Neighborhood Change and Tipping Points

William Cunningham*

November 11, 2024

Abstract

This paper applies Schelling's model of segregation to neighborhood choice. I estimate neighborhood choice as a function of past neighborhood Minority shares for various demographic groups. A conditional instrumental variables estimator allows for unbiased estimation of preferences for Minority shares by addressing endogeneity concerns present in my theoretical model. Preference parameters for every US state are estimated. Google trends is used to identify a correlation between within-group preferences and hateful and xenophobic rhetoric. Furthermore, neighborhood specific turning point simulations are estimated for Texas census tracts. An investigation of turning points in cities reveals implications for inner city populations. The results of this paper suggest that neighborhood demand is sensitive to within-group preferences, and varies by age.

JEL Classification: D2, J3, M5, R3, R4

*Department of Economics, Southern Methodist University. E-mail: cunninghamw@smu.edu

1 Introduction

Although residential segregation has decreased since the Civil Rights era, it remains a persistent challenge in contemporary American society. As recently as 2023, policies combating neighborhood segregation have been passed. [1] Such policies are motivated by a long literature documenting segregation's negative impacts on welfare. Logan and Stults [2022] notes "residential segregation, especially between Whites and blacks, is a durable feature of America's urban landscape". This is especially concerning given segregation may result in welfare differences between demographic groups. Welfare differences may be driven by differential access to amenities. For instance, Card and Krueger [1992] note school segregation results in long-run Black-White wage gaps. While many overt segregative policies are illegal in the US, segregation persists, especially in cities.

This paper studies the dynamics of segregation in the US. More specifically, I study how the proportion of minorities in a location's population influences future population's choice of residential location. To accomplish this, I employ a location choice model popularized by Schelling [1971]. Schelling's model casts current location choice as a function of the past racial composition of a neighborhood, and amenities available at that location. This agent-based models showed how mild individual preferences for same-group neighbors could produce complete segregation through cascading effects. Preferences over race may composed of direct discrimination, and preferences over expected endogenous amenities correlated with race. Amenities are qualities of a location that make it attractive or pleasant to live in. Examples of amenities can take the form of natural geographic features, such as a beach or mountain view, or man-made features, such as schools and shopping centers. Estimating the effect of a neighborhood's past racial composition through the lens of this model provides us with a tool to study voluntary segregation.

A key feature of this approach is the existence of tipping points. A tipping point may be defined as the minority share in a neighborhood for which the trend of demographics in that neighborhood changes direction. For segregation, this would look like a neighborhood trending towards a low(high) minority share of population in a neighborhood being shocked and then rending towards a high(low minority share of population. Taking this theoretical model to the data poses challenges. First and foremost, past Minority population shares for neighborhoods are correlated with current populations. This is for two reasons. First, the individuals comprising past Minority shares may still be present in the current demographic makeup of a neighborhood. Second, Schelling's model has neighborhood amenities impacting past and current location

[1] For instance, the Notice of Proposed Rulemaking (NPRM) "Affirmatively Furthering Fair Housing" was published last year and aims to combat housing segregation (U.S. Department of Housing and Urban Development [2023]).

choice for individuals. For instance, a neighborhood amenity such as a park may persist through time. This impacts individuals' choices on where to locate through time as well. An estimation of the effect of past neighborhood Minority population shares on current population would then pick up on the effect of the park's persistent existence. Caetano and Maheshri [2017] face a similar issue in their school choice paper. Their solution is a conditional instrumental variable approach is used to control for persistent amenities, and identify exogenous variation in past Minority shares.

In this paper, I apply the conditional instrumental variables approach to estimate preferences for past Minority shares by state. I find evidence of within-group preferences in location choice, indicating homophily. Such preferences provide an endogenous explanation for the persistence of segregation in US cities to this day. Estimation of homophily varies by state. To unpack this variation, I regress racial preference parameters on measures of racial animus and xenophobia from Google trends data by state. I also adapt Caetano and Maheshri's estimation of tipping points procedure to estimate equilibria for neighborhood Minority shares. This estimation approach allows for a more granular prediction of tipping points at the census tract level. This is an advantage over the literature's common approaches using estimations at the metropolitan statistical area level, like that of Card et al. [2008]'s foundational work in this field. As noted by Krugman [1991] and other works in the urban literature, economic activity varies with the density of cities. When mapped, this paper's predicted segregation dynamics in conjunction with the urban literature's findings on density paint a worrying picture of worsening segregation and welfare outcomes for Minority groups. This paper's contribution is in applying the conditional instrumental IV approach to the entirety of the US, predicting segregation dynamics for neighborhoods, and unpacking estimated racial preferences.

The study of residential segregation and neighborhood tipping points has been central to urban economics since Schelling's seminal work demonstrating how individual preferences can lead to aggregate patterns of segregation even when no individual explicitly desires such outcomes. This literature has evolved to encompass rich empirical investigations of actual tipping behavior, theoretical refinements incorporating market forces, and policy analyses examining interventions aimed at promoting integration.

Schelling's framework has been extensively utilized by subsequent scholars. Sethi and Somanathan [2004]) incorporate heterogeneous preferences to this framework and demonstrate how income disparities interact with preferences for neighborhood racial composition. This paper marks an explicit attempt at linking the segregation puzzle to the role of cities. Zhang [2011] extended the basic model to include housing prices, showing how market forces can either amplify or dampen segregating tendencies depending on the distribution of income across groups. While these papers are important to our understanding of segregation, they do not provide empirical evidence for the existence of tipping, or provide estimates for preferences for the racial composition of locations. This paper, along with others in the empirical literature, add to this discussion through providing such evidence.

Card et al. [2008] provide the first systematic empirical evidence for tipping behavior in U.S. metropolitan areas between 1970 and 2000. They found that many White neighborhoods experienced sharp outflows of White residents when Minority shares exceeded certain thresholds. This is achieved through a multi-step root finding method for polynomial functions of Minority share in a given area and through a regression-discontinuity design. This is evident in the long line of papers adhering to Card's methodology. While most papers of this literature focus on the US, Aldén et al. [2015] use Card's regression discontinuity approach to study residential mobility in Europe. Böhlmark and Willén [2020] even utilize estimated tipping points as an instrument to study the labor market impact of immigration. One disadvantage of Card's approach is in its lack of granularity, being only applicable to larger geographies such as an Metropolitan Statistical Area(MSA). This is due to comparing each location's Minority shares to a greater area's average Minority share. A weakness of this approach is in it's large geographical unit of estimation. Estimation at the MSA level loses relevent detail's about race in cities at smaller geographical levels. Despite this shortcoming, Card et al's seminal paper is a useful benchmark for all future work on neighborhood tipping points. I improve on this weakness through recent empirical techniques from the school segregation literature.

To apply Schelling's segregation model to finer geographies, this paper uses a methodology detailed in Caetano and Maheshri [2017]. Endogeneity threatens identification of population responses to past Minority shares. Past persistent amenities unobserved by the econometrician present an omitted variables bias. Caetano and Maheshri study segregation in school choice using a conditional instrumental variables approach. They find within-group preferences for school choice, and they are able to estimate tipping points at the school level. A key facet of their approach is the imposition of a cohort structure, allowing for the construction of an unbiased conditional instrumental variable. I further detail this empirical approach in section 3, and apply it to location choice. This paper estimates the aggregate effect of minority shares on future populations. This approach does not make a distinction between direct racial preferences and preferences for amenities endogenous to the presence of those Minorities. I contribute this process by going one step further and regress coefficients from these estimations on measures of racial animus. Data on Google search frequency provides a rich empirical setting to test the role of racial animus in location choice.

This paper is organized as follows, section 2 details the data and empirical setting for this paper.

Section 3 is broken into two subsections, the first of which concerns estimation of preferences for Minority shares in neighborhoods, and the second uses these estimates to study the dynamics of segregation. Section 4 details the results of these approaches, and section 5 concludes.

2 Data

Data on population counts are from the US Census Bureau. I collect data from the Decennial Census for the years 2000, 2010, and 2020 [2]. This study uses census tracts as the primary unit of analysis, representing neighborhoods in the theoretical framework. This distinction is an important one, as a major contribution of this paper is the estimation of tipping points at a more granular geography than a Metropolitan Statistical Area. Defining neighborhoods has been an ever present issue to applied economists. Tiebout [1956]'s seminal paper conceptualized neighborhoods as local jurisdictions where people "vote with their feet" by choosing where to live based on their preferences for public goods and services. More recently, Durlauf [2004] in his chapter "Neighborhood Effects" in the Handbook of Regional and Urban Economics provides a comprehensive treatment. He defines neighborhoods as "the geographic and social space in which externalities between individuals are prevalent and persistent." Adao et al. [2019] and Ahlfeldt et al. [2015] find that spillovers are highly localized. A census tract tends to house a few thousand people and is the largest geography that doesn't envelop whole cities in the way a county or Metropolitan Statistical Area would U.S. Census Bureau [2022]. To study city dynamics, which the literature has identified as the source of the most potent segregation in the US, census tracts are an appropriate choice of observation unit.

I restrict the sample to working age population, 25-65, to avoid confounding factors. Populations are aggregated into White, and non-White or Minority demographic groups[3]. Furthermore, I group populations into 10-year age cohorts, making four cohorts in total. This results in eight total population groups that will be used in this paper's analysis. Data from all 50 US states is collected, though this paper's primary analysis will focus on Texas. I provide sample summary statistics in the Appendix in Table 6. One can note that restricting my sample to the working age population of the US cuts the largest observation in my sample to around 5000 people. It is also worth noting that Texas has large a substantial non-White population, large metropolitan areas for study, and historical non-White communities making it an ideal candidate for the featured analysis of this paper.

[2] This paper's analysis occurs at the 2020 census tract level. To convert data measured using 2000 and 2010 tract definitions, this paper employs geographic crosswalks from Manson et al. [2024].

[3] Non-White population is constructed as the sum of Hispanic and Black populations. This is to focus on demographic groups that are consistently studied in the segregation and population dynamics literature.

Additional controls are gathered for robustness checks. Income and education of one's peers may influence neighborhood choice, as shown by Diamond [2016]. I use a conditional IV to implicitly control for these variables in most cases. Direct measures of human capital and productivity are also included as robustness checks. This data is gathered from ACS 5-year estimates from 2010. Income and education are measured through the proportion of a census tract with a bachelors degree or greater and median income. These are some of the same controls used in Card et al. [2008] and other studies of population dynamics. From this same source, I collect total tract population to address the threat of more populated tracts acting differently than less populated tracts.

I use google trends to collect data on racial animus and xenophobia in the United States[4]. Following Stephens-Davidowitz [2014], I collect data on searches for hateful terms. In addition, I include results for searches for affirmative action. "Slur1" corresponds to racist rhetoric while "Slur2" corresponds to xenophobia. In my results section, I regress the results for preferences for minority groups on the prevalence of these searches for each US state.

3 Methodology

3.1 Location Choice

The identification strategy models location choice as a function of lagged neighborhood characteristics to estimate the causal effect of past demographic composition on residential decisions. This modeling choice follows a long history of the literature on segregation stemming from Schelling [1971]. Let $n_{j,t}^k$ be the number of people of demographic, $k \in (w, m)$, living in location, j, at time, t. How do individuals make location choices? I assume individuals make myopic decisions based on the demographic composition and available amenities at location j in the prior period, $t-1$. Furthermore, I assume individuals face some moving costs between locations, so that individuals slowly adjust to working locations over time. This assumption plays a crucial role later in my methodological approach. Given these assumptions, the log of $n_{j,t}^k$ is given by the following:

$$log(n_{j,t}^k) = \beta^k s_{j,t-1} + \gamma^k A_{j,t-1} + u_{j,t}^k \tag{1}$$

Where $s_{j,t-1}$ is the Minority group, $m's$ population share in location j at time $t-1$, i.e.: $s_{j,t-1} = \frac{n_{j,t-1}^m}{n_{j,t-1}}$. $A_{j,t-1}$ are the amenities observed by individuals as they make a decision on where to locate. This amenity term acts as a catch-all for a variety of factors that may influence location choice. Examples of $A_{j,t-1}$ range

[4] Exact search terms are available upon request

from local school or road quality, to the kind of restaurants located in j. Finally, $u^k_{j,t}$ is a location-time-demographic specific error term.

The parameter of interest in this paper is the vector β^k, which identifies whether a particular k may have different location demand responses to the Minority group's $t-1$ population shares. This allows for Whites and minorities to react differently to the demographic composition of a given location.

An issue presents itself when one considers that $A_{j,t-1}$ may be unobserved by the econometrician. For example, a local highway constructed before $t-1$ may be differentially preferred by individuals from m and w. In this case, both our dependent variables, current populations, and past shares which are constructed from dependent variables in time $t-1$, $n^k_{j,t-1}$, are correlated with the existence of this highway. More formally $cov(log(n^k_{j,t}), A_{j,t-1} \neq 0$, $cov(log(n^k_{j,t-1}), A_{j,t-2} \neq 0$, $cov(A_{j,t-1}, A_{j,t-2} \neq 0$. If these correlations are due to the same amenities in $t-1$ and $t-2$ I face omitted variables bias.

Let $A_{j,t-1} = A_{j,t-2} + \alpha_{j,t-1}$, where $\alpha_{j,t-1}$ is a White-noise shock, and $v^k_{j,t} = \gamma^k A_{j,t-1} + u^k_{j,t}$ be a composite error term. This assumption aims to capture variation in amenities over time, with some amenities persisting through $A_{j,t-2}$ and others being formed or destroyed through $\alpha_{j,t-1}$. Identification of β^k through $log(n^k_{j,t}) = \beta^k s_{j,t-1} + v^k_{j,t}$ is subject to omitted variables bias from amenities that persist from $t-2$ to $t-1$ in $A_{j,t-2}$. How can persistent amenities be controlled for to allow for unbiased identification of β^k?

This will be accomplished by first introducing cohorts: $n^k_{j,t-1} = \sum_{g=G_j} n^{g,k}_{j,t-1}$, where $G_j = \underline{g}, ..., \overline{g}$ denotes age cohort. Rewrite (1) as the location demand for a cohort-demographic group g, k:

$$logn^{g,k}_{j,t} = \beta^{g,k} s_{j,t-1} + \gamma^{g,k} A_{j,t-1} + u^{g,k}_{j,t} \tag{2}$$

Using the definition of $A_{j,t-1}$:

$$logn^{g,k}_{j,t} = \beta^{g,k} s_{j,t-1} + \gamma^{g,k}(A_{j,t-2} + \alpha_{j,t-1}) + u^{g,k}_{j,t} \tag{3}$$

Now it can be noted that $logn^{g,k}_{j,t}$ and $s_{j,t-1}$ are correlated with $A_{j,t-2}$. However, the g cohort of t is not present in $t-1$ and the g of $t-1$ is not present in t. The overlap of the t and $t-1$ cohorts are all $logn^{g,k}_{j,t}$, save for the oldest, g. I control for relevant $logn^{g,k}_{j,t-1}$ with $C^{g,k}_{j,t-1}$:

$$logn^{g,k}_{j,t} = \beta^{g,k} s_{j,t-1} + \underbrace{\sum_{i=\underline{g}}^{\overline{g}-1} (\alpha^{g,k}_{i,k} logn^{ik}_{j,t-1} + \alpha^{g,k}_{i,k'} logn^{i,k'}_{j,t-1})}_{C^{g,k}_{j,t-1}} + \gamma^{g,k}(A_{j,t-2} + \alpha_{j,t-1}) + u^{g,k}_{j,t} \tag{4}$$

Endogeneity is still a threat to identification due to the assumption that individuals face moving costs and don't immediately adjust their location. In this case, past α may threaten unbiased identification of

$\beta^{g,k}$ as their population effect persists in the dependent variable and $s_{j,t-1}$. An appropriate IV would be something that influences $s_{j,t-1}$ directly, but not $logn^{g,k}_{j,t}$. The IV $s^{\overline{g}-1k'}_{j,t-2}$ satisfies this requirement as this cohort has aged out of the dependent variable in t, and therefore only impacts $logn^{g,k}_{j,t}$ through $s_{j,t-1}$! I modify 4 by collecting unobservable amenities into the composite error term, $\tilde{u}^{g,k}_{j,t}$, and arrive at:

$$logn^{g,k}_{j,t} = \beta^{g,k} s_{j,t-1} + \underbrace{\sum_{i=\underline{g}}^{\overline{g}-1} (\alpha^{g,k}_{i,k} logn^{ik}_{j,t-1} + \alpha^{g,k}_{i,k'} logn^{i,k'}_{j,t-1})}_{C^{g,k}_{j,t-1}} + \tilde{u}^{g,k}_{j,t} \tag{5}$$

which may be estimated via 2SLS and instrumenting for $s_{j,t-1}$ with $s^{\overline{g}-1k'}_{j,t-2}$. This is the primary specification of my analysis. Later robustness checks may pull location peer effects and county fixed effects out of the error term, yielding the new error term $\tilde{u'}^{g,k}_{j,t}$:

$$logn^{g,k}_{j,t} = \beta^{g,k} s_{j,t-1} + \underbrace{\sum_{i=\underline{g}}^{\overline{g}-1} (\alpha^{g,k}_{i,k} logn^{ik}_{j,t-1} + \alpha^{g,k}_{i,k'} logn^{i,k'}_{j,t-1})}_{C^{g,k}_{j,t-1}} + X_{jt} + z_{c,t} + \tilde{u'}^{g,k}_{j,t} \tag{6}$$

Where X_{jt} is a vector of controls and $z_{c,t}$ is a county fixed effect for which each j belongs to some c. 6 may then be estimated in the same manner as 5. To better visualize the conditional IV approach, Table 1 shows the role each cohort plays in estimating 5 and the propagation of a past transitory shock. Consider the example of a public park closure prior between 1970 and 1980 in location j. This shock would affect those who were around to see it, and moving costs ensure that some of the population that chose j in 1980 don't move due to its closure. The remnants of this group drives exogenous variation in the IV cohort. This approach yields unbiased estimates of the impact of Minority shares on location choice. In the next section, I use these estimates to simulate segregation dynamics.

Transitory Amenity Visualization				
Time	**25-35**	**35-45**	**45-55**	**55-65**
2020	Dependent Variables			
2010	Control	Control	Control	
2000	–	–	IV	
1990	–			
1980	*			

Table 1: Visualization of transitory amenities across age groups and time periods. Blue cells indicate presence of amenities, * denotes the impact to the IV cohort in 1980.

3.2 Simulations

Following identification of $\beta^{g,k}$ in 5 for all eight demographic groups [5], I estimate the response of $n^{g,k}_{j,t}$ to hypothetical change in past Minority shares. This is necessary to plot out the dynamics of populations and their tendencies to fall into segregation. Given within-group preferences for the population of a location, neighborhoods populations will trend towards stable equilibria with near-homogeneous demographics. This process reveals the primary result of Schelling-type models, locations tend to self-segregate without policy intervention.

To illustrate segregative tract population dynamics, I follow Caetano and Maheshri [2017] to simulate the evolution of current neighborhood Minority shares in response to changes in past Minority shares. This protocol begins with calculating counterfactual tract populations using:

$$n^{g,k}_{j,t}(s) = exp(logn^{g,k}_{j,t} + \hat{\beta}^{g,k}(s - s_{j,t-1})) \tag{7}$$

Where $s = .001, .002, ..., .999, 1$. s is a simulated shock to past Minority shares and its iteration over zero to one allows for the construction of a location's future Minority share response to changes in $s_{j,t-1}$. $\hat{\beta}^{g,k}$ is a vector of eight key parameters obtained through estimation of 5, one for each g and k pair. As a result, each round of iteration over s results in eight values of $n^{g,k}_{j,t}(s)$. Therefore, simulated Minority share, $S_{j,t}(s)$, is given by:

$$S_{j,t}(s) = \frac{\sum_{g=\underline{g}}^{\overline{g}} n^{g,m}_{j,t}(s)}{\sum_{g=\underline{g}}^{\overline{g}} (n^{g,w}_{j,t}(s) + n^{g,m}_{j,t}(s))} \tag{8}$$

I identify equilibria as the values of $S_{j,t}(s)$ for which $S_{j,t}(s) = s$. Graphically, these equilibria may be identified as values of $S_{j,t}(s)$ where $S_{j,t}(s)$ crosses a 45-degree line on a plot of $S_{j,t}(s)$ on s. For example, Figure 1 shows a location with three equilibrium. The point $(0, 0)$ is a stable equilibrium with a low Minority share, $(1, 1)$ is a stable equilibrium with a high Minority share, and $(.5, .5)$ is an unstable equilibrium, aptly named the "tipping point". In addition to estimating equilibria, I construct confidence intervals of their location through a bootstrap process. This is achieved through draws from multivariate normal distributions of each $\beta^{g,k}$ and estimates of coefficients contained in corresponding $C^{g,k}_{j,t-1}$ terms.

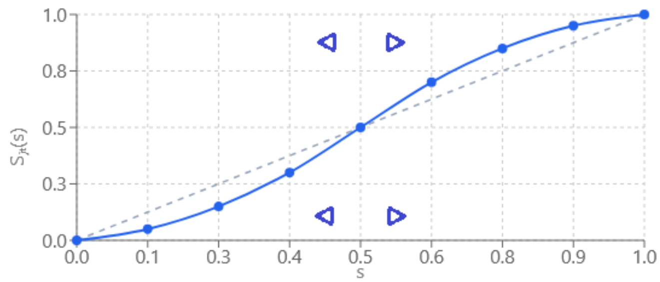

FIGURE 1: Sample Minority share Dynamics

TABLE 2: Estimation of the impact of past Minority share on future populations

	Model				
Dependent Variable	1	2	3	4	5
White Population					
log(25-34 years old)	-2.425*	-4.410***	-3.824***	-3.911***	-3.471***
	(0.996)	(0.613)	(0.604)	(0.632)	(0.516)
log(35-44 years old)	-2.934**	-5.000***	-4.961***	-4.663***	-4.740***
	(0.982)	(0.925)	(0.862)	(0.648)	(0.524)
log(45-54 years old)	-2.623**	-4.355***	-3.915***	-3.964***	-3.921***
	(0.865)	(0.655)	(0.572)	(0.497)	(0.421)
log(55-64 years old)	-2.572**	-3.967***	-3.464***	-3.522***	-3.277***
	(0.998)	(0.890)	(0.726)	(0.730)	(0.564)
Minority Population					
log(25-34 years old)	2.503**	1.596***	1.823***	1.203***	1.184***
	(0.851)	(0.453)	(0.511)	(0.318)	(0.346)
log(35-44 years old)	2.421***	0.997*	0.906***	0.616	0.363
	(0.584)	(0.431)	(0.260)	(0.446)	(0.285)
log(45-54 years old)	1.818**	0.814***	1.104***	0.305	0.315
	(0.667)	(0.204)	(0.169)	(0.225)	(0.178)
log(55-64 years old)	2.168***	1.678***	1.461***	1.234***	0.876***
	(0.637)	(0.308)	(0.180)	(0.323)	(0.205)
Method	OLS	2SLS	2SLS	2SLS	2SLS
Errors			Conley		
Persistent Amenity Controls			Y		
Additional Controls	N	N	Y	N	Y
County FE	N	N	N	Y	Y

Note: Standard errors in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 3: Estimation of Model 2 for White, 25-34

Dependent Variable: log_White_25_34_2020
Endogenous: Minority_Share_2010; **Instrument:** Minority_Share_45_54_2000

Variable	Estimate	Std. Error	t value	Pr(> $\|t\|$)
(Intercept)	2.480035	0.213922	11.593194	2.2e-16^***
fit_Minority_Share_2010	-4.409604	0.612996	-7.193523	6.9909e-13^***
log_Minority_25_34_2010	0.202793	0.100346	2.020945	4.3325e-02^*
log_Minority_35_44_2010	0.306261	0.079551	3.849872	1.1926e-04^***
log_Minority_45_54_2010	-0.124484	0.048604	-2.561222	1.0452e-02^*
log_White_25_34_2010	0.245316	0.056157	4.368358	1.2705e-05^***
log_White_35_44_2010	0.064555	0.020037	3.221751	1.2801e-03^**
log_White_45_54_2010	0.012646	0.023935	0.528356	5.9727e-01
Observations			6,800	
RMSE			2.80208	
Adj. R^2			0.302839	

Note: *p<0.05; **p<0.01; ***p<0.001
Standard errors: Conley (160.934km)
F-statistic (1st stage): 12,097.5 (df = 1; 6,792), p < 2.2e-16
Wu-Hausman: 102.9 (df = 1; 6,791), p < 2.2e-16

4 Results

4.1 Estimating the impact of past Minority share

Table 2 presents my main results. Across all specifications, positive within-group preferences are identified for Texas, corroborating the literature's results on this topic, including that of Schelling [1971] and Card et al. [2008]. Each cell displays an estimate of $\beta^{g,k}$ from equation 5 and corresponds to its own regression whose dependent variable is listed in the left column of the table. Model 2 is the result of the conditional IV estimation 5. This paper's next subsection will use the eight $\beta^{g,k}$ from this specification to perform simulations. The first coefficient of Model 2 may be interpreted as the following: for a given census tract, a one unit increase in past Minority share results in roughly a -4.41 percent change in the population of the White, 25-34 cohort ten years later. The magnitude of within-group preferences is similar across cohorts, albeit with White population's reactions to past Minority shares being slightly greater than that of Minority groups. -4.41% is an economically large number in this context. when applied to the largest census tract, this would equate to a - 10 percentage point change in minority share resulting in the addition

[5] Eight demographic groups arise from their being two racial cohorts, white and non-white, and four age cohorts.

of 20 more white individuals moving to that area. This in turn would modify the minority share again, and the magnitude of the effect of a drop in Minority share would compound as time advanced.

Model 1 shows an OLS estimation of equation 5 . Comparing Model 1's results to the coefficients of Model 2 reveals a pattern. OLS estimates for White cohorts are consistently smaller than 2SLS estimates, while estimates for Minority cohorts are consistently larger. This comes as no surprise given this paper's justification for using a Conditional IV approach in the previous section. Such a result is consistent with transitory amenities unobserved by the econometrician being positively correlated with White location choice and past Minority shares in those locations. These same amenities are likely negatively correlated with future Minority location choice[6]. The literature on gentrification provides an explanation for these results. Freeman [2005] details a common pattern of high-income Whites moving to historically Minority-dominated neighborhoods in search of amenities and displacing Minority residents. Instrumenting for $s_{j,t-1}$ with $s^{\overline{g}-1k'}_{j,t-2}$ solves this problem by only using variation in transitory amenities that do not impact current populations but did impact past Minority shares.

The results of Model 2 for the White, 25-34 cohort are shown in their entirety in Table 3. The six controls of $C^{g,k}_{j,t-1}$ appear in this table, and capture the effect persistent amenities that would impact the dependent variable and the instrument, $s^{\overline{g}-1k'}_{j,t-2}$. Direct interpretation of $C^{g,k}_{j,t-1}$ as a function of persistent amenities is not possible as these controls also are correlated with components of the dependent variable, this confounds these variables interpretations. The F-statistic of 2SLS' first stage is shown at the bottom of the table. This estimation's large F-statistic indicates that the instrument is certainly statistically relevant. Such a result directly assuages concern that variation in past transitory amenities that only impact $s^{\overline{g}-1k'}_{j,t-2}$ are too small to make an impact on Minority shares in $t-1$. This is also evidence that the assumption of moving costs is plausible. Without moving costs, individuals driving the exogenous variation in $s^{\overline{g}-1k'}_{j,t-2}$ would immediately relocate following a transitory amenity shock, and the instrument would have no relevance. Models 3 and 5 provide additional evidence that $C^{g,k}_{j,t-1}$ sufficiently controls for persistent location characteristics. These regressions include additional controls for the total population, education, and income of the population of tracts in 2010. These changes don't result in substantial changes in coefficients from Models 2 to 3.

Locations may see geographic correlation in their population dynamics. This is evident in the existence of cities in Texas, showing "things that are near each other tend to be more related than things that are far apart" (Miller [2004]). These correlations may inflate the statistical significance of this paper's

[6]See Wooldridge [2010] for further explanation.

estimates. To account for this, I follow Conley [1999] and adjust standard errors to account for geographic autocorrelation [7]. It may also be the case that amenities spill over tract borders and result in location choices near, but not in the same tract as desired amenities. Adao et al. [2019] and Ahlfeldt et al. [2015] provide evidence that such spillovers are highly localized. In light of this, Models 4 and 5 include county fixed effects. Such additions have little impact on the sign and magnitude of this paper's estimates. This is evidence that the transitory amenity shocks of interest are highly localized, and Model 2 is a sufficient estimation. The addition of county fixed effects provide an additional benefit of allowing this paper's model to be recast as a logit model if necessary. This is possible as the county fixed effects act as adding the county mean's of log populations to the right hand side of 6. The dependent variable may then be rearranged to be the share of a cohort who chooses to locate in a specific tract over other tracts in that county. While this is not the chief aim of this paper, this does allow for comparison of this model to other logistic location choice models like that of Diamond [2016].

Having selected Model 2 for Texas simulations, I also estimate the model for every US state. Results of within-group preferences are common and consistent across all cohorts. Exceptions to this rule tend to be low population states, like Alaska, which are outliers to the general trend. Texas lands well within the expected results for most states. This makes Texas a great candidate to represent the potential of this method in the next section[8].

To better understand these coefficients, I follow Stephens-Davidowitz [2014]. I regress select search term frequency by state on my estimated coefficients, Table 4 shows these results. Coefficients indicating a preference for minority groups tend to be negatively correlated with measures of racial animus and xenophobia for Whites and positively correlated for Minorities. These results are consistent with the interpretation the beta coefficients. The use of intolerant language is positively correlated with within group preferences. Searches for "Affirmative Action" also seem to carry this trend.

4.2 Simulations

Population dynamics are simulated for all census tracts within Texas. Simulated Minority share responses to past shares are plotted in blue while current Minority shares as of 2020 are given by vertical dashed lines. Current Simulations result in one of three kinds of dynamics. The first of these is a single, low Minority share crossing point. Figure 2 provides an example of such a location for a tract in Palestine TX.

[7] Errors correlations are allowed to exist over a distance of 100 miles (160.963km). While this is computationally expensive, it more than addresses any concerns of error correlations over space.

[8] I include a map of results for the White 35 to 44 cohort for the US in Figure 10 in the Appendix

Table 4: Regression Results for Racial Attitude Measures

	Dependent Variables		
Variables	Slur1	Slur2	Affirmative Action
White 25-34	-8.073*	-6.635	-1.625
	(4.591)	(4.530)	(4.979)
Minority 25-34	2.670***	2.409**	2.335**
	(0.931)	(0.919)	(1.010)
White 35-44	0.110	-0.624	-4.270
	(3.841)	(3.790)	(4.165)
Minority 35-44	3.376***	2.829**	3.065**
	(1.103)	(1.088)	(1.196)
White 45-54	-4.618	-5.088	-4.728
	(3.671)	(3.622)	(3.981)
Minority 45-54	-0.842	-1.105	-1.209
	(0.890)	(0.879)	(0.966)
White 55-64	-5.952	-3.781	-8.573*
	(4.271)	(4.214)	(4.632)
Minority 55-64	1.208**	0.957**	1.011*
	(0.478)	(0.471)	(0.518)
Observations	49	49	49
RMSE	25.9	25.6	28.1

Note: Standard errors in parentheses
*** $p<0.01$, ** $p<0.05$, * $p<0.1$

This figure suggests that Minority share in this location will decrease in the future, and no amenity shock would result in a tipping behavior. These locations are common in places where neighboring White tracts and the trend of the county heavily favor an increase in White population. The second type of dynamics observed in this paper appear constitute a single, high crossing point. Figure 3 shows Crystal City TX which already appears near its predicted high Minority share equilibrium. The figure indicates that little change is expected in this tract. The distribution of single crossing equilibria is shown in 7. It is not surprising that these crossings occur near 0 and 1. This is an expected result of the Schelling-type model.

Triple crossings displaying dynamics shown in Figure 1 may be seen in Figures 5 and 6. The first example of these displays a tract in University Park, TX. This tract displays an already low Minority share in 2020 that is expected to trend towards an even lower Minority share if left untouched by exogenous shocks. The second example of a triple crossing shows the opposite. 2020's Minority share is greater than the estimated tipping point, indicating that Minority share will increase. The distribution of triple crossing equilibria is shown in Figure **??**. Low Minority share stable equilibria are shown in red, high Minority share stable equilibria are shown in blue, and tipping points are shown in yellow. Like the single crossing tracts, the stable equiliria for triple crossing tracts cluster near 0 and 1. These locations are neighborhoods for which

	White		Minority	
Statistics	25–34	35–44	25–34	35–44
Minimum	−5.8790	−7.7496	−8.6340	−15.2350
Q1	−2.8940	−2.8855	1.8780	1.7260
Median	−1.0970	−1.5342	2.8410	2.7780
Mean	−1.4760	−1.8755	3.8840	3.3650
Q3	−0.1470	−0.4855	4.0940	4.6510
Maximum	5.9380	0.3387	42.1580	23.9390
	White		**Minority**	
Statistics	45–54	55–64	45–54	55–64
Minimum	−7.9965	−5.7234	−19.0710	−49.6250
Q1	−3.1342	−3.0902	1.4290	1.6230
Median	−1.9457	−1.4983	2.7240	2.7680
Mean	−2.0722	−1.9419	3.2610	2.2560
Q3	−0.4954	−0.7716	4.4760	4.8450
Maximum	0.4903	0.2491	35.6930	29.4280

TABLE 5: State estimates of betas for each cohort

an exogenous shock to Minority share may result in a complete reversal of population dynamics, and a new trend towards a low or high stable equilibrium.

To better understand where Texas neighborhoods are trending towards, I calculate the difference between a tract's Minority share and its tipping point as $D = s_{j,t} - s*$ where $s*$ is the location's estimated tipping point. The results of this exercise are shown in Figure 8. Most tracts see Minority shares near their estimated tipping points in 2020. This means exogenous shocks are playing a large role in these neighborhoods. My model predicts in the absence of these shocks, neighborhoods would trend away from their tipping points and towards stable equilibria. Figure 9 displays neighborhoods distance to their tipping points for Dallas County. Observing such differences graphically illustrates predictions for what areas of a city may see reversal of trends in Minority, and how close those locations are to this possibility. A location with a difference near zero could see the trajectory of its population change with the addition of a shock to Minority shares brought about by policy or outside investment. The literature on urban density and wages notes spatial concentration of economic activity in urban centers (Krugman [1991]). These dense locations for the city of Dallas appear on the northern half of the city center. Human capital spillovers detailed in seminal works like Moretti [2004] therefore would be predicted to disproportionately benefit areas that are trending towards low Minority shares[9].

[9] In the appendix, I include a replication of Card et al. [2008] for Texas in Figure 11 and the results of this paper's approach for Texas in 12.This illustrates a stark difference in granularity each approach provides. Furthermore, Card's approach loses granularity within cities, where it matters the most

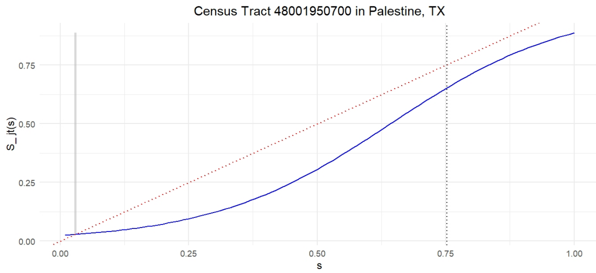

FIGURE 2: 1 Equilibrium, Low Crossing

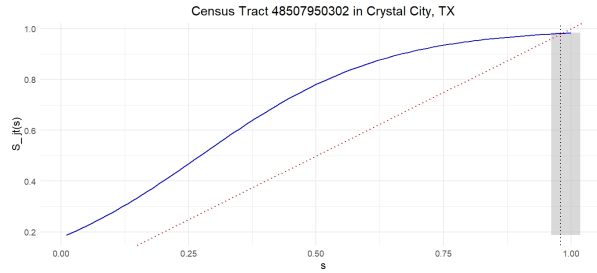

FIGURE 3: 1 Equilibrium, High Crossing

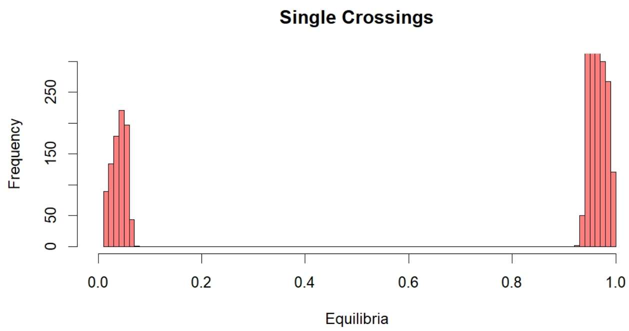

FIGURE 4: Distribution of Single Crossing Equilibria

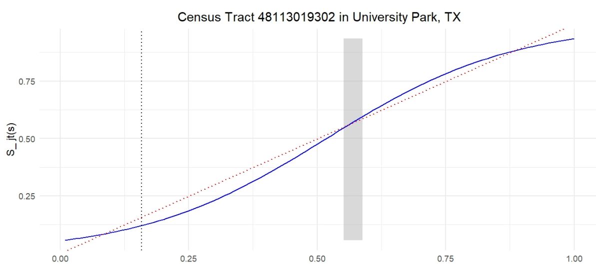

FIGURE 5: 3 Equilibrium, Low Minority Share

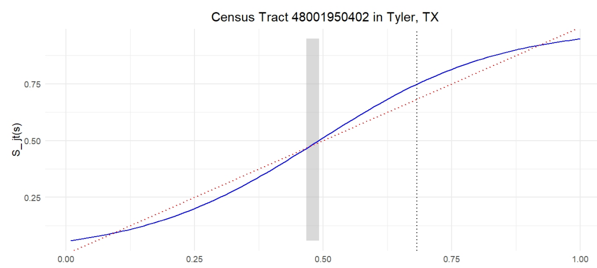

FIGURE 6: 3 Equilibrium, High Minority Share

5 Conclusion

This paper contributes to our understanding of neighborhood segregation dynamics by applying a conditional instrumental variables approach to estimate location choice preferences at the census tract level. My results indicate there is strong evidence of within-group preferences in location choice across all demographic groups in Texas, with White populations showing slightly stronger responses to neighborhood racial composition than Minority populations. These preferences are robust to various specifications and controls, including geographic fixed effects and socioeconomic factors. Estimation of other state coefficients reveal similar patterns. These coefficients are correlated with searches associated with harmful rhetoric and anti-minority sentiment.

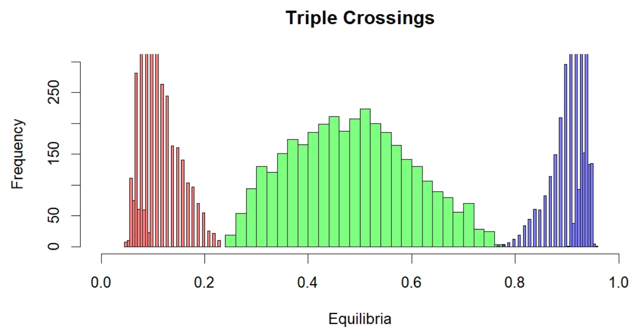

FIGURE 7: Distribution of triple crossing equilibria

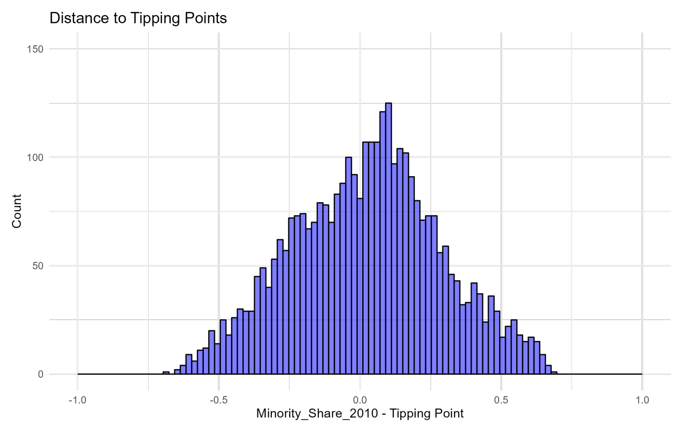

FIGURE 8: Distribution of D

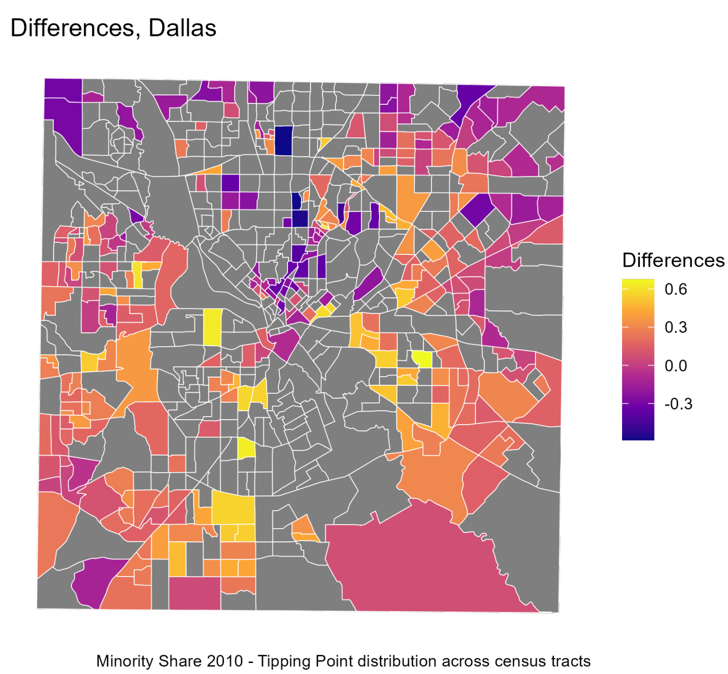

FIGURE 9: Map of D in Dallas

Simulations using these estimates of neighborhood population dynamics reveals three distinct patterns of equilibria: stable low Minority share, stable high Minority share, and triple-crossing points with unstable middle equilibria. The distribution of these equilibria suggests that many Texas neighborhoods are near potential tipping points, making them susceptible to demographic shifts from exogenous shocks to Minority shares. This finding has particular significance for urban policy, as it indicates that targeted interventions could have significant effects in neighborhoods near their tipping points. The granular nature of this analysis at the census tract level provides several advantages over previous MSA-level studies. Most notably, it allows for the identification of within-city variation in segregation dynamics, revealing how different neighborhoods within the same metropolitan area may be trending toward different equilibria. My analysis Dallas County illustrates this point, showing how neighborhoods trending toward lower Minority shares often coincide with areas of higher economic activity and potential wage premiums from human capital spillovers.

A significant limitation of this study is a lack of direct observations of amenities. Direct observation of amenities would allow for alternative instrument construction, and remove the need for many robustness checks. The results of this paper are valid in the context of the modern US. Sample selection and modeling choices were made with this setting in mind. Future work should look to expand this method to other areas and more directly model amenities.

These findings have important policy implications. First, I posit that anti-segregation policies may be most effective when targeted at neighborhoods near their tipping points, where small interventions could prevent neighborhoods from tipping toward a new equilibrium. Second, the correlation between areas trending toward lower Minority shares and economic opportunity highlights the need for policies that ensure equal access to high-productivity urban areas.

The persistence of residential segregation remains a critical policy challenge in American cities. This paper's findings suggest that understanding the granular dynamics of neighborhood change is crucial for designing effective interventions. Without such interventions, the self-reinforcing nature of location preferences may continue to perpetuate patterns of residential segregation and unequal access to economic opportunity.

References

Rodrigo Adao, Costas Arkolakis, and Federico Esposito. Spatial linkages, global shocks, and local labor markets: Theory and evidence. 2019.

Gabriel M Ahlfeldt, Stephen J Redding, Daniel M Sturm, and Nikolaus Wolf. The economics of density: Evidence from the berlin wall. *Econometrica*, 83(6):2127–2189, 2015.

Lina Aldén, Mats Hammarstedt, and Emma Neuman. Ethnic segregation, tipping behavior, and native residential mobility. *International Migration Review*, 49(1):36–69, 2015.

Anders Böhlmark and Alexander Willén. Tipping and the effects of segregation. *American Economic Journal: Applied Economics*, 12(1):318–347, 2020.

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

David Card and Alan B Krueger. School quality and black-white relative earnings: A direct assessment. *The quarterly journal of Economics*, 107(1):151–200, 1992.

David Card, Alexandre Mas, and Jesse Rothstein. Tipping and the dynamics of segregation. *The Quarterly Journal of Economics*, 123(1):177–218, 2008.

Timothy G Conley. Gmm estimation with cross sectional dependence. *Journal of econometrics*, 92(1):1–45, 1999.

Rebecca Diamond. The determinants and welfare implications of us workers' diverging location choices by skill: 1980–2000. *American Economic Review*, 106(3):479–524, 2016.

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

Lance Freeman. Displacement or succession? residential mobility in gentrifying neighborhoods. *Urban affairs review*, 40(4):463–491, 2005.

Paul Krugman. Increasing returns and economic geography. *Journal of political economy*, 99(3):483–499, 1991.

John R Logan and Brian J Stults. Metropolitan segregation: No breakthrough in sight. Technical report, Brown University, 2022.

Steven Manson, Jonathan Schroeder, David Van Riper, Tracy Kugler, and Steven Ruggles. National historical geographic information system: Version 19.0, 2024. URL http://doi.org/10.18128/D050.V19.0. Minneapolis, MN: IPUMS.

Harvey J Miller. Tobler's first law and spatial analysis. *Annals of the association of American geographers*, 94(2):284–289, 2004.

Enrico Moretti. Human capital externalities in cities. In *Handbook of regional and urban economics*, volume 4, pages 2243–2291. Elsevier, 2004.

Thomas C Schelling. Dynamic models of segregation. *Journal of mathematical sociology*, 1(2):143–186, 1971.

Rajiv Sethi and Rohini Somanathan. Inequality and segregation. *Journal of Political Economy*, 112(6): 1296–1321, 2004.

Seth Stephens-Davidowitz. The cost of racial animus on a black candidate: Evidence using google search data. *Journal of Public Economics*, 118:26–40, 2014.

Charles M Tiebout. A pure theory of local expenditures. *Journal of political economy*, 64(5):416–424, 1956.

U.S. Census Bureau. Geography program: Glossary, April 2022. URL https://www.census.gov/programs-surveys/geography/about/glossary.html. Last revised April 11, 2022.

U.S. Department of Housing and Urban Development. Affirmatively furthering fair housing (AFFH), 2023. URL https://www.hud.gov/AFFH. Accessed: November 4, 2024.

Jeffrey M Wooldridge. *Econometric analysis of cross section and panel data*. MIT press, 2010.

Junfu Zhang. Tipping and residential segregation: a unified schelling model. *Journal of Regional Science*, 51(1):167–193, 2011.

Statistic	All 2020	Min. Share 2010	log Min. 45-54 2000	log White 25-34 2010
Min	1.00	0.0000	−34.5388	−34.5388
Q1	28.00	0.2784	−0.7349	0.2555
Median	52.55	0.4937	0.4055	1.2684
Mean	88.42	0.5268	−0.4444	0.8192
Q3	98.11	0.7857	1.3797	2.1017
Max	4868.20	1.0000	5.6331	6.5550

Table 6: Summary statistics by demographic group and year

6 Appendix

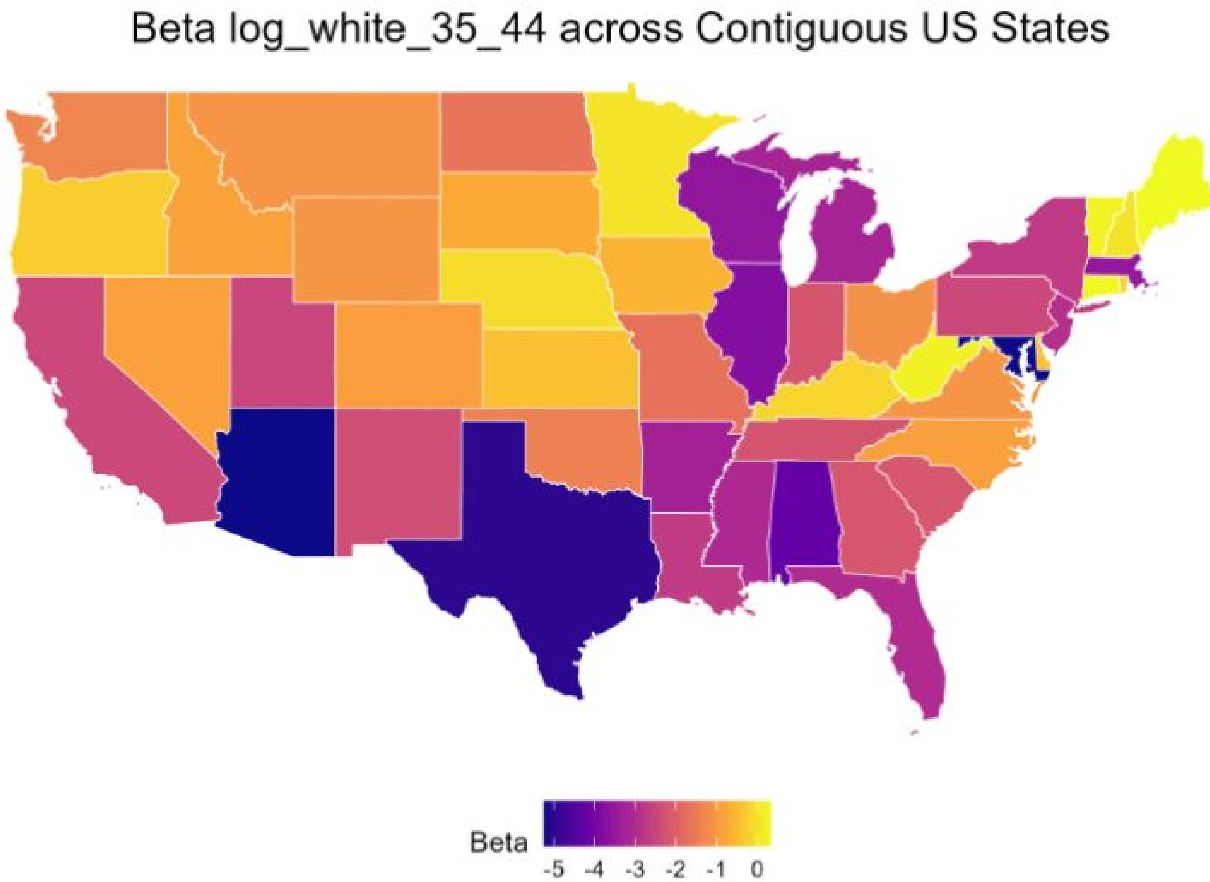

Figure 10: Beta estimates for every state

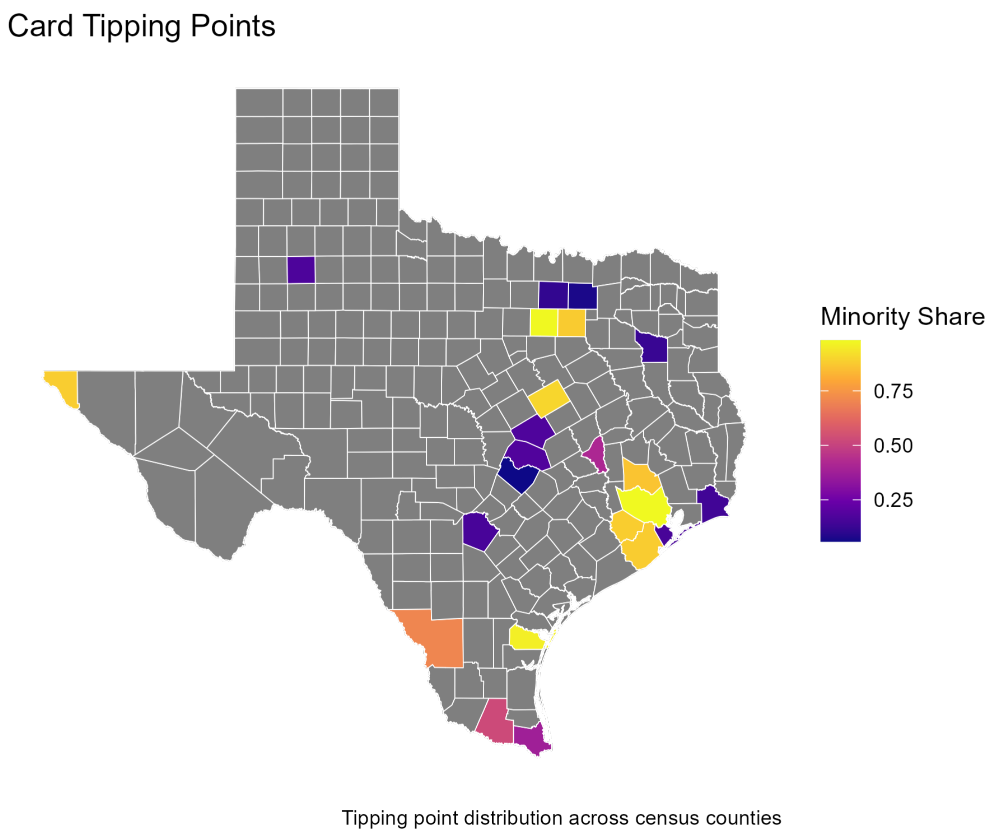

FIGURE 11: Card replication results for large Texas metropolitan areas

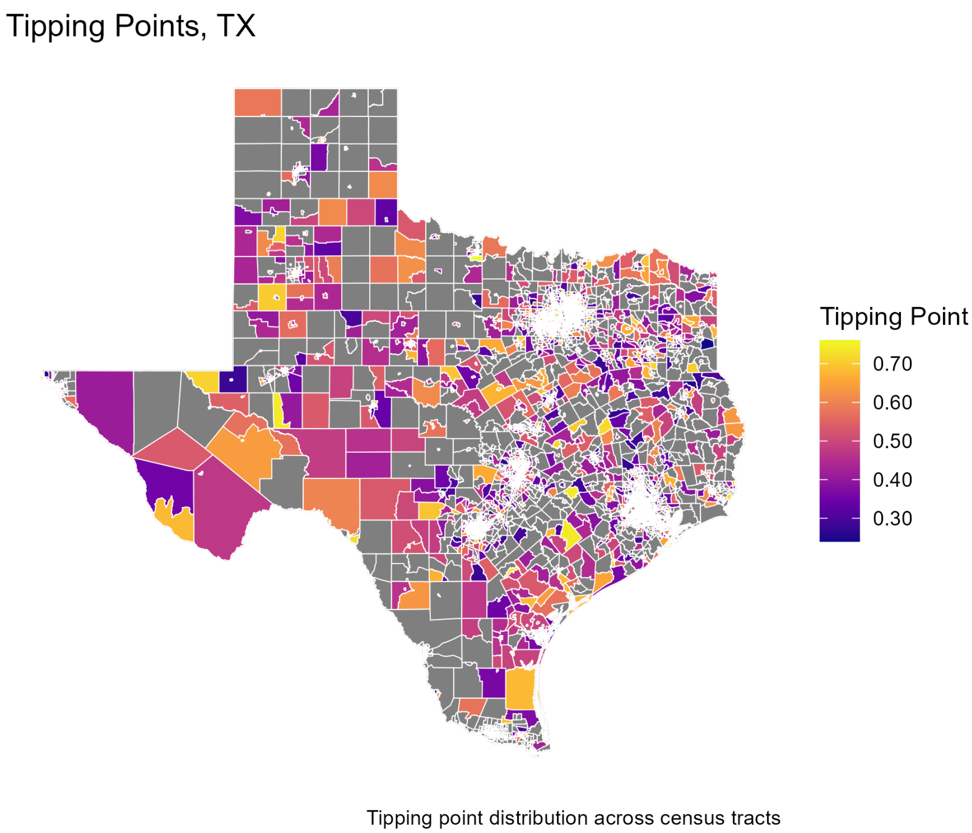

FIGURE 12: Tract estimates for Texas