Figures (c) and (d) plots residuals from regressing the urbanization frontier index instrument on the set of controls from Equation 2 and enumeration district (E.D.) fixed effects. Figures (d) plots the residual for one E.D. The controls include log of distance, log of slope, percent of wetlands, log of distance to heavy industry area, and log of distance to commercial zoning area. The city boundary of Minneapolis, St. Louis Park, Edina, Robbinsdale, Richfield, and Hopkins is marked in black and central business district (CBD) is indicated in a black dot.

To construct the urbanization frontier index, we first divide the city from the CBD into r radians of equal size. Within a radian r, we define $f_j \equiv \frac{\text{Built}_{r,1930}(z_j)}{\text{Lots}_{r,1930}(z_j)}$, where $\text{Built}_{r,1930}(z_j)$ is the number of houses built by 1930, and $\text{Lots}_{r,1930}(z_j)$ is the total number of buildable lots by 1930 within a given buffer zone z_j. A buffer zone for development j is within the same radian r but on the side that is closer to the city center and represents the neighboring developments for the development j. Note that the buffer zone is defined separately for each development j. Our preferred specification sets $r = 10$ and uses buffer size $z_j = 1000$ meters. Figure 2a plots the urbanization frontier index for one radian segment and Figure 2b plots the urbanization frontier index for all developments in Minneapolis metro area where the time period is $1930 - 1940$.[7]

The exclusion restriction is that conditional on enumeration district fixed effects, distance to the CBD, and location controls for the developments, the exact location of a development and the share of built within a development in the previous time period (by 1930) was random. Thus, the changes in the individual fortunes of developers and builders that had purchased lots from farmers and other landowners provides the underlying random variation.[8] To see the underlying variation that our identification strategy is capturing, we plot the residuals from regressing the urbanization frontier index instrument, f_j, on distance to the CBD, location controls, and enumeration district fixed effects. Figure 2c plots the residuals for all developments in Minneapolis metro area and Figure 2d plots the residuals for one enumeration district. As can be seen from both these figures, there is random variation across developments even within an enumeration district.

The first stage is given by the following model:

$$\text{Cov-Share}_{j(i)} = \alpha_0 + \alpha_1 f_j + \alpha_2 \text{Dist-CBD}_j + \alpha_3 X_j + \eta_{k(j)} + \varepsilon_{j(i)}. \tag{2}$$

Observe that this instrument not only exploits temporal variation in the addition of racial

[7]We deliberately use a ten-year gap in the construction of our instrument to allow for the time housing construction can take.

[8]The commercial and residential development in Minneapolis metro area until mid-1910s was focused very close the CBD. As the large milling large milling district starting collapsing by early 1920s, the commercial and residential development started moving away from the CBD (Danbom, 2003). Using survey data from the 1960s and 1970s from the neighboring Olmsted County, Minnesota, Pyle (1985) finds the most common reason for farm sales for housing development in the area is idiosyncratic personal shocks related to health and finances.

covenants, it also ensures that development demographic differences within an enumeration district operate only through the addition of racial covenants on the frontier of the new development.

3.2 Reduced-Form Results

Table 1 reports the coefficients α_1 and β_1 from Equations 2 and 1. The first-stage estimate indicate that our instrument works as expected and the higher the urbanized built-up area in a buffer zone z_j, the higher the probability of adding racial covenants to that development j. In particular, if the urbanization frontier index goes from 0 to 1, the probability of a development being covenanted increases by 0.18 percentage points (column 1). We also find that having a fully covenanted development (Cov-Share$_{j(i)}$ goes from 0 to 1) reduced the probability of non-White households by 9.8 percentage points (column 2). Non-White households include Black, Asian, Hispanic, and Native American households as classified in the census, regardless of the country of birth.

However, in the early to mid-20th century, the groups of households facing discrimination extended beyond just racial minorities. To explore this issue, we refer to Homer Hoyt's 1933 list of immigrants, where "the most favorable come first in the list and those exerting the most detrimental effect appear last" (Hoyt, 2000). Based on Hoyt's classification, we can categorize White immigrants, classified as those born in a country other than the US, into one of six groups, ordered from "most favorable" to "least favorable": Commonwealth (UK, Canada, Australia, and New Zealand) and Western Europe as most "favorable", Northern Europe and Southern Europe as medium "favorable", and Eastern Europe and Others (which includes immigrants from all other regions) as least "favorable".[9] Having a fully covenanted development increased the likelihood of first-generation White immigrant households from Western Europe by 5.1 percentage points (see column 6). This suggests that White immigrant households from Western Europe were not facing discrimination; rather, they were preferred as suggested by rank of one in Hoyt's list. We did not find any statistically significant effects for White immigrant households from the Commonwealth countries, Northern or Southern Europe. However, having a

[9]See Appendix A.1, where we provide Homer Hoyt's list and show which countries are in each of these categories.

fully covenanted development decreased the likelihood of Eastern European immigrant households, who were placed near the bottom of Hoyt's list, by 8.3 percentage points (see column 8). Thus, racial covenants were not only used as a discriminatory tool against racial minorities, but also to promote certain White ethnic minorities, which during the early-to-mid 20th century meant favoring White Western European immigrant households over White households from many other countries.

The effects of covenants on racial and ethnic sorting are not driven by the head of household's occupation, their ownership or marital status, or the number of members and working members in the household. As can be seen in columns 3, 7, and 9, controlling for these householder characteristics provides statistically similar effects and even increases the statistical significance in the case of White Eastern European immigrant households. In addition, note that the OLS coefficient for all demographic groups is significantly lower than the IV estimates (Table B.3), indicating a downward bias in the OLS estimation. The downward bias in the estimation aligns with the hypothesis that racial covenants were located in transition, low-middle-income, and middle-income regions that were likely of lower to middle-quality level within the larger metropolitan area.

The baseline results presented above are based on radians $r = 10$, $z_j = 1000$ meters, and the number of houses built by 1930 as a measure to construct the urbanization index. However, the results remain robust when we consider alternative time periods, such as the number of houses built by 1925 or 1910 (see Table B.1, Panels A and B), as well as varying buffer zones that decrease as the distance to the CBD increases $z_j = 2000 - 500$ meters (Table B.1, Panel C). Furthermore, the results are robust even when we control for the mean build year for each development (see Table B.2, Panel A) and when we make alternative assumptions about matching a handful number of households to different developments (see TableB.2, Panel B). Additionally, we conduct a placebo test, which uses a neighbor buffer zone that faces away from the CBD, rather than towards it, as in the baseline analysis. As can be seen in Table B.2 Panel C, we do not find statistically significant effects under this placebo test.

Moving forward, based on the results from Table 1, we define minority households as those that are non-White, regardless of their birthplace, as well as White households

Table 1: Covenants and Sorting in the Housing Markets

	First-Stage	Non-White		White Northern Europe Imm.		White Western Europe Imm.		White Eastern Europe Imm.	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Share Covenanted		-0.098***	-0.093***	0.058	0.069	0.051**	0.052**	-0.083	-0.119*
		(0.034)	(0.036)	(0.124)	(0.121)	(0.024)	(0.024)	(0.069)	(0.070)
Share Urban Index	0.180***								
	(0.021)								
Log Distance CBD	✓	✓	✓	✓	✓	✓	✓	✓	✓
Location Controls	✓	✓	✓	✓	✓	✓	✓	✓	✓
Household Controls			✓		✓		✓		✓
Enumeration District FE	✓	✓	✓	✓	✓	✓	✓	✓	✓
N	163,538	163,538	163,535	163,538	163,535	163,538	163,535	163,538	163,535
First-Stage F-Stat	62.32	73.04	72.02	73.04	72.02	73.04	72.02	73.04	72.02
Mean Dep. Variable	0.039	0.009	0.009	0.120	0.120	0.006	0.006	0.054	0.054

Note: This table presents results from Equations 2 and 1 with robust standard errors. Non-White is an indicator for non-White household (Black, Asian, Hispanic, or Native American regardless of country of birth). White Northern Europe Immigrants (Imm.), Western and Eastern Europe Immigrants is an indicator variable for White households born in these regions. See Appendix A.1, where we provide the list of which countries are in each of these categories. Share Covenanted represents the share of racial covenants in a development. Share Urban Index refers to the urban frontier index instrument. Log Distance CBD is the log of the distance in meters to Central Business District (CBD). Location controls include log of slope, percent of wetlands, log of distance to heavy industry zoning area, log of distance to commercial zoning area, log of distance to the nearest lake, and log of distance to the nearest streetcar. Household controls include household's ownership status, indicator for occupation, marital status, and the number of members and working members of the household. Enumeration District FE refers to fixed effects. Mean dependent (Dep.) variables provide means for indicator variables. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

born in Eastern Europe and "Other" countries. All other households are classified as non-minority households.[10] Under this classification, minority households account for

[10] Although we do not find statistically significant negative effects of racial covenants on White households

7.3% of households and 8.2% of the population in the Minneapolis metro area in 1940.

4. Model

4.1 Ignoring Restricted Choice Sets

We first show how ignoring restrictions on the household's choice sets can bias household preference estimates. To do so, we rely on a simplified economy that is populated with a minority group, whose households face restrictions in their choice sets and can only choose to live in a restricted set of neighborhoods.

We start by proving a general result that holds under mild assumptions and helps us quantify the sign of the bias of a mis-specified residential choice model. Concretely, consider the following model:

$$V_j^{m,i} = \beta^m \log \text{cov}_j + \epsilon_j^i$$

where $V_j^{m,i}$, is the utility of type-m (minority) household i gets from living in neighborhood j. Utility depends on neighborhood j's presence of covenants cov and an idiosyncratic component that is drawn from Type-I extreme value distribution ϵ_j^i. Each household chooses the neighborhood that gives her the highest positive utility. We denote $\mathcal{J}$ as the set of all possible neighborhoods. We denote the choice set of minority households (the latent choice set) as $\mathcal{J}_0$. $\mathcal{J}_1$ is the set of restricted neighborhoods that are not available to the minority group, so $\mathcal{J} = \mathcal{J}_0 \cup \mathcal{J}_1$. Since non-minority households are not face any restrictions in their neighborhood choices, their choice set is always $\mathcal{J}$.

We use maximum likelihood (MLE) to estimate β^m. If the choice set of the minority group is correctly specified, given our distributional assumption on idiosyncratic shocks, the probability of choosing neighborhood j is:

$$\mathbb{P}_j^m(\beta^m) = \frac{\exp\left(\beta^m \log \text{cov}_j\right)}{\sum_{l \in \mathcal{J}_0} \exp\left(\beta^m \log p_l\right)}.$$

born in "Other" countries of the world, we believe this is due to the small number of such households. In our baseline structural estimation, these households are included in the minority category.

The log-likelihood of the minority group is given by:

$$L^m(\beta^m) = \sum_{j \in \mathcal{J}_0} N_j^m \log \mathbb{P}_j^m(\beta),$$

where N_j^m is the number of minority households living in the j neighborhood in the data. Denote $\hat{\beta}^m \equiv \arg\max L^m(\beta)$ as the solution of the correctly specified choice set.

On the other hand, if the econometrician wrongly assumes that the choice set for minority households is given by $\mathcal{J}$ instead of $\mathcal{J}_0$, the minority m choice probabilities are given by:

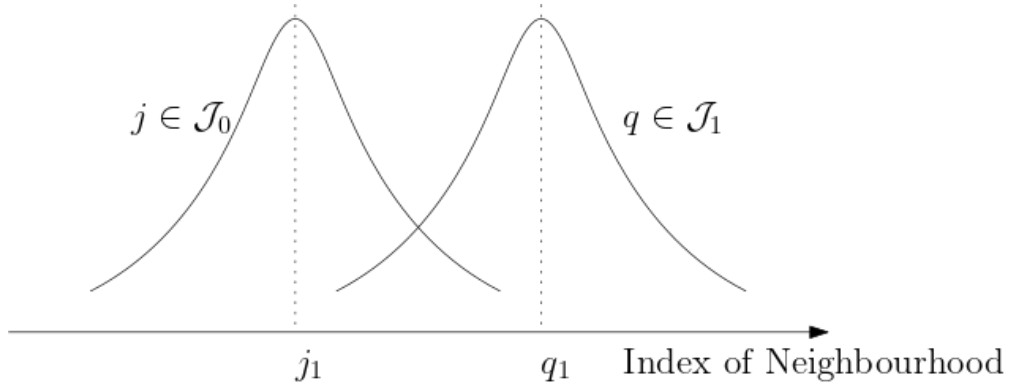
$$\tilde{\mathbb{P}}_j^m(\beta^m) = \frac{\exp\left(\beta^m \log \text{cov}_j\right)}{\sum_{l \in \mathcal{J}} \exp\left(\beta^m \log p_l\right)}.$$

In the case where the choice set is wrongly given by $\mathcal{J}$, the log-likelihood is given by:

$$\tilde{L}^m(\beta^m) = \sum_{j \in \mathcal{J}} N_j^m \log \tilde{\mathbb{P}}_j(\beta^m) = \sum_{j \in \mathcal{J}_0} N_j^m \log \tilde{\mathbb{P}}_j(\beta^m),$$

where the second equality follows as $N_j^m = 0$ for all $j \in \mathcal{J}_1$. Denote $\tilde{\beta^m} \equiv \arg\max \tilde{L}^m(\beta^m)$ as the solution of the problem with the mis-specified choice set.

Figure 3: Overlapping Covenants Distributions



Next, we assume differences in the characteristics of neighborhoods, namely covenants, across choice sets. Motivated by our prior evidence, we assume that neighborhoods that are not in the choice set of minorities, $\mathcal{J}_1$, have more covenants, on average:

Assumption 1. *Neighborhoods in the choice set for minorities have less covenants, on average:*

$$\frac{1}{N_0}\sum_{j\in\mathcal{J}_0} cov_j \leq \frac{1}{N_1}\sum_{j\in\mathcal{J}_1} cov_j,$$

where N_0 and N_1 are the number of houses in $\mathcal{J}_0$ and $\mathcal{J}_1$, respectively.

We allow for the range of the conditional covenants distributions to have a non-empty intersection, namely that they do not have a disjoint support (see Figure 3), but this overlap cannot be too large. Concretely, let's denote $[j_1, q_1]$ the interval where the range of the distributions for $\mathcal{J}_0$ and $\mathcal{J}_1$ intersect. Namely, $\text{cov}_j \leq \text{cov}_{q_1}$, $\forall j \in \mathcal{J}_0$ and $\text{cov}_{j'} \geq \text{cov}_{j_1}$, $\forall j' \in \mathcal{J}_1$. We assume that the mass of houses in the restricted choice set $\mathcal{J}_1$ inside $[j_1, q_1]$ is smaller than the mass of houses in the set $\mathcal{J}$, inside $[j_1, q_1]$:

Assumption 2. *Let N_0^s and N_1^s denote the number of observations of the distribution of covenants inside the range $[j_1, q_1]$, conditional on choice sets $\mathcal{J}_0$ and $\mathcal{J}_1$, respectively. We assume that $\frac{N_0^s}{N_0} \geq \frac{N_1^s}{N_1}$.*

Denote the minimum and maximum values of covenants in the constrained choice set $\mathcal{J}_1$ as cov_q^{min} and cov_q^{max}, respectively. Denote the minimum and maximum values of covenants in the latent choice set $\mathcal{J}_0$ as cov_j^{min} and cov_j^{max}, respectively. Define $\underline{\beta} \equiv \frac{\log \frac{N_0^s}{N_0}\frac{N_1}{N_1^s}}{\log \frac{\text{cov}_{q_1}}{\text{cov}_q^{min}}\frac{\text{cov}_j^{max}}{\text{cov}_j^{min}}}$, $\overline{\beta} \equiv \frac{\log \frac{N_0^s}{N_0}\frac{N_1}{N_1^s}}{\log \frac{\text{cov}_q^{min}}{\text{cov}_q^{max}}\frac{\text{cov}_j^{min}}{\text{cov}_{j_1}}}$. It is easy to see that under assumption 2, $\underline{\beta} < 0$ and $\overline{\beta} > 0$.

Theorem 1. *For $\beta \in (\underline{\beta}, \overline{\beta})$, under assumptions 1 and 2, $\tilde{\beta}^m < \hat{\beta}^m$. In other words, with mis-specified choice sets, the MLE estimate of the covenant sensitivity for the minority group has a downward bias.*

Proof. See proof in Appendix C.1. □

The intuition for the result above is straightforward. If the econometrician wrongly assumes that minority households can choose from all neighborhoods but only observes choices from a restricted choice set whose conditional mean is lower, the MLE estimator will wrongly infer that minority households dislike covenants more than they actually do, hence, producing an estimate of β has a downward bias.

Finally, we obtain the opposite result if we instead assume that neighborhoods in the choice set of minorities have *more* covenants:

Theorem 2. *Assume that houses in the choice set for minorities have more covenants on average:*

$$\frac{1}{N_0}\sum_{j\in\mathcal{J}_0} cov_j \geq \frac{1}{N_1}\sum_{j\in\mathcal{J}_1} cov_j,$$

and that the mass of the overlap between distributions is not too large:

$$\frac{N_0^s}{N_0} \leq \frac{N_1^s}{N_1}.$$

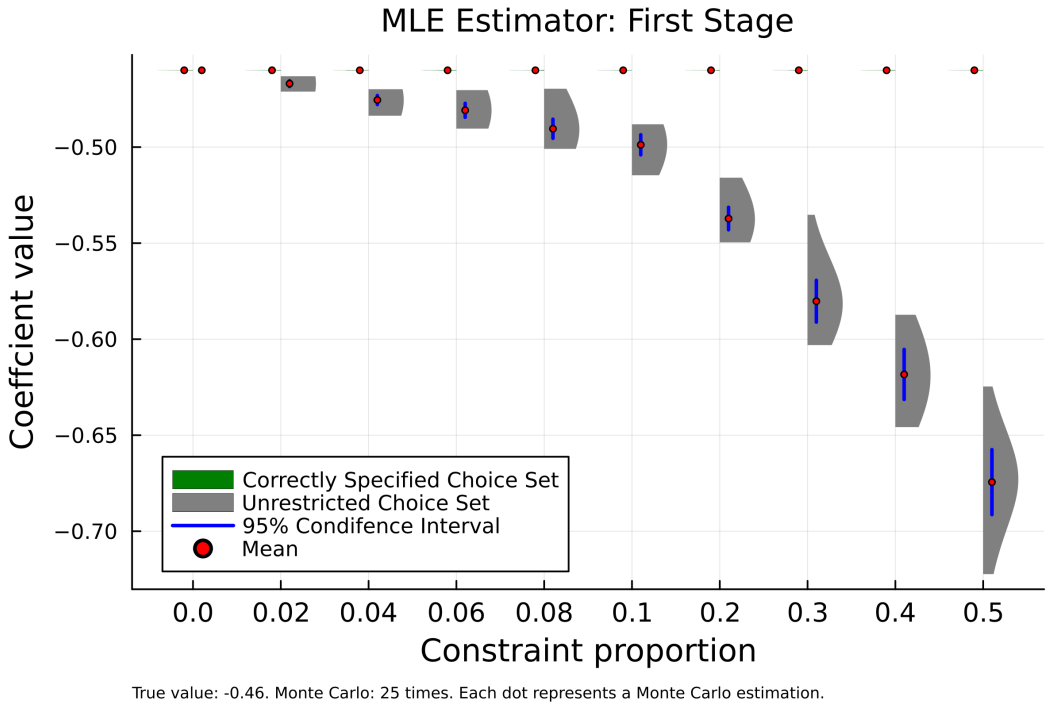
Then, for $\beta \in (\underline{\beta}, \overline{\beta})$, it follows that $\tilde{\beta}^m > \hat{\beta}^m$. In other words, under mis-specified choice sets, the MLE estimate of the covenant sensitivity for the minority group will have an upward bias.

4.1.1 Monte Carlo Simulations

In this section, we quantify how the size of the bias in the mis-specified model increases as the number of neighborhoods that are restricted to the minority households increases. To do so, we rely on Monte Carlo simulations designed as follows. First, we simulate a residential choice model with two groups that have heterogeneous preferences for housing characteristics. Recall that minority households face restrictions in their choice sets and can only choose to live in a restricted set of neighborhoods. Non-minority households do not face any restrictions. In the neighborhoods that are accessible to minority households, we simulate that presence of covenants is 50% lower, on average. For each data draw, we estimate coefficients for minority households for a mis-specified model with an unrestricted choice set and a correctly specified model with the correct choice set.

Figure 4 plots the distribution of estimates for minority m households. When the true value of $\beta^m = -0.46$, the unrestricted model estimates (in grey) are lower than the true value, i.e., the mis-specified unrestricted model delivers estimates with a downward bias, which will in turn imply that minority households have a lower willingness-to-pay for covenants than they actually do. The intuition for this downward bias is as follows. Recall that in the places where there is no restrictions, houses are less likely to have covenants,

Figure 4: True and Estimated Value of Preferences for Covenants



Note: This figure plot the distribution of estimates for minority households m. Mean and 95% confidence intervals are depicted in blue. The true value of $\beta^m = -0.46$. The vertical axis plots the coefficient values while the horizontal axis plots the share of developments that are restricted to the minority households.

on average. Therefore, if we wrongly specify the choice sets and assume that minority households can choose to live everywhere but we only see them choosing places with fewer covenants, our estimator would incorrectly infer that minority households are very negatively affected by this characteristic. Additionally, note that the difference between the true and estimated β^m increases as we increase the number of restricted neighborhoods for minority households. Also note that the bias disappears under the correctly specified model (in green) where the restrictions in the choice set for minority households are known.

4.2 A Two-Sided Matching Model of the Housing Market

Having established the importance of choice-set mis-specification, we now proceed to present a two-sided matching model with selective admissions of households into neighborhoods. Households in our model belong to different types, namely, minorities and

non-minorities, with heterogeneous preferences for neighborhood characteristics across types. On the supply side, we model developments owned by real-estate developers who admit households into their development based on an admission rule. Importantly, we assume that this admission rule is a function of household's race and ethnicity. This model of selective admissions introduces variation, and thus, potential restriction in the neighborhoods that can be accessed by different demographic groups.

4.2.1 Demand Model

Households, indexed by i, belong to two types k: minorities (m) and non-minorities (nm). In what follows, we fix the non-minority group as our baseline group and denote with indicator $m(i)$ where i belongs to the minority group.[11] The indirect utility for household i from residing in development j is given by:

$$U_{ij} = (\beta_x + \beta^m \mathbb{1}\{k(i) = m\})X_j^d + \xi_j + \epsilon_{ij} \tag{3}$$

$$= \delta_j + \beta^{m(i)} X_j^d + \epsilon_{ij}, \tag{4}$$

where X_j^d are observable characteristics of development j that enter the demand function, ξ_j are unobservable demand shocks, $\beta^{m(i)} = \beta^m \mathbb{1}\{k(i) = m\}$, and $\delta_j = \beta X_j^d + \xi_j$. In other words, δ_j represents the average utility of the non-minority group and that β^m are just the deviations in preferences of the minority group relative to the non-minority group. The characteristics of development j that enter demand, X_j^d, include development housing price p_j, the share of houses in the development with covenants Cov-Share$_j$, and the share of minority households m_j. Thus, households have preferences not only on the presence of covenants in their neighborhood but also on the racial and ethnic makeup of their neighbors. Finally, ϵ_{ij} is an idiosyncratic shock associated with household i's choice for development j. If ϵ_{ij} is an i.i.d. type-I extreme value error, the choice probability for development $j \in O$ conditional on choice set O is given by:

[11] This model requires the definition of a baseline group. If not, observe that the combination;

$$\tilde{\delta}_j = \delta_j + cX_j, \qquad \tilde{\beta}^{nm} = \beta^{nm} + c, \qquad \tilde{\beta}^m = \beta^m + c,$$

gives observationally equivalent outcomes.

$$\mathbb{P}_j(\boldsymbol{\delta}, \boldsymbol{\beta}|O) = \frac{\exp(v_{ij}(\boldsymbol{\delta}, \boldsymbol{\beta}))}{1 + \sum_{j' \in O} \exp(v_{ij'}(\boldsymbol{\delta}, \boldsymbol{\beta}))},$$

where (δ, β) is the vector of demand parameters and $v_{ij}(\boldsymbol{\delta}, \boldsymbol{\beta}) \equiv \delta_j + \beta^{m(i)} X_j^d$. Importantly, $\mathbb{P}_j(\boldsymbol{\delta}, \boldsymbol{\beta}|O) = 0$ if $j \notin O$.

4.2.2 Development Admission Rule

We model discrimination and restricted choice sets as development admission rules, following Abaluck and Adams-Prassl (2021), Agarwal and Somaini (2022) and Gandhi (2019). Developer j admits household i according to the following rule:

$$\eta_i \geq \alpha X_j^s + \alpha^{m(i)} X_j^s, \tag{5}$$

where X_j^s are the development characteristics that enter in the admission rule and η_i is household's i unobservable idiosyncratic component related to her admissions. Note that this admission rule depends on household characteristics as we allow for the admission rule parameters α to differ across different demographic groups. We also assume that the outside option is always in the choice set of every household, regardless of demographic type.

From equation (5), we can see that if i is admitted into development j and $\alpha X_j^s + \alpha^{m(i)} X_j^s \geq \alpha X_{j'}^s + \alpha^{m(i)} X_{j'}^s$, then household i must also be admitted into development j'.[12] This transitivity result implies we only have to consider $J + 1$ potential choice sets that can be ordered as follows:[13]

$$\{0\}, \{0, 1\}, \ldots, \{0, 1, ..., J\},$$

[12]Observe that this transitivity results relies on the assumption that the unobservable component that enters into admission rules, η_i, is common across developments. This assumption may fail if households of a group are favored differently by different developments in a way that is unobservable to the econometrician. One example of this could be family ties between households and developers. While we can relax this assumption, it is helpful for computational reasons. However, we believe that this unobservable match-specific component does not play a big role given that our model already allows developments to have different preferences over households based on observable components, such as the race, ethnicity, income, occupation, number of household members, age profile of the household, and the characteristics of the development.

[13]Otherwise, we would have to consider the full power set of $\{0, 1, ..., J\}$, which has 2^J elements.

where we order developments according to:

$$\alpha X_J^s + \alpha^{m(i)} X_J^s \geq \alpha X_{J-1}^s + \alpha^{m(i)} X_{J-1}^s \geq ... \geq \alpha X_1^s + \alpha^{m(i)} X_1^s.$$

Intuitively, the larger j is, the larger $\alpha X_j^s + \alpha^{k(i)} X_j^s$ and the stricter the admission rule. Therefore, for household i the probability that the set $[j] \equiv \{0, 1, ..., j\}$ is her choice set, O, is given by:

$$\begin{aligned}\mathbb{P}(O = [j]|\boldsymbol{\alpha}) &= \mathbb{P}(\alpha X_{j+1}^s + \alpha^{m(i)} X_{j+1}^s > \eta_i \geq \alpha X_j^s + \alpha^{m(i)} X_j^s) \\ &= F_\eta\Big(\alpha X_{j+1}^s + \alpha^{m(i)} X_{j+1}^s\Big) - F_\eta\Big(\alpha X_j^s + \alpha^{m(i)} X_j^s\Big),\end{aligned}$$

where F_η is the cumulative distribution function of the idiosyncratic payoff error term η_i.

4.2.3 Match probabilities

Putting everything together, the probability of an observed match s_{ij} for individual i is given by:

$$\begin{aligned}\mathbb{P}(s_{ij} = 1|\boldsymbol{\delta}, \boldsymbol{\beta}, \boldsymbol{\alpha}) &= \sum_{O \in 2^{|J|}} \mathbb{P}(O|\boldsymbol{\alpha}) \mathbb{P}_j(\boldsymbol{\delta}, \boldsymbol{\beta}|O) \\ &= \sum_{n=1}^{J} \mathbb{P}([n]|\boldsymbol{\alpha}) \mathbb{P}_j(\boldsymbol{\delta}, \boldsymbol{\beta}|[n]) \\ &= \sum_{n \geq j}^{J} \mathbb{P}([n]|\boldsymbol{\alpha}) \mathbb{P}_j(\boldsymbol{\delta}, \boldsymbol{\beta}|[n]) \\ &= \sum_{n \geq j}^{J} \left(F_\eta\Big(\alpha X_{n+1}^s + \alpha^{m(i)} X_{n+1}^s\Big) - F_\eta\Big(\alpha X_n^s + \alpha^{m(i)} X_n^s\Big) \right) \frac{\exp(v_{ij}(\boldsymbol{\delta}, \boldsymbol{\beta}))}{\sum_{j'=1}^{n} \exp(v_{ij'}(\boldsymbol{\delta}, \boldsymbol{\beta}) + 1}, \end{aligned} \tag{6}$$

where the third equality follows from the fact that $j \notin [n]$ if $j > n$.

5. Estimation of the Two-sided Matching Model

5.1 Estimation Procedure

The estimation relies on a outer-loop (first step) and an inner-loop (second step) following Berry et al. (1993) and Bayer et al. (2007). We extend their procedure to allow for the estimation of latent choice sets. The estimation algorithm proceeds as follows:

Inner Loop For given (α, β), we construct the Maximum Likelihood function following these steps:

(a) Order developments for each $m = 0, 1$ according to admission rules :

$$\Pi_{m1}(\alpha) \leq ... \leq \Pi_{mJ}(\alpha),$$

where $\Pi_j(\alpha; m) \equiv \alpha X_j^s + \alpha^m X_j^s$.

(b) Given the previous ordering, for each choice set $[n]$, construct choice set probabilities:

$$\mathbb{P}(O = [n]|\alpha; m) = F_\eta\Big(\Pi_{n+1}(\alpha; m)\Big) - F_\eta\Big(\Pi_n(\alpha; m)\Big).$$

(c) Compute the probability of a match following a modified version of the contraction mapping in Berry et al. (1993) and Bayer et al. (2007) to recover δ_j:[14]

i. Start with a guess of δ_j, say δ_j^0.

ii. For every step, k, compute match probabilities for all minority and non-minority households given α, β, and δ_j^k:

$$\mathbb{S}_j(\alpha, \beta, \delta^k; m) = \sum_{n \geq j}^{J} \mathbb{P}(O = [n]|\alpha; m) \cdot \frac{\exp(v_{mj}(\beta, \delta_j^k))}{\sum_{j'=1}^{n} \exp(v_{mj'}(\beta, \delta_j^k)) + 1}.$$

iii. Compute the model-predicted match probability for development j by sum-

[14] Berry et al. (1993) show that for given observed shares $s_j \equiv \frac{1}{N}\sum_i s_{ij}$ and $\beta^{m(i)}$, there is a unique vector of δ_j's that makes choice probabilities equal to s_j and that these δ_j's can be recovered using a contraction mapping.

ming across groups:

$$\mathbb{S}_j(\alpha, \beta, \delta^k) = \sum_m \mathbb{S}_j(\alpha, \beta, \delta^k; m) \cdot \frac{N^m}{N}.$$

iv. Update δ_j^k using:

$$\delta_j^{k+1} = \delta_j^k + \log s_j - \log \mathbb{S}_j(\alpha, \beta, \delta^k).$$

v. Repeat until convergence.

(d) Given α, β, and the resulting $\delta_j(\alpha, \beta)$, compute the log-likelihood:

$$\mathcal{L}(\alpha, \beta) = \sum_i \sum_j s_{ij} \log \mathbb{S}_j(\alpha, \beta; m(i)).$$

Outer Loop For every evaluation of the outer loop, we proceed as follows:

(a) Maximize the log-likelihood function to estimate α, β and the resulting δ_j.

(b) We estimate the remaining θ^{nm} by GMM using the following moments:

$$\mathbb{E}[G(\alpha, \beta)] = \mathbb{E}[\xi_j \cdot z_j] = 0,$$

where z_j are either instruments for the endogenous variables — prices, covenants, and demographic composition — or the variables that are assumed to be orthogonal to ξ_j. Concretely, we maximize the following objective function:

$$\max_{\alpha,\beta} G(\alpha, \beta)' W G(\alpha, \beta),$$

where $G(\alpha, \beta)$ is the vector of empirical moments and W is a weighting matrix.

5.2 Identification

In this section, we explain the sources of variation that identify our parameters. We do so step by step. Recall that in the first step (inner loop) we estimate the parameters governing the admission rule α, the differences in preference parameters for minority households

β as well as the utility levels δ_j of the non-minority group. In the second step (outer loop) , after recovering δ_j in the first step, we identify the parameters of the non-minority households.

5.2.1 Identification of the First Step: Admission Rules versus Preferences

We now discuss the variation in the data that allows us to separately identify admission-rule parameters $\boldsymbol{\alpha}$ from demand parameters $\boldsymbol{\beta}$ in our model. However, to make progress, we first need to make several normalizations. First, as pointed out by Train (2009), the level of utility is not identified. Thus, we normalize the value of the outside option equal to zero ($\delta_j = 0$) for all demographic groups. Similarly, because the scale of a logit model is not identified, we also normalize the variance of idiosyncratic shocks ϵ_{ij} to be equal to $\frac{\pi^2}{3}$. On the other side of the model, it is also easy to see that we need to normalize the variance of η. We fix $\sigma_\eta^2 = 1$.

To show that our model is identified, we need to show that there's an injective mapping from the parameters to the data. In what follows, we show that such an injective mapping exists. We start with a simplified example where we only have one demographic group, with utility given by $v_{ij} = \beta X_j^d + \epsilon_{ij}$, and an admission rule that is increasing in X_j^s, a one-dimensional vector.[15] Under this monotonicity condition, we know that we can order developments according to increasing X_j^s:

$$X_1^s \leq \ldots \leq X_J^s.$$

Given this ordering, we know that:

$$\begin{aligned}
s_j &= \mathbb{P}(s_{ij} = 1 | \boldsymbol{\delta}, \boldsymbol{\beta}, \boldsymbol{\alpha}) \\
&= \sum_{n \geq j}^{J} \left(F_\eta\left(\alpha X_{n+1}^s\right) - F_\eta\left(\alpha X_n^s\right) \right) \frac{\exp(\beta X_j^d)}{\sum_{j'=1}^{n} \exp(\beta X_{j'}^d) + 1} \\
&= \sum_{n \geq j+1}^{J} \left(F_\eta\left(\alpha X_{n+1}^s\right) - F_\eta\left(\alpha X_n^s\right) \right) \frac{\exp(\beta X_j^d))}{\sum_{j'=1}^{n} \exp(\beta X_{j'}^d) + 1} + \left(F_\eta\left(\alpha X_{j+1}^s\right) - F_\eta\left(\alpha X_j^s\right) \right) \frac{\exp(\beta X_j^d)}{\sum_{j'=1}^{j} \exp(\beta X_{j'}^d) + 1}
\end{aligned}$$

[15] This assumption can be relaxed but we keep it for the ease of exposition. In the estimation methods, it allows us to deal with large J without having to list all subsets of the power set of $\mathcal{J} = \{0, \ldots, J\}$.

$$= s_{j+1} \exp\left(\beta(X_j^d - X_{j+1}^d)\right) + \left(F_\eta\left(\alpha X_{j+1}^s\right) - F_\eta\left(\alpha X_j^s\right)\right) \frac{\exp(\beta X_j^d)}{\sum_{j'=1}^{j} \exp(\beta X_{j'}^d) + 1}. \tag{7}$$

Observe that we can write (7) for all j, forming a system of $J+1$ equations in the observed shares of matches $\{s_j\}_{j=1}^J$:

$$\begin{aligned}
s_0 &= s_1 \exp(-\beta X_1^d) + F_\eta\left(\alpha X_1^s\right) \\
&\vdots \\
s_j &= s_{j+1} \exp\left(\beta(X_j^d - X_{j+1}^d)\right) + \left(F_\eta\left(\alpha X_{j+1}^s\right) - F_\eta\left(\alpha X_j^s\right)\right) \frac{\exp(\beta X_j^d)}{\sum_{j'=1}^{j} \exp(\beta X_{j'}^d) + 1} \\
&\vdots \\
s_J &= \left(1 - F_\eta\left(\alpha X_J^s\right)\right) \frac{\exp(\beta X_J^d)}{\sum_{j'=1}^{J} \exp(\beta X_{j'}^d) + 1}.
\end{aligned}$$

The system above shows that, under a specific functional form assumption on F_η, we can identify the parameters α and β, as we have a system of $J+1$ equations and two unknown variables. In this simplified model with two parameters, having only one location and the outside option is enough. Generalizing this argument, we can identify up to $J+1$ parameters using the previous system of equations.

The intuition behind the identification is as follows. In a logit model with full consideration, we have the independence of irrelevant alternatives (IIA) property. In other words, the relative probability of two neighborhoods is only a function of how those neighborhoods compare in their characteristics space. Formally, under full consideration:

$$\frac{s_j}{s_{j'}} = \exp(\alpha(X_j^d - X_{j'}^d)).$$

Observe that this function corresponds to the first summand of the components of the previous system of equations. Therefore, under the assumptions of this model, the relative shares of j and j' also move with X_k^d with $k \neq j, j'$ only in the absence of full consideration. Moreover, our model also implies that variation in X_j^d and $X_{j'}^d$ should move s_j relative to

s_{j+1} in a log-linear way. Therefore, any departure from that log-linear relationship should indicate lack of full consideration.[16].

If we have k demographic groups, we just expand our system of equations across demographic groups. For the baseline group, we have

$$
\begin{aligned}
s_0^0 &= s_1^0 \exp(-\delta_1) + F_\eta\left(\alpha X_1^s\right) \\
&\vdots \\
s_j^0 &= s_{j+1}^0 \exp\left(\delta_j - \delta_{j+1}\right) + \left(F_\eta\left(\alpha^k X_{j+1}^s\right) - F_\eta\left(\alpha^k X_j^s\right)\right)\frac{\exp(\delta_j)}{\sum_{j'=1}^{j}\exp(\delta_{j'}) + 1} \\
&\vdots \\
s_J^0 &= \left(1 - F_\eta\left(\alpha^k X_J^s\right)\right)\frac{\exp(\delta_j)}{\sum_{j'=1}^{J}\exp(\delta_{j'}) + 1},
\end{aligned}
$$

while for all k that is not the baseline group we have:

$$
\begin{aligned}
s_0^k &= s_1^k \exp(-\beta^k X_1^d) + F_\eta\left(\alpha^k X_1^s\right) \\
&\vdots \\
s_j^k &= s_{j+1}^k \exp\left(\beta^k (X_j^d - X_{j+1}^d)\right) + \left(F_\eta\left(\alpha^k X_{j+1}^s\right) - F_\eta\left(\alpha^k X_j^s\right)\right)\frac{\exp(\delta_j + \beta^k X_j^d)}{\sum_{j'=1}^{j}\exp(\delta_{j'} + \beta^k X_{j'}^d) + 1} \\
&\vdots \\
s_J^k &= \left(1 - F_\eta\left(\alpha^k X_J^s\right)\right)\frac{\exp(\delta_j + \beta^k X_J^d)}{\sum_{j'=1}^{J}\exp(\delta_{j'} + \beta^k X_{j'}^d) + 1},
\end{aligned}
$$

These equations together, constitute a $(J+1)\cdot K$ system of equations, where K is the number of demographic groups. That implies that we can identify up to $(J+1)\cdot K$ parameters.

[16] Agarwal and Somaini (2022) show that for a family of very general models that nests ours, a sufficient condition for identification is the existence of two shifters that independently move utility and choice sets, respectively.

Adding match-specific admission rule unobserved heterogeneity. Finally, we can write a similar system of equations when we add match-specific admission rule unobserved heterogeneity. Assume for simplicity our baseline model with only one group and one covariate entering in each side of the market. Assume that the admission rule is given by

$$\eta_{ij} \geq \alpha X_j^s.$$

In this more general model, we can write the probability of having choice set C as follows:

$$C = \prod_{j \in C} F_\eta(\alpha X_j^s) \prod_{j \notin C} (1 - F_\eta(\alpha X_j^s)),$$

where η_{ij} is i.i.d. distributed according to F_η. Following a similar argument as above, we have that for all j:

$$\begin{aligned} s_0 &= s_j \exp(-\beta(X_j^d)) + \sum_{\{C | j \notin C\}} \mathbb{P}(C) \frac{1}{1 + \sum_{k \in C} \exp(\beta X_k^d)} \\ &= s_j \exp(-\beta(X_j^d)) + \sum_{\{C | j \notin C\}} \prod_{j \in C} F_\eta(\alpha X_j^s) \prod_{j \notin C} (1 - F_\eta(\alpha X_j^s)) \frac{1}{1 + \sum_{k \in C} \exp(\beta X_k^d)}. \end{aligned}$$

Observe that under parametric assumptions of F_η, we have a system of equations J in two unknowns, α and β. It is easy to generalize this system of equations to any s_j and $s_{j'}$.

5.2.2 Identification of the Second Step: Baseline Group Preferences

To recover the preferences for the baseline group levels β we project δ_j on X_j^d. Note that the unobservable component of utility ξ_j enter additively in δ_j. Therefore, the identification assumptions required by OLS would give biased estimates if there are observable characteristics that could be potentially correlated with ξ_j. In our case, these characteristics are housing prices, the minority share, as well as the share of covenants in a development. To overcome this challenge, we construct moments based on instruments orthogonal to ξ_j.

First, we instrument house prices using the slope within development j, following Saiz (2010), who documents significant increase in construction costs as slope increases over 10

and 15 degrees. The identifying assumption is that the steep slope instrument shifts the supply by affecting construction costs and thus prices, but does not affect the demand for housing, especially in the relatively flat geography of Minneapolis metro area. Second, we instrument minority share m_j by constructing a shift-share IV à la **?**Boustan (2010); Tabellini (2020):

$$Z_{c,1940} = \frac{1}{\hat{P}_c, 1940} \sum_r w_{cr}(1900) O^{-M}_{r,1930-40} \tag{8}$$

where $\hat{P}_c, 1940$ is the predicted city c population; w_{cr} is the share of households from country r living in city in 1900; and $O^{-M}_{r,1930-40}$ is the sum of the yearly quota between 1930 and 1940 for region r specified by the US Immigration Acts of 1921 and 1924. Third, we instrument the share of covenants using our frontier urbanization index explained in Section 3.

5.3 Results

In this section, we present *preliminary* estimation results of the two-sided housing market model with selective admissions. Table 2 shows the preference parameters for both demand and admissions rules.

From a qualitative point of view, all parameters go in the direction that we expect. Notably, we obtain a smaller estimate of α^m. That $\alpha^m < \alpha^{nm}$ means that being a minority household translates into a penalty in the admissions threshold rule, and thus, will be less likely to be admitted inside the developments. Concretely, we can compute the difference in the admissions probabilities across development for minority and non-minority households. These differences are plotted on Figure 5. On average, we see that a minority household is on average 26% less likely to be admitted into a development.

Also, the estimate of α^{nm}_{cov} is significantly positive, compared to the close to zero estimate of α^m_{cov}. This positive estimate means developments with a higher covenant share are more likely to admit the non-minority households. This result reaffirms our reduced-form results that the highly covenanted developments give more priority to non-minority households when choosing their residents.

Moving on to the demand parameters we find that all households have a negative

Table 2: Two-step Maximum Likelihood Estimation Results

	(1) Price		(2) Price + Covenants Demand		(3) Price + Covenants Both	
Supply Side Parameters:						
α^m	0.074	(0.009)	-0.075	(0.017)	-1.389	(0.053)
α^{nm}	0.434	(0.002)	0.304	(0.003)	-0.575	(0.005)
α^m_{cov}					-0.001	(0.426)
α^{nm}_{cov}					0.497	(0.002)
Demand Side Parameters:						
β^m_{price}	-1.048	(0.015)	-0.988	(0.019)	-0.989	(0.015)
$\beta^m_{covenants}$			-0.452	(0.017)	1.667	(0.119)
β^{nm}_{price}	-0.543	(0.005)	-0.521	(0.005)	-0.527	(0.005)
$\beta^{nm}_{covenants}$			-0.273	(0.019)	2.423	(0.019)
Log-lik.	1465297.551		1465246.038		1465239.351	
	N_j	3356	N^{nm}	579969	N^m	30116

Note: Standard errors in parentheses. Minority definition: all non-White households + White first-generation immigrant households from Eastern Europe, Middle East, Asia, Africa, South America, and other Islands territories.

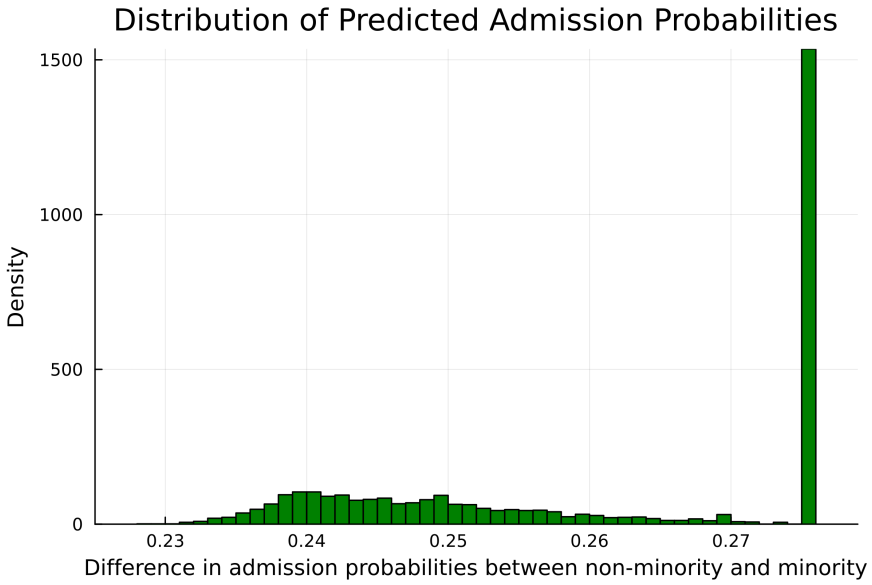
coefficient on prices. For all the specifications, we obtain a more negative β^m, meaning that minority households are slightly more price-sensitive than non-minority households. This result is in line to what previous studies have found (Bayer et al., 2007; Almagro et al., 2024).

We interpret the rest of the utility parameters on characteristics by computing the willingness-to-pay (WTP) across the demographic groups.[17] We find relative to non-minorities, minorities prefer to live in developments with a smaller covenant share. Moreover, for specification (2), where we only introduce demand side preference for develop-

[17]WTP for a 1% increase in a variable x that enters in logs or shares is computed as follows:

$$WTP^k_x = \frac{\alpha^x_p}{\alpha^k_p}.$$

Figure 5: Differences in Admission Probabilities between Non-Minority & Minority Households



ment's covenant share, we find both groups dislike living in developments with covenants. Minorities and non-minorities are willing to pay 45.7% and 52.4% higher rent to live in a development with 1% percentage point less of covenanted houses, respectively.

Finally, we compare our results to a standard approach that ignores the admission constraints and choice set restrictions as shown by Table 3. The first thing we note is that price coefficient is 18.6% larger in magnitude for minority households, and remains generally same for non-minority households, as indicated by Theorems 1 and 2. Ignoring the choice set restrictions affects the minority households' preferences more, since the penalty for minority households in the admissions threshold rule is captured in the minorities downward-biased preference. The preferences for covenants become negative. In light of our theoretical results, the change in sign is intuitive: a model that ignores choice set restrictions would predict that minoritiy households dislike more living in neighborhoods with higher shares of covenants, when in reality those covenants create restrictions that prevent them from moving to those locations.

Table 3: Estimation Results Comparison - With and Without Choice Set Restrictions

	Demand only				Demand and Supply			
	Price Only		Price + Covenants		Price Only		Price + Covenants	
Supply Side Parameters:								
α^{m}					0.074	(0.009)	-1.389	(0.053)
α^{nm}					0.434	(0.002)	-0.575	(0.005)
α^{m}_{cov}							-0.001	(0.426)
α^{nm}_{cov}							0.497	(0.002)
Demand Side Parameters:								
β^{m}_{price}	-1.271	(0.014)	-1.173	(0.015)	-1.048	(0.015)	-0.989	(0.015)
$\beta^{m}_{covenants}$			-0.617	(0.030)			1.667	(0.119)
β^{nm}_{price}	-0.538	(0.005)	-0.516	(0.005)	-0.543	(0.005)	-0.527	(0.005)
$\beta^{nm}_{covenants}$			-0.274	(0.019)			2.423	(0.019)
Log-lik.	1465707.642		1465611.658		1465297.551		1465239.351	
$N_j = 3356$			$N^{nm} = 579969$				$N^{m} = 30116$	

Note: Standard errors in parentheses. Minority definition: all non-White households + White first-generation immigrant households from Eastern Europe, Middle East, Asia, Africa, South America, and other Islands territories.

6. Conclusion

In this paper, we study the extent to which observed segregation is caused by restrictions or preferences. To do this, we have created a novel dataset that links historical census data with household choices and household characteristics with racial covenants, which were the primary *de jure* form of discrimination in the first half of the 20th century in the US. We provide evidence that racial covenants causally affected racial and ethnic sorting in the Minneapolis metro area in 1940. Motivated by our evidence that restrictions created by covenants did have their intended effect, we use a simplified model to show how ignoring choice set restrictions may severely bias preference parameters for neighborhood characteristics. Next, we propose an augmented two-sided version of the traditional discrete choice models of residential choice, where we model restrictions arising either

from implicit or explicit *de jure* discrimination as selective admissions on the supply side. We use our newly-collected to estimate our model finding that minority households had lower probabilities of having developments in the choice sets, even after controlling for the presence of covenants. We also find that ignoring restrictions on choices would have delivered significantly biased estimates for the share of covenants that for minority households, a result that aligns with our theoretical predictions. We take our results as suggestive evidence that ignoring implicit discrimination would have over-estimated the role that *de-jure* discrimination would have played in shaping the spatial distribution of minority households, and thus, segregation.

References

Aaronson, Daniel, Daniel Hartley, and Bhashkar Mazumder, "The effects of the 1930s HOLC "redlining" maps," *American Economic Journal: Economic Policy*, 2021, *13* (4), 355–392.

Abaluck, Jason and Abi Adams-Prassl, "What do consumers consider before they choose? Identification from asymmetric demand responses," *The Quarterly Journal of Economics*, 2021, *136* (3), 1611–1663.

Agarwal, Nikhil and Paulo J Somaini, "Demand Analysis under Latent Choice Constraints," Technical Report, National Bureau of Economic Research 2022.

Albright, Alex, Jeremy A Cook, James J Feigenbaum, Laura Kincaide, Jason Long, and Nathan Nunn, "After the burning: the economic effects of the 1921 Tulsa Race Massacre," Technical Report, National Bureau of Economic Research 2021.

Almagro, Milena, Eric Chyn, and Bryan A Stuart, "Neighborhood Revitalization and Inequality: Evidence from Chicago's Public Housing Demolitions," 2024.

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

_ , Marcus D Casey, W Ben McCartney, John Orellana-Li, and Calvin S Zhang, "Distinguishing Causes of Neighborhood Racial Change: A Nearest Neighbor Design," Technical Report, National Bureau of Economic Research 2022.

Berry, Steven T, James A Levinsohn, and Ariel Pakes, "Automobile prices in market equilibrium: Part I and II," 1993.

Boustan, Leah Platt, "Was postwar suburbanization "white flight"? Evidence from the black migration," *The Quarterly Journal of Economics*, 2010, *125* (1), 417–443.

Caetano, Gregorio and Vikram Maheshri, "A unified empirical framework to study segregation," Technical Report, Mimeo., University of Georgia 2021.

Card, David, "Immigrant inflows, native outflows, and the local labor market impacts of higher immigration," *Journal of Labor Economics*, 2001, *19* (1), 22–64.

Chetty, Raj, Nathaniel Hendren, and Lawrence F Katz, "The effects of exposure to better neighborhoods on children: New evidence from the moving to opportunity experiment," *American Economic Review*, 2016, *106* (4), 855–902.

Christensen, Peter and Christopher Timmins, "Sorting or steering: The effects of housing discrimination on neighborhood choice," *Journal of Political Economy*, 2022, *130* (8), 2110–2163.

__ and __ , "The damages and distortions from discrimination in the rental housing market," *The Quarterly Journal of Economics*, 2023, *138* (4), 2505–2557.

Chyn, Eric and Lawrence F Katz, "Neighborhoods matter: Assessing the evidence for place effects," *Journal of Economic Perspectives*, 2021, *35* (4), 197–222.

Coeurdacier, Nicholas, Florian Oswald, and Marc Teignier, "Structural Change, Land Use and Urban Expansion," 2022.

Collins, William J and Robert A Margo, "Residential segregation and socioeconomic outcomes: When did ghettos go bad?," *Economics Letters*, 2000, *69* (2), 239–243.

Crawford, Gregory S, Rachel Griffith, and Alessandro Iaria, "A survey of preference estimation with unobserved choice set heterogeneity," *Journal of Econometrics*, 2021, *222* (1), 4–43.

Crooks, James A, "Racial Covenant Cases," *The Georgetown Law Journal*, 1948, *37*, 514.

Danbom, David B, "Flour power: The significance of flour milling at the falls," *Minnesota History*, 2003, *58* (5/6), 269–285.

Durlauf, Steven N, "Neighborhood effects," *Handbook of regional and urban economics*, 2004, *4*, 2173–2242.

Ellen, Ingrid and Justin Steil, *The dream revisited: Contemporary debates about housing, segregation, and opportunity*, Columbia University Press, 2019.

FHA, "Underwriting Manual: Underwriting and Valuation Procedure Under Title II of the National Housing Act," 1938.

Fishback, Price V, Jessica LaVoice, Allison Shertzer, and Randall Walsh, "Race, risk, and the emergence of federal redlining," Technical Report, National Bureau of Economic Research 2020.

Gandhi, Ashvin, "Picking your patients: Selective admissions in the nursing home industry," *Available at SSRN 3613950*, 2019.

Hoyt, Homer, *One hundred years of land values in Chicago: The relationship of the growth of Chicago to the rise of its land values, 1830-1933*, Beard Books, 2000.

Jones-Correa, Michael, "The origins and diffusion of racial restrictive covenants," *Political Science Quarterly*, 2000, *115* (4), 541–568.

Jr, Samuel L Myers, "Racial discrimination in housing markets: Accounting for credit risk," *Social Science Quarterly*, 1995, pp. 543–561.

Katz, Arnold J et al., "Imputing Rents to Owner-Occupied Housing by Directly Modelling Their Distribution," *WP2017-7), BEA Working Paper*, 2017.

Logan, John, "Map USA, a human-mapping project (1940s–2010)," 2017.

Logan, Trevon D and John M Parman, "Segregation and homeownership in the early twentieth century," *American Economic Review*, 2017, *107* (5), 410–14.

Lowe, R Gordon, "Racial restrictive covenants," *Alabama Law Review*, 1948, *1*, 15.

Pyle, Lizbeth A, "The land market beyond the urban fringe," *Geographical Review*, 1985, pp. 32–43.

Rothstein, Richard, *The Color of Law: A Forgotten History of How Our Government Segregated America*, illustrated edition ed., New York ; London: Liveright, May 2017.

Saiz, Albert, "The geographic determinants of housing supply," *The Quarterly Journal of Economics*, 2010, *125* (3), 1253–1296.

Schelling, Thomas C, "Models of segregation," *The American Economic Review*, 1969, *59* (2).

Shertzer, Allison, Tate Twinam, and Randall P. Walsh, "Race, Ethnicity, and Discriminatory Zoning," *American Economic Journal: Applied Economics*, July 2016, *8* (3), 217–46.

Sood, Aradhya and Kevin Ehrman-Solberg, "Long Shadow of Housing Discrimination: Evidence from Racial Covenants," *SSRN*, 2023.

Tabellini, Marco, "Gifts of the immigrants, woes of the natives: Lessons from the age of mass migration," *The Review of Economic Studies*, 2020, *87* (1), 454–486.

Tiebout, Charles M, "A pure theory of local expenditures," *Journal of political economy*, 1956, *64.*

Train, Kenneth E, *Discrete choice methods with simulation*, Cambridge university press, 2009.

Trounstine, Jessica, *Segregation by design: Local politics and inequality in American cities*, Cambridge University Press, 2018.

Sorting by Choice? Separating Preferences from Discrimination in Housing Markets

by Milena Almagro and Aradhya Sood

ONLINE APPENDIX

A. Data Appendix

A.1 Census Microdata

The core individual-level census dataset focuses on Hennepin County, Minnesota, in the year 1940. The raw data files were obtained from IPUMS USA. The census data we obtain from IPUMS is full-count, i.e. there exists an observation for every individual surveyed by the Census Bureau in Hennepin County in 1940. The dataset contains fullcount data for all the census years within 1900-1940. We work on this entire dataset and then retain only the data corresponding to 1940, as that is the year of primary concern. Since we are mainly concerned with the housing market and location decisions, all measures that we use are household-level, determined by the variable for the head of the household. First, the raw data files were merged with enumeration district (ED) boundary files This resulted in a dataset of 561,528 individual records in 1940. Then, data processing included the cleaning and standardizing coding of key demographic variables related to race, educational attainment, home ownership, and house values. Additional variables relating to aggregate household composition were then generated, such as the number of families, couples, and multi-generational homes. We also incorporate census data on individuals' cities and counties of residence 5 years ago to measure internal migration.

Any missing variable data was inferred. Parental birthplaces were inferred based on other family members. Each individual's immigration status (first generation, second generation, etc.) was categorized based on their own, and their parents' birthplaces. Missing home values and rent amounts were imputed using the mean for that household's ED and year. All monetary variables were then harmonized to be consistent real values in 1940 dollar terms.

In order to classify the national background of respondents, we group individuals into

six categories based on their birthplace.

US: All states, Native American

Commonwealth: Scotland; Wales; United Kingdom; Ireland; Canada, Australia, New Zealand

Western Europe: Belgium; France; Liechtenstein; Luxembourg; Switzerland; Italy

Northern Europe: Denmark; Finland; Iceland; Norway; Sweden

Southern Europe: Albania; Greece; Portugal; Spain

Eastern Europe: Bulgaria; Austria; Hungary; Poland; Romania; Yugoslavia; Central Europe; Estonia; Latvia; Other USSR/Russia

Other Immigrants: Turkey; Syria; Israel/Palestine; Lebanon; China; Japan; Korea; Indonesia; Philippines; India; Asia Minor; Asia (unspecified); Africa (unspecified); Puerto Rico; Mexico; Central America; Cuba; West Indies; South America

In addition to the main census dataset, we utilize restricted-access ED geographic identifiers to link households to polygons.

A.2 Property Records

To analyze property-level characteristics, we utilize a dataset of all parcels in Hennepin County in 2019. Extensive cleaning was done to standardize land use types, street names, and address formats. We utilize the parcel data both for analyzing property-level characteristics (such as lot size, year built, development name, house built area, etc.) and for geocoding the census microdata. The parcels were also spatially joined to the ED boundary files.

A.2.1 Land Use Records

To measure proximity to industrial areas, we incorporate a shapefile of industrial-zoned land parcels. This dataset provides a 1940 land-use zoning map for the following areas only : St. Louis Park, Minneapolis, Edina, Richfield, Hopkins, and Robbinsdale. We are interested in distance to jobs from each development, so we calculate the distance to heavy and light industrial zones and commercial zones for each development. Distance from each residential parcel to the nearest industrial parcel was calculated.

A.3 Development Maps

We obtained development-level data on 6,800 separate developments located in the greater Minneapolis area of Hennepin County. Historical household addresses are assigned to a historical development built in 1940. The data is is first aggregated based on the 2019 Hennepin County Assesor dataset with relevant development information. We then use historical information including historical street grid, household addresses in the 1940 census, and the aforementioned covenant dataset to adjust the map into a historical version of the development map. From the current day parcel dataset, we group parcels with the same development. This generate 5,730 non-contiguous current-day developments in the six cities for which we have streetgrid data. Based on the historical streetgrid information, we eliminate all the gap between parcels within the same development and combine the nearby parcels into single polygons. Since many developments are still multipolygons (separated by other developments) even after reducing the gap between streets, we have separate them into contiguous polygons, so that a development may have multiple subdevelopments (e.g. developments 1A, 1B, 1C). All developments that have at least one parcel built before 1940 are included in the dataset. Other current day developments are also retained because houses are torn down and rebuilt and current day assessor data would only record the latest building year. Based on current day information, we are able to recover about 6000 continguous developments. In addition, there can also be cases where there used be parcels but buildings were torn down not rebuilt (e.g. houses torn down to built a highway). These geometries are manually created based on historical street grid, plat maps, or covenant geometry.

When imputing these historical graphs to fill in the remaining gaps, some historical developments were added with no overlapping of the current data. Many, however, were overwritten in this process. We prioritized historical development additions over those originally generated from the 2019 data, and thus, many of these development geometries were altered or replaced entirely to fit a more accurate historical mapping. This leaves us with a final total of 5,257 contiguous developments, where 4,986 can be traced back to a modern-day development (although they were historically retrofitted in the data, possibly including their geometry) and 271 completely new historical additions. Including the

industrial and infrastructural data discussed in previous sections, we finalize the dataset at this level by connecting it with wetland percentage, elevation, bedrock, and soil-related data.

A.4 Enumeration of District Boundaries

To analyze the data at fine geographic levels, we utilize census enumeration district (ED) boundary shapefiles for Hennepin County. These boundary files allow us to map the census microdata to spatial units for analysis. ED matching enables us to include ED fixed effects to conduct our main analysis at the subdivison level. Extensive data cleaning was performed to standardize the ED numbering systems across years and align the boundaries with other geographic datasets.

A.5 Street and Transportation Networks

Historical street network data is used to geocode addresses and analyze spatial connectivity: We utilize data generously provided by Brown University's Map USA Database Logan (2017) and supplement the data with manually collected streetgrid data for Minneapolis in 1940, and the available suburbs at that time. We note that streetgrid data is only available in six cities in Hennepin County: St. Louis Park, Minneapolis, Edina, Richfield, Hopkins, and Robbinsdale and proceed in adding those six cities to our Minneapolis data.

The street grid files were collected, digitized, and cleaned to extract street names, types (e.g. street, avenue, boulevard), and directional prefixes/suffixes. They also include interpolated address number ranges for geocoding.

A.6 Racial Covenants

Our key variable of interest is the prevalence of racially-restrictive covenants. We utilize a historical development-level dataset of all racial covenants executed in Hennepin County from 1910-1955, linked to 2019 parcel boundaries. We obtain parcel information from the 2019 Hennepin County assessor data. We match the covenant information to parcels within each developmemt and aim to recover share of covenants within each development. Parcels are tagged with a binary variable indicating the presence of a racial covenant, as well as the date of covenant execution.

There are 25818 covenant observations in the entire Hennepin County, of which we

identify 25165 unique covenant observations based on the following attributes: language used on covenants, addition, city, block, lot, grantor, and date of execution. Of those, there are 19,668 unique observations in the six cities we focus in this paper mentioned previously. Covenants are matched to parcels based on common attributes (Addition, City, Block, and Lot), and geographical location (spatial match). Around 90% of the covenants were matched with a relevant parcels. The remaining were matched manually (by GIS experts). Since the parcels and covenant can cover different geographic geometries, we allow for one parcel to be matched to multiple covenants if the parcel size is larger than the historical covenants, and we allow for multiple parcels to be matched to a single covenant if historically the covenant is larger than current parcel geography. After identifying which parcels have been covered under covenants and at which dates, we calculate the share of covenants in the developments at different time periods.

The covenants are then connected to our developments data to obtain parcel- and development-level information for each covenant in the six cities of focus within Hennepin County. Roughly 85% were matched via parcel-related attribute matches (also spatially, if there were multiple contiguous developments matched to a single covenant using the attributes). The remaining matches were completed with the assistance of a combination of location, historical development names, and parcel-level attributes.

A.7 Address Geocoding

To enable household-level analysis, we cleaned, geocoded, and matched 1940 households to historic (1940) developments. This process began with the raw input file of 1940 census addresses. It contained 180,424 household-level address records in Hennepin County with variables for the raw address string household ID. There were 164,729 raw addresses in six cities for which streetgrid datafiles were available. Extensive cleaning of the raw address strings was conducted in order to extract house numbers, directional prefixes, street names, and street types. This involved a combination of string parsing, pattern matching, and fuzzy text matching using the cleaned street name data as a reference. After this, missing address components (especially house numbers) were imputed based on adjacent records. The raw file contained 75.1% addresses with full addresses before cleaning (street addresses and house number), which improved to 96.1% coverage of full

addresses after cleaning. 95.6% of the data entries contained street address information before imputation, which improved to 96.9% after imputation. Meanwhile, 78.1% of entries contained house number information before imputation, which was improved to 98.3% after imputation.

Table A.1: Quality of Imputation

Address Information	Raw	%	After Imputation	%
Street Name	157,547	95.64	159,633	96.91
House Number	128,654	78.10	162,010	98.35
Full Addresses	123,768	75.13	158,355	96.13
Total	164,578	100.00	164,578	100.00

Note: Full Addresses means this address contains both street name and house number information.

The cleaned addresses were then geocoded by matching with the parcel and street grid files. Matches were attempted based on ED, house number, street name, street type, and direction. Both exact and fuzzy matching was used. 40.6% addresses were matched using this process. The rest were given to GIS experts to manually geocode using historical streetgrid mapping data. The output of the geocoding process is a dataset containing the original household id, the cleaned and geocoded address fields, and the matched parcel identifier. Of the 164,578 addresses in our final cleaned dataset, we were able to geocode 72,882 (44%) outselves prior to utilizing GIS experts.

The final step in the data construction process involved extensive spatial data processing to link all of the datasets together. Spatial joins between the parcel, ED boundary, street grid, industrial zoning, wetland geography, and covenant datasets were used to merge datasets. For each parcel, we calculated the percent of area overlapping each ED (for mapping between 1940 and 2019 geographies) and the distance to the nearest industrial-zoned parcel. Finally, data was aggregated from parcel-level data (land use, covenant status, etc.) to the development level based on the spatial crosswalks. The output is an development-level dataset for 1940 Hennepin County that contains, for each ED, the geometry defining the geographic boundary; the aggregate population and housing characteristics from the

1940 census microdata; parcel-level distributions of land use, property characteristics, and covenant prevalence; and spatial metrics like distance to industrial areas.

A.8 Summary Statistics for Developments and Enumeration Districts

Table A.2: Enumeration Districts (ED) vs Development Size

Percentile	Development Size (m^2)	ED Size (km^2)
25th	4,642	2.090
50th	14,023	5.853
75th	40,329	15.515
Mean	77,016	17.001
SD	223,801	21.832

Table A.3: Development Summary Statistics

Variables	P_{25}	P_{50}	P_{75}	Mean	SD
Number of Households	4	16	81	426	2,874
Minority Share	0.00	0.00	0.20	0.15	0.24
log impute hval4	3.987	4.731	5.223	4.624	1.040
Build Year	1933	1935	1937	1935	2.87
Commute Distance (m)	0	24.84	299.7	368.8	836.3
Distance to CBD (m)	4,813	6,824	8,673	6,791	2,474
Distance to Heavy Industry (m)	355.2	879.3	1,626	1,121	1,043
Distance to Lake (m)	313.2	660.4	1,077	735.4	526.6
Distance to Streetcar (m)	4.84	324.7	803.3	565.7	725.1
Wetland PE	0.00	0.00	0.00	3.97	14.52
Instrument (new construction)	0.450	0.473	0.490	0.467	0.064
Inverse Instrument (new construction)	0.455	0.476	0.491	0.476	0.072

Table A.4: Additional Summary Statistics

Variables	P_{25}	P_{50}	P_{75}	Mean	SD
Panel A : Development Level					
Number of Households	3	12	43	47.7	123.2
Minority Share	0.000	0.157	0.286	0.203	0.222
Share Non-White	0.000	0.000	0.000	0.011	0.076
Share Black	0.000	0.000	0.000	0.009	0.070
Share Western European	0.000	0.000	0.000	0.011	0.056
Share Northern European	0.000	0.089	0.200	0.137	0.184
Share Eastern European & Middle East	0.000	0.000	0.013	0.033	0.108
Share Southern European	0.000	0.000	0.000	0.002	0.023
Share Immigrant Other	0.000	0.000	0.016	0.020	0.076
Panel B : Enumeration Districts (ED) Level					
Share Non-White	0.000	0.000	0.003	0.009	0.040
Share Black	0.000	0.000	0.000	0.008	0.040
Share Western European	0.005	0.009	0.014	0.011	0.014
Share Northern European	0.070	0.122	0.191	0.137	0.080
Share Eastern European & Middle East	0.007	0.014	0.029	0.042	0.091
Share Southern European	0.000	0.000	0.004	0.003	0.004
Share Immigrant Other	0.010	0.016	0.024	0.018	0.013
Number of Developments	7.0	11.0	21.0	16.2	13.8

B. Empirical Evidence

B.1 Additional Reduced Form Evidence

Table B.1: Robustness Table 1

Panel A: 1925-1940			
	Non-White	White Western European	White Eastern European
Share Covenanted	-0.372**	0.090	-0.424*
	(0.151)	(0.083)	(0.240)
Observations	163,535	163,535	163,535
First-Stage F-Stat	17.65	17.65	17.65
Mean Dep. Variable	0.010	0.006	0.054
Panel B: Cumulative			
	Non-White	White Western European	White Eastern European
Share Covenanted	-0.093***	0.051**	-0.119*
	(0.036)	(0.024)	(0.070)
Observations	163,535	163,535	163,535
First-Stage F-Stat	72.02	72.02	72.02
Mean Dep. Variable	0.010	0.006	0.054
Panel C: Shrinking			
	Non-White	White Western European	White Eastern European
Share Covenanted	-0.163**	0.086**	-0.224*
	(0.068)	(0.043)	(0.125)
Observations	163,560	163,560	163,560
First-Stage F-Stat	25.65	25.65	25.65
Mean Dep. Variable	0.010	0.006	0.054

Note: Western European immigrants were born in Western Europe, similarly for Eastern Europe & Middle East immigrants. Share Covenanted represents the share of covenants in a development. Controls are inlcuded as follows: Log Distance CBD is the log of the distance in meters to Central Business District (CBD), which is at St. Anthony Falls. Location controls include altitude, log of slope, percent of wetlands, log of distance to heavy industry area, log of distance to commercial zoning area, log of distance to the nearest lake, and log of distance to the nearest streetcar. Household controls include household's ownership status, income imputed from occupation, marital status, and the number of members and working members of the household. Enumeration District FE refers to ED fixed effects. Mean dependent (Dep.) variables provide means for indicator variables and non-log values. Robust standard errors in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table B.2: Robustness Table 2

Panel A: Build Year Control			
	Non-White	White Western European	White Eastern European
Share Covenanted	-0.056	0.070**	-0.192*
	(0.035)	(0.034)	(0.098)
Observations	147,309	147,309	147,309
First-Stage F-Stat	41.43	41.43	41.43
Mean Dep. Variable	0.010	0.006	0.054
Panel B: Ranks 1 & 2			
	Non-White	White Western European	White Eastern European
Share Covenanted	-0.090**	0.051**	-0.119*
	(0.035)	(0.024)	(0.070)
Observations	163,535	163,535	163,535
First-Stage F-Stat	74.83	74.83	74.83
Mean Dep. Variable	0.010	0.006	0.054
Panel C: Placebo			
	Non-White	White Western European	White Eastern European
Share Covenanted	0.008	0.026	-0.155**
	(0.016)	(0.023)	(0.068)
Observations	163,908	163,908	163,908
First-Stage F-Stat	92.47	92.47	92.47
Mean Dep. Variable	0.010	0.006	0.054

Note: Western European immigrants were born in Western Europe, similarly for Eastern Europe & Middle East immigrants. Share Covenanted represents the share of covenants in a development. Controls are inlcuded as follows: Log Distance CBD is the log of the distance in meters to Central Business District (CBD), which is at St. Anthony Falls. Location controls include altitude, log of slope, percent of wetlands, log of distance to heavy industry area, log of distance to commercial zoning area, log of distance to the nearest lake, and log of distance to the nearest streetcar. Household controls include household's ownership status, income imputed from occupation, marital status, and the number of members and working members of the household. Enumeration District FE refers to ED fixed effects. Mean dependent (Dep.) variables provide means for indicator variables and non-log values. Robust standard errors in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table B.3: Robustness Table 3 - OLS

	First-Stage	Non-White	Northern Europe	Western Europe	Eastern Europe
Share Covenanted		-0.005***	-0.011***	-0.001	-0.021***
		(0.034)	(0.004)	(0.001)	(0.002)
Urban Index	-0.247***				
	(0.000)				
Observations	163,540	164,653	164,653	164,653	164,653
Mean Dep. Variable	0.120	0.010	0.120	0.006	0.054

Note: Western European immigrants were born in Western Europe, similarly for Eastern European & Middle Eastern immigrants. Share Covenanted represents the share of covenants in a development. Controls are included as follows: Log Distance CBD is the log of the distance in meters to the Central Business District (CBD), which is at St. Anthony Falls. Location controls include altitude, log of slope, percent of wetlands, log of distance to heavy industry area, log of distance to commercial zoning area, log of distance to the nearest lake, and log of distance to the nearest streetcar. Household controls include household's ownership status, income imputed from occupation, marital status, and the number of members and working members of the household. Enumeration District FE refers to ED fixed effects. Mean dependent (Dep.) variables provide means for indicator variables and non-log values. Robust standard errors in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

B.2 Shift-Share Instrument Construction

To address the potential endogeneity in minority household location decisions, we construct a modified version of the shift-share instrument following Boustan and Tabellini. The instrument combines three key components: predicted city growth rates from non-Minnesota metropolitan areas, historical migrant settlement patterns, and immigration quotas from the Immigration Act of 1924. We use the 1930 and 1940 IPUMS ACS censuses to construct our instrument. First, we calculate the predicted 1940 population for each of the six cities in our sample, c, using the average growth rate of non-Minnesota metropolitan areas V between 1930 and 1940:

$$\hat{P}_{c,1940} = P_{c,1930}(1 + g_{1930-1940}) \tag{9}$$

where $g_{1930-1940}$ represents the average population growth rate across all non-Minnesota metropolitan areas.

The 1924 quota gives annual quota numbers for a set of sending regions R. For each sending region $r \in R$ and city c, we calculate historical settlement patterns using 1930

data:

$$\omega_{cr} = \frac{M_{cr,1930}}{\sum_{i \in \{V \cup c\}} M_{ir,1930}} \tag{10}$$

where $M_{cr,1930}$ represents the migrant population from region r in city c in 1930.

We then predict the number of new migrants in each city by combining these settlement patterns with immigration quotas:

$$\hat{M}_{c,1940} = \frac{\sum_r \omega_{cr} \cdot Q_r \cdot 10}{h_c} \tag{11}$$

where Q_r represents the annual quota for region r under the Immigration Act of 1924, multiplied by 10 to cover the decade. h_c represents the average number of residents per household in city c in 1930, as we convert from the individual count of the quota to the unit of measure for this paper, the household.

For the addition-level instrument, we distribute the predicted city-level migrants across additions based on each addition's share of city households:

$$\hat{M}_{ac,1940} = \hat{M}_{c,1940} \cdot \frac{H_{ac,1940}}{\sum_a H_{ac,1940}} \tag{12}$$

where $H_{ac,1940}$ represents the household count in addition a of city c.

Finally, we construct our addition-level instrument as:

$$Z_{ac,1940} = \frac{M_{ac,1930} + \hat{M}_{ac,1940}}{\hat{P}_{ac,1940}} \tag{13}$$

where $M_{ac,1930}$ represents the existing migrant population in addition a of city c in 1930, and $\hat{P}_{ac,1940}$ is the predicted 1940 population of the addition, calculated by distributing the predicted city population proportionally to addition household counts. This instrument satisfies the exclusion restriction as it relies on two plausibly exogenous components: national immigration quotas that were set without consideration of local conditions, and historical settlement patterns from 1930 that are unlikely to be correlated with addition-level changes in the 1930s after controlling for city characteristics.

C. Model Appendix

C.1 Proof of Theorem 1

Proof. By definition of $\tilde{\beta}$ and $\hat{\beta}$, we have the following inequalities:

$$\tilde{L}^m(\hat{\beta}^m) < \tilde{L}^m(\tilde{\beta}^m), \quad L^m(\tilde{\beta}^m) < L^m(\hat{\beta}^m)$$

Combining both inequalities, we get the following equivalence relations:

$$\begin{aligned}
&\tilde{L}^m(\hat{\beta}^m) - L^m(\hat{\beta}^m) < \tilde{L}^m(\tilde{\beta}^m) - L^m(\tilde{\beta}^m) \iff \\
&\sum_{j\in\mathcal{J}0} N_j \log\Big(\frac{\tilde{\mathbb{P}}j(\hat{\beta}^m)}{\mathbb{P}j(\hat{\beta}^m)}\Big) < \sum_{j\in\mathcal{J}0} N_j \log\Big(\frac{\tilde{\mathbb{P}}j(\tilde{\beta}^m)}{\mathbb{P}j(\tilde{\beta}^m)}\Big) \iff \\
&\sum_{j\in\mathcal{J}0} N_j \log\Big(\frac{\sum_{k\in\mathcal{J}0}\exp(\hat{\beta}^m \log p_k)}{\sum_{k\in\mathcal{J}}\exp(\hat{\beta}^m \log p_k)}\Big) < \sum_{j\in\mathcal{J}0} N_j \log\Big(\frac{\sum_{k\in\mathcal{J}0}\exp(\tilde{\beta}^m \log p_k)}{\sum k\in\mathcal{J}\exp(\tilde{\beta}^m \log p_k)}\Big) \iff \\
&\log\Big(\frac{\sum_{k\in\mathcal{J}0}\exp(\hat{\beta}^m \log p_j)}{\sum_{k\in\mathcal{J}}\exp(\hat{\beta}^m \log p_k)}\Big) < \log\Big(\frac{\sum_{k\in\mathcal{J}0}\exp(\tilde{\beta}^m \log p_k)}{\sum_{k\in\mathcal{J}}\exp(\tilde{\beta}^m \log p_k)}\Big) \iff \\
&\frac{\sum_{j\in\mathcal{J}_0}\exp(\hat{\beta}^m \log p_j)}{\sum_{j\in\mathcal{J}}\exp(\hat{\beta}^m \log p_j)} < \frac{\sum_{j\in\mathcal{J}_0}\exp(\tilde{\beta}^m \log p_j)}{\sum_{j\in\mathcal{J}}\exp(\tilde{\beta}^m \log p_j)}
\end{aligned}$$

Define the following function:

$$F(\beta) = \sum_{j\in\mathcal{J}_0} \tilde{\mathbb{P}}_j(\beta) = \frac{\sum_{j\in\mathcal{J}_0}\exp\left(\beta \log p_j\right)}{\sum_{j\in\mathcal{J}}\exp(\beta \log p_j)}.$$

Under this definition, the above inequalities can be written as:

$$F(\hat{\beta}) < F(\tilde{\beta}).$$

Taking derivatives of $F(\beta)$, we get:

$$\begin{aligned}\frac{\partial F(\beta)}{\partial \beta} &= \sum_{j\in\mathcal{J}_0} \tilde{\mathbb{P}}_j(\beta)\Big\{\log p_j - \sum_{j\in\mathcal{J}} \tilde{\mathbb{P}}_j(\beta)\log p_j\Big\} \\ &= \sum_{j\in\mathcal{J}_0} \tilde{\mathbb{P}}_j(\beta)\log p_j - F(\beta)\sum_{j\in\mathcal{J}} \tilde{\mathbb{P}}_j(\beta)\log p_j.\end{aligned}$$

From the last expression, it is easy to see that[1]

$$\frac{\partial F(\beta)}{\partial \beta} < 0 \iff \sum_{j\in\mathcal{J}_0} \mathbb{P}_j(\beta)\log p_j - \sum_{j\in\mathcal{J}} \tilde{\mathbb{P}}_j(\beta)\log p_j < 0$$

Observe that

$$\begin{aligned}&\sum_{j\in\mathcal{J}_0} \mathbb{P}_j(\beta)\log p_j - \sum_{j\in\mathcal{J}} \tilde{\mathbb{P}}_j(\beta)\log p_j \\ =& \sum_{j\le j_1, j\in\mathcal{J}_0} [\mathbb{P}_j(\beta) - \tilde{\mathbb{P}}_j(\beta)]\log p_j + \sum_{j_1<j<q_1, j\in\mathcal{J}_0} [\mathbb{P}_j(\beta) - \tilde{\mathbb{P}}_j(\beta)]\log p_j \\ &- \sum_{j_1\le q\le q_1, q\in\mathcal{J}_1} \tilde{\mathbb{P}}_q(\beta)\log p_q - \sum_{q>q_1, q\in\mathcal{J}_1} \tilde{\mathbb{P}}_q(\beta)\log p_q \\ <& \sum_{j\le j_1, j\in\mathcal{J}_0} [\mathbb{P}_j(\beta) - \tilde{\mathbb{P}}_j(\beta)]\log p_j + \sum_{j_1<j<q_1, j\in\mathcal{J}_0} [\mathbb{P}_j(\beta) - \tilde{\mathbb{P}}_j(\beta)]\log p_j \\ &- \sum_{j_1\le q\le q_1, q\in\mathcal{J}_1} \tilde{\mathbb{P}}_q(\beta)\log p_{j_1} - \sum_{q>q_1, q\in\mathcal{J}_1} \tilde{\mathbb{P}}_q(\beta)\max_{j\in\mathcal{J}_0}\log p_j\end{aligned}$$

The inequality follows from assumption 1.

The probability should satisfy

$$\sum_{j\le j_1, j\in\mathcal{J}_0} [\mathbb{P}_j(\beta) - \tilde{\mathbb{P}}_j(\beta)] + \sum_{j_1<j<q_1, j\in\mathcal{J}_0} [\mathbb{P}_j(\beta) - \tilde{\mathbb{P}}_j(\beta)] = \sum_{j_1\le q\le q_1, q\in\mathcal{J}_1} \tilde{\mathbb{P}}_q(\beta) + \sum_{q>q_1, q\in\mathcal{J}_1} \tilde{\mathbb{P}}_q(\beta)$$

[1]Notice that $\frac{\tilde{\mathbb{P}}_j(\beta)}{F(\beta)} = \frac{\tilde{\mathbb{P}}_j(\beta)}{\sum_{k\in\mathcal{J}_0}\tilde{\mathbb{P}}_k(\beta)} = \mathbb{P}_j(\beta)$.

Replace $\sum_{q>q_1,q\in\mathcal{J}_1}\tilde{\mathbb{P}}_q(\beta)$ in the above inequality, we get

$$
\begin{aligned}
&\sum_{j\in\mathcal{J}_0}\mathbb{P}_j(\beta)\log p_j-\sum_{j\in\mathcal{J}}\tilde{\mathbb{P}}_j(\beta)\log p_j\\
<&\sum_{j\le j_1,j\in\mathcal{J}_0}[\mathbb{P}_j(\beta)-\tilde{\mathbb{P}}_j(\beta)][\log p_j-\max_{j\in\mathcal{J}_0}\log p_j]+\sum_{j_1<j<q_1,j\in\mathcal{J}_0}[\mathbb{P}_j(\beta)-\tilde{\mathbb{P}}_j(\beta)][\log p_j-\max_{j\in\mathcal{J}_0}\log p_j]\\
&-\sum_{j_1\le q\le q_1,q\in\mathcal{J}_1}\tilde{\mathbb{P}}_q(\beta)[\log p_{j_1}-\max_{j\in\mathcal{J}_0}\log p_j]\\
<&\sum_{j_1<j<q_1,j\in\mathcal{J}_0}[\mathbb{P}_j(\beta)-\tilde{\mathbb{P}}_j(\beta)][\log p_j-\max_{j\in\mathcal{J}_0}\log p_j]\\
&+\Big\{\sum_{j\le j_1,j\in\mathcal{J}_0}[\mathbb{P}_j(\beta)-\tilde{\mathbb{P}}_j(\beta)]-\sum_{j_1\le q\le q_1,q\in\mathcal{J}_1}\tilde{\mathbb{P}}_q(\beta)\Big\}[\log p_{j_1}-\max_{j\in\mathcal{J}_0}\log p_j]
\end{aligned}
$$

The first term is negative[2]. If we show the terms in the brace is negative, then the above formula is negative.

Recall that $\mathbb{P}_j(\beta)=\frac{\exp(\beta\log p_j)}{\sum_{j\in\mathcal{J}_0}\exp(\beta\log p_j)}$ and $\tilde{\mathbb{P}}_j(\beta)=\frac{\exp(\beta\log p_j)}{\sum_{j\in\mathcal{J}}\exp(\beta\log p_j)}$, the terms in the brace can be written as:

$$
\begin{aligned}
&\sum_{j\le j_1,j\in\mathcal{J}_0}[\mathbb{P}_j(\beta)-\tilde{\mathbb{P}}_j(\beta)]-\sum_{j_1\le q\le q_1,q\in\mathcal{J}_1}\tilde{\mathbb{P}}_q(\beta)\\
=&\sum_{j<j_1,j\in\mathcal{J}_0}\frac{\exp(\beta\log p_j)}{\sum_{j\in\mathcal{J}_0}\exp(\beta\log p_j)}-\sum_{j<j_1,j\in\mathcal{J}_0}\frac{\exp(\beta\log p_j)}{\sum_{j\in\mathcal{J}}\exp(\beta\log p_j)}-\sum_{j_1\le q\le q_1,q\in\mathcal{J}_1}\frac{\exp(\beta\log p_q)}{\sum_{j\in\mathcal{J}}\exp(\beta\log p_j)}\\
=&\frac{\sum_{j<j_1,j\in\mathcal{J}_0}\exp(\beta\log p_j)\sum_{q\in\mathcal{J}_1}\exp(\beta\log p_q)-\sum_{j_1\le q\le q_1,q\in\mathcal{J}_1}\exp(\beta\log p_q)\sum_{j\in\mathcal{J}_0}\exp(\beta\log p_j)}{\sum_{j\in\mathcal{J}_0}\exp(\beta\log p_j)\sum_{j\in\mathcal{J}}\exp(\beta\log p_j)}\\
=&\frac{\sum_{j<j_1,j\in\mathcal{J}_0}\exp(\beta\log p_j)\sum_{q>q_1,q\in\mathcal{J}_1}\exp(\beta\log p_q)}{\sum_{j\in\mathcal{J}_0}\exp(\beta\log p_j)\sum_{j\in\mathcal{J}}\exp(\beta\log p_j)}\\
&-\frac{\sum_{j_1\le q\le q_1,q\in\mathcal{J}_1}\exp(\beta\log p_q)\sum_{j_1\le j\le q_1,j\in\mathcal{J}_0}\exp(\beta\log p_j)}{\sum_{j\in\mathcal{J}_0}\exp(\beta\log p_j)\sum_{j\in\mathcal{J}}\exp(\beta\log p_j)}\\
=&\sum_{j<j_1,j\in\mathcal{J}_0}\mathbb{P}_j(\beta)\sum_{q>q_1,q\in\mathcal{J}_1}\tilde{\mathbb{P}}_q(\beta)-\sum_{j_1\le j\le q_1,j\in\mathcal{J}_0}\mathbb{P}_j(\beta)\sum_{j_1\le q\le q_1,q\in\mathcal{J}_1}\tilde{\mathbb{P}}_q(\beta)\\
=&\sum_{q>q_1,q\in\mathcal{J}_1}\tilde{\mathbb{P}}_q(\beta)-\sum_{j_1\le j\le q_1,j\in\mathcal{J}_0}\mathbb{P}_j(\beta)\sum_{q\in\mathcal{J}_1}\tilde{\mathbb{P}}_q(\beta)
\end{aligned}
$$

[2]Note that for all $j\in\mathcal{J}_0$, $\mathbb{P}_j(\beta)>\tilde{\mathbb{P}}_j(\beta)$ for all β

The last line can be written into:

$$\begin{aligned}
&\sum_{q>q_1, q\in\mathcal{J}_1} \tilde{\mathbb{P}}_k(\beta) - \sum_{j_1\le j\le q_1, j\in\mathcal{J}_0} \mathbb{P}_j(\beta) \sum_{q\in\mathcal{J}_1} \tilde{\mathbb{P}}_q(\beta) \\
&= \sum_{q\in\mathcal{J}_1} \tilde{\mathbb{P}}_q(\beta)[\frac{\sum_{q>q_1, q\in\mathcal{J}_1} \tilde{\mathbb{P}}_q(\beta)}{\sum_{q\in\mathcal{J}_1} \tilde{\mathbb{P}}_q(\beta)} - \sum_{j_1\le j\le q_1, j\in\mathcal{J}_0} \mathbb{P}_j(\beta)] \\
&= \sum_{q\in\mathcal{J}_1} \tilde{\mathbb{P}}_q(\beta)[\sum_{q>q_1, q\in\mathcal{J}_1} \overline{\mathbb{P}}_q(\beta) - \sum_{j_1\le j\le q_1, j\in\mathcal{J}_0} \mathbb{P}_j(\beta)]
\end{aligned}$$

where $\overline{\mathbb{P}}_q(\beta) = \frac{\exp(\beta \log p_q)}{\sum_{q'\in\mathcal{J}_1} \exp(\beta \log p_{q'})}$.

It is clear that:

$$\sum_{q>q_1, q\in\mathcal{J}_1} \overline{\mathbb{P}}_q(\beta) < \sum_{j_1\le j\le q_1, j\in\mathcal{J}_0} \mathbb{P}_j(\beta) \iff \sum_{j\in\mathcal{J}_0} \mathbb{P}_j(\beta) \log p_j - \sum_{j\in\mathcal{J}} \tilde{\mathbb{P}}_j(\beta) \log p_j < 0 \iff \tilde{\beta} < \hat{\beta}$$

For $\beta \in [0, \overline{\beta})$, we have:

$$\begin{aligned}
\sum_{j_1\le q\le q_1, q\in\mathcal{J}_1} \overline{\mathbb{P}}_q(\beta) &= \frac{\sum_{j_1\le q\le q_1, q\in\mathcal{J}_1} \exp(\beta \log p_q)}{\sum_{q\in\mathcal{J}_1} \exp(\beta \log p_q)} > \frac{N_1^s \exp(\beta \log p_q^{min})}{N_1 \exp(\beta \log p_q^{max})} \\
\sum_{j<j_1, j\in\mathcal{J}_0} \mathbb{P}_j(\beta) &= \frac{\sum_{j<j_1, j\in\mathcal{J}_0} \exp(\beta \log p_j)}{\sum_{j\in\mathcal{J}_0} \exp(\beta \log p_j)} < \frac{N_0^s \exp(\beta \log p_{j_1})}{N_0 \exp(\beta \log p_j^{min})} \\
\frac{N_1^s \exp(\beta \log p_q^{min})}{N_1 \exp(\beta \log p_q^{max})} \Big/ \frac{N_0^s \exp(\beta \log p_{j_1})}{N_0 \exp(\beta \log p_q^{min})} &= \frac{N_1^s N_0}{N_1 N_0^s} \exp\big(\beta \log \frac{p_q^{min}}{p_q^{max}} \frac{p_j^{min}}{p_{j_1}}\big) \\
&> \frac{N_1^s N_0}{N_1 N_0^s} \exp\big(\frac{\log \frac{N_0^s}{N_0} \frac{N_1}{N_1^s}}{\log \frac{p_q^{min}}{p_q^{max}} \frac{p_j^{min}}{p_{j_1}}} \log \frac{p_q^{min}}{p_q^{max}} \frac{p_j^{min}}{p_{j_1}}\big) \\
&> 1
\end{aligned}$$

So $\sum_{j_1\le q\le q_1, q\in\mathcal{J}_1} \overline{\mathbb{P}}_q(\beta) > \sum_{j<j_1, j\in\mathcal{J}_0} \mathbb{P}_j(\beta)$ for $\beta \in [0, \overline{\beta})$. This is equivalent to $\sum_{q>q_1, q\in\mathcal{J}_1} \overline{\mathbb{P}}_q(\beta) < \sum_{j_1\le j\le q_1, j\in\mathcal{J}_0} \mathbb{P}_j(\beta)$ for $\beta \in [0, \overline{\beta})$.

For $\beta \in (\underline{\beta}, 0)$, we have:

$$\sum_{j_1 \le q \le q_1, q \in \mathcal{J}_1} \overline{\mathbb{P}}_q(\beta) = \frac{\sum_{j_1 \le q \le q_1, q \in \mathcal{J}_1} \exp(\beta \log p_q)}{\sum_{q \in \mathcal{J}_1} \exp(\beta \log p_q)} > \frac{N_1^s \exp(\beta \log p_{q_1})}{N_1 \exp(\beta \log p_q^{min})}$$

$$\sum_{j < j_1, j \in \mathcal{J}_0} \mathbb{P}_j(\beta) = \frac{\sum_{j < j_1, j \in \mathcal{J}_0} \exp(\beta \log p_j)}{\sum_{j \in \mathcal{J}_0} \exp(\beta \log p_j)} < \frac{N_0^s \exp(\beta \log p_j^{min})}{N_0 \exp(\beta \log p_j^{max})}$$

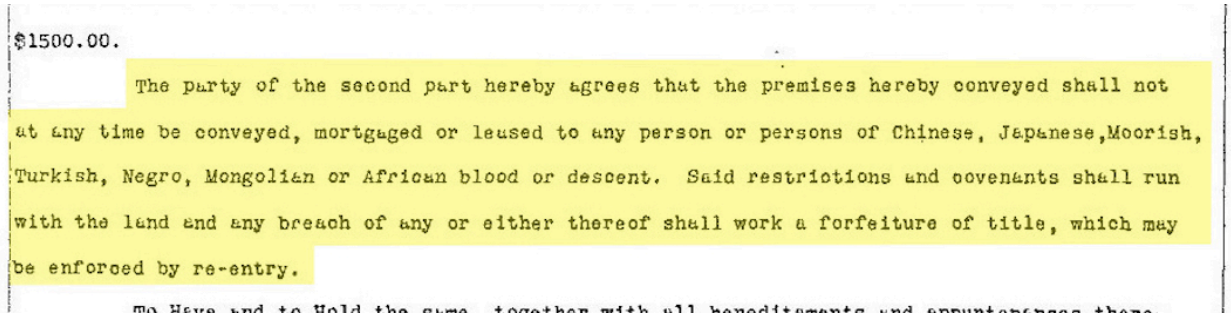
$$\begin{aligned}
\frac{N_1^s \exp(\beta \log p_{q_1})}{N_1 \exp(\beta \log p_q^{min})} \Big/ \frac{N_0^s \exp(\beta \log p_j^{min})}{N_0 \exp(\beta \log p_j^{max})} &= \frac{N_1^s N_0}{N_1 N_0^s} \exp\left(\beta \log \frac{p_{q_1}}{p_q^{min}} \frac{p_j^{max}}{p_j^{min}}\right) \\
&> \frac{N_1^s N_0}{N_1 N_0^s} \exp\left(\frac{\log \frac{N_0^s}{N_0} \frac{N_1}{N_1^s}}{\log \frac{p_{q_1}}{p_q^{min}} \frac{p_j^{max}}{p_j^{min}}} \log \frac{p_{q_1}}{p_q^{min}} \frac{p_j^{max}}{p_j^{min}}\right) \\
&> 1
\end{aligned}$$

So $\sum_{q > q_1, q \in \mathcal{J}_1} \overline{\mathbb{P}}_q(\beta) < \sum_{j_1 \le j \le q_1, j \in \mathcal{J}_0} \mathbb{P}_j(\beta)$ also holds for $\beta \in (\underline{\beta}, 0)$.

So within the range of $(\underline{\beta}, \overline{\beta})$, the inequality $\sum_{q > q_1, q \in \mathcal{J}_1} \overline{\mathbb{P}}_q(\beta) < \sum_{j_1 \le j \le q_1, j \in \mathcal{J}_0} \mathbb{P}_j(\beta)$ holds, so we have $\tilde{\beta} < \hat{\beta}$. □

D. Additional Figures and Tables

Figure D.1: Example 1 of a Racial Covenant Deed

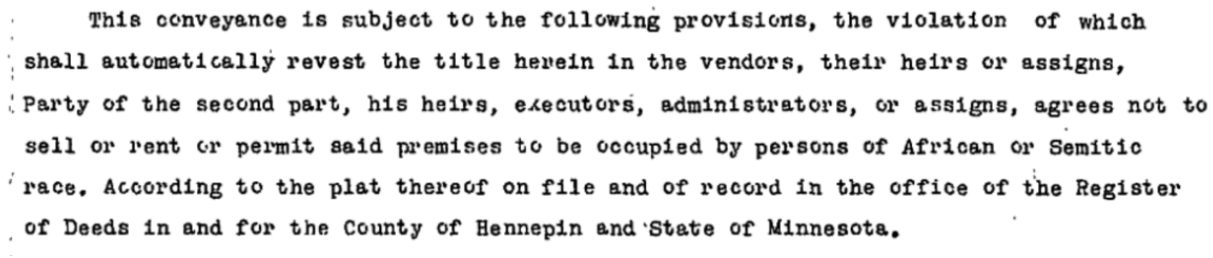


$1500.00.

The party of the second part hereby agrees that the premises hereby conveyed shall not at any time be conveyed, mortgaged or leased to any person or persons of Chinese, Japanese,Moorish, Turkish, Negro, Mongolian or African blood or descent. Said restrictions and covenants shall run with the land and any breach of any or either thereof shall work a forfeiture of title, which may be enforced by re-entry.

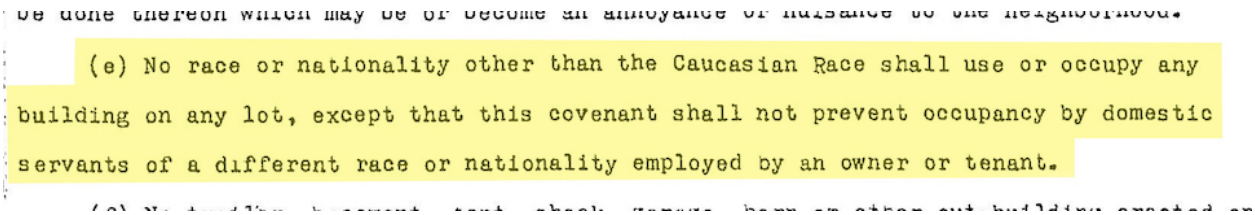
To Have and to Hold the same, together with all hereditaments and appurtenances there-

Figure D.2: Example 2 of a Racial Covenant Deed



This conveyance is subject to the following provisions, the violation of which shall automatically revest the title herein in the vendors, their heirs or assigns, Party of the second part, his heirs, executors, administrators, or assigns, agrees not to sell or rent or permit said premises to be occupied by persons of African or Semitic race. According to the plat thereof on file and of record in the office of the Register of Deeds in and for the County of Hennepin and State of Minnesota.

Figure D.3: Example 3 of a Racial Covenant Deed



be done thereon which may be or become an annoyance or nuisance to the neighborhood.

(e) No race or nationality other than the Caucasian Race shall use or occupy any building on any lot, except that this covenant shall not prevent occupancy by domestic servants of a different race or nationality employed by an owner or tenant.

(f) No tent, basement, tent, shack, garage, barn or other out building erected on

Figure D.4: Newspaper advertisement of Lake of Isles development

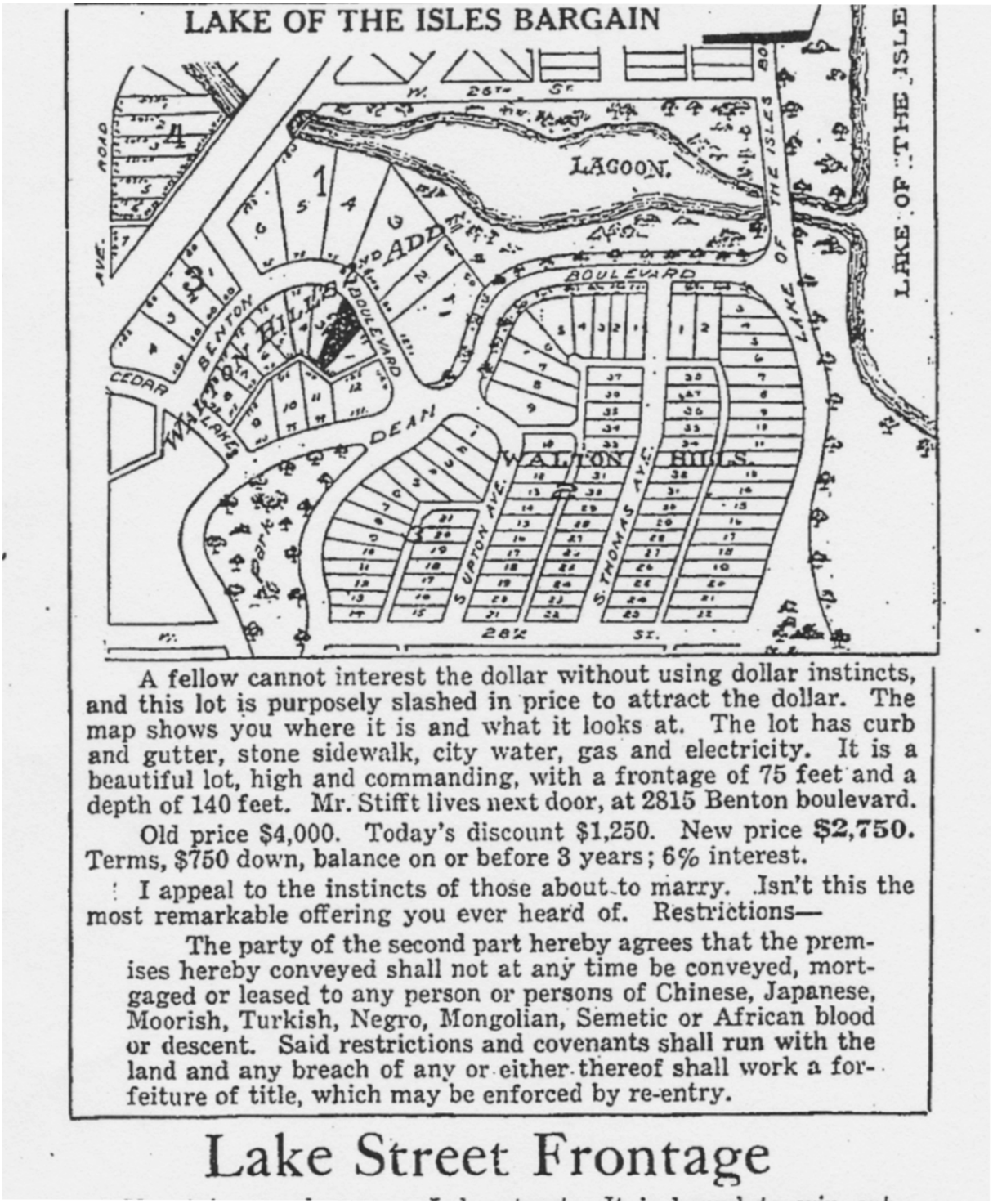


Figure D.5: Example of Development Data from 1940 Minneapolis City Atlas

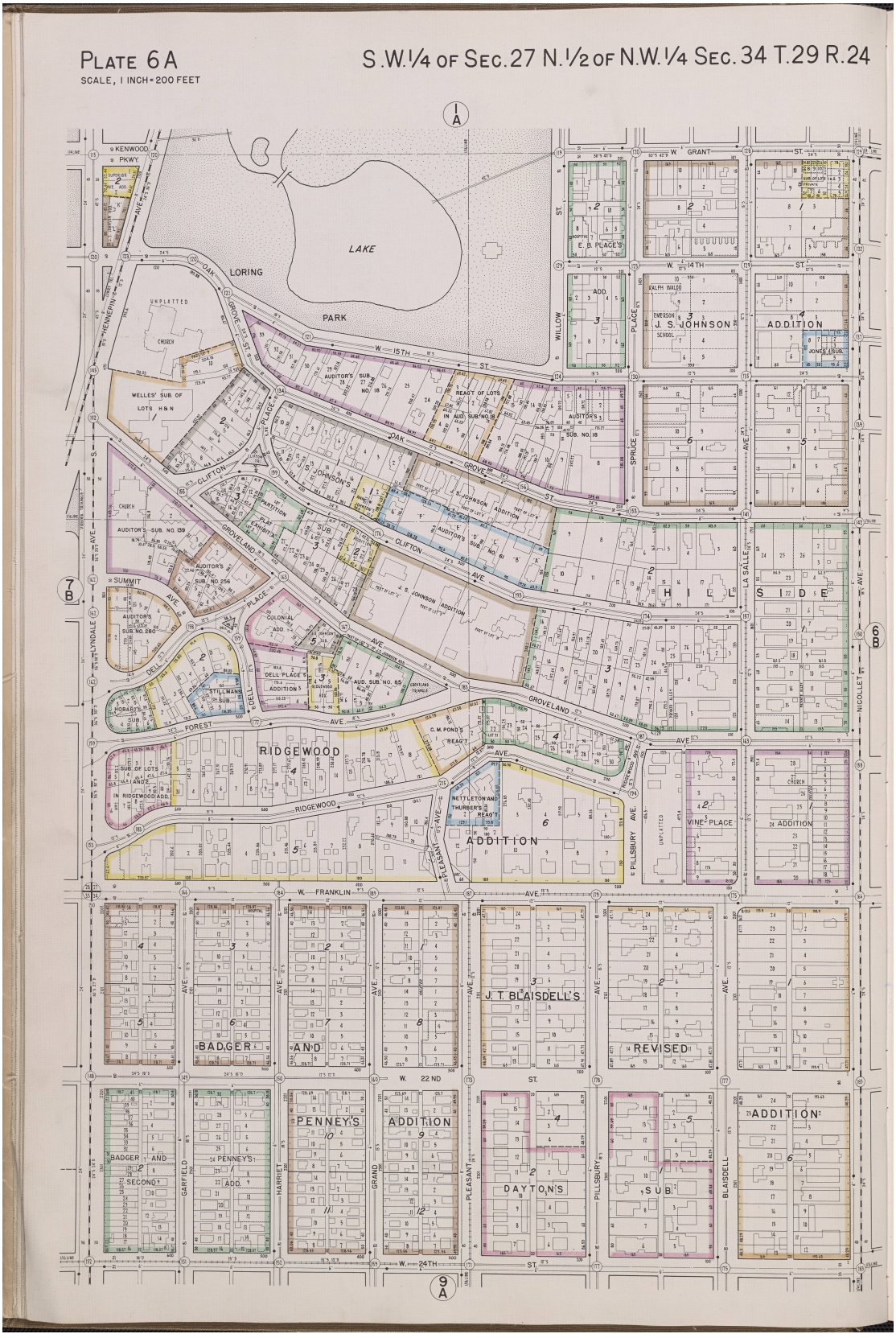


Figure D.6: House Value, Rental Price, and Imputed Rental Value Distributions

