Universidad del Rosario
Facultad de Economía

Racial and spatial interaction for neighborhood dynamics in Chicago

Carlos Castro
Cristhian Rodríguez

SERIE DOCUMENTOS DE TRABAJO

No. 189

Mayo de 2016

Racial and spatial interaction for neighborhood dynamics in Chicago*

Carlos Castro† Cristhian Rodriguez ‡

March 2, 2017

Abstract

We look at the empirical validity of Schelling's models for racial residential segregation applied to the case of Chicago. Most of the empirical literature has focused exclusively the single neighborhood model, also known as the tipping point model and neglected a multi-neighborhood approach or a unified approach. The multi-neighborhood approach introduced spatial interaction across the neighborhoods, in particular we look at spatial interaction across neighborhoods sharing a border. An initial exploration of the data indicates that spatial contiguity might be relevant to properly analyze the so call tipping phenomena of predominately non-Hispanic white neighborhoods to predominantly minority neighborhoods within a decade. We introduce an econometric model that combines an approach to estimate tipping point using threshold effects and a spatial autoregressive model. The estimation results from the model disputes the existence of a tipping point, that is a discontinuous change in the rate of growth of the non-Hispanic white population due to a small increase in the minority share of the neighborhood. In addition, we find that racial distance between the neighborhood of interest and the surrounding neighborhoods has an important effect on the dynamics of racial segregation in Chicago.

*The authors would like to thank Eugenio Merlano for excellent research assistance at the initial phase of the project.

†Universidad del Rosario. email: carlos.castro@urosario.edu.co.

‡Universidad del Rosario.

JEL Classification: J15, R23.
Keywords: racial residential segregation, tipping point, spatial econometrics, Chicago.

1 Introduction

Does spatial interaction play a role in explanting residential segregation in Chicago? Many metropolitan areas in the United States display substantial racial segregation and substantial variation in incomes and house prices across neighborhoods. Racial segregation, in particular, is a striking trait of U.S. cities. Iceland, Weinberg, and Steinmetz (2002) report that 64% of the black population would have needed to change residence for all U.S. neighborhoods to become fully integrated in the year 2000.
Perhaps one of the most well known theories to explain the dynamics of US cities with respect to racial segregation is Schelling's bounded-neighborhood model, also known as neighborhood tipping point model. In the tipping point model, an all white neighborhood, where households have heterogenous preference over minorities begins a process of evolving in to a full minority neighborhood after an initial small number of minorities households moves into the neighborhood. Once the process is started (with a small initial shock) the share of minorities is above the tipping point and the process is irreversible. Therefore, tipping point is identified at the level where the share of minorities in a given neighborhood creates a clear break with respect to the growth of the white population.
The literature contains two explicit empirical assessments of Schelling's racial tipping point model. On one hand, Card et al. (2008) finds strong evidence of tipping behavior in US cities, choosing Chicago a one of the prime examples. In the paper they estimate differences in tipping point in other US cities as well as the covariates that determine these differences. Tipping points are estimated between 5% to 20% of the minority share. On the other hand, Easterly (2009) presents a negative view, concluding that tipping is not an important driver of neighborhood change in the data. Easterly (2009) finds evidence of a tipping point only in a small subset of neighborhoods with low population density and a small minority share. Most recently, Blair (2016) provides a method for estimating census tract tipping points and updates the estimates finding that tipping points have increase from 13% in 1970 to 42% in 2010. Furthermore, he shows that the tipping behavior is more correlated

to outside options for white household than difference in racial attitudes. From the empirical debate we could distinguish two different phenomena: an increase in white neighborhoods in the suburbs that is not necessarily racially driven and an spanning predominately minority neighborhoods where tipping behavior is more likely to be the dominant factor. It is important to understand the extend of both phenomena to ascertain the empirical validity of low tipping point behavior identified by Card et al. (2008) for the city of Chicago.

From a theoretical point of view and a number of Schelling's publications, he postulates two models for explaining segregation, that introduce the notion of a tipping point like behavior: the first model for segregation dynamics is based on a single neighborhood, whereas a second model is in a multi-neighborhood setting (checkerboard model). The checkerboard model is based on an initial set of preference over the racial composition of the immediate neighbors for a grid of white and minority agents. A simulation begins by an initial choice of new residential location of the agents that are dissatisfied with their initially assigned location. The results from the model show that even though agents have moderate preference for same color neighbors, the outcome after the simulation has ended so that no agent is found to be discontent with his residential choice, is complete residential segregation. The similarity with respect neighborhood tipping point model is that a small initial shock can lead to complete segregation.

Zhang (2011) proposes a unified framework that theoretically encompasses both approaches, a unified Schelling model. The model is based on checkerboard model and rigorously introduces the idea of tipping in this multi-neighborhood setting. The key insights form both models is that it does not take extreme preferences over the racial composition of neighbors or that a small initial shock (such as an increase in the income of minorities) combined with heterogenous racial preferences over racial composition within a neighborhood, to produce complete segregation. Zhang (2011) uses these insight to develop a mathematical model that indicates that the outcome of complete segregation can be considered as a coordination problem due to the fact that when an agent decides to move out of a predominantly white neighborhood he does not take into account the effect on his neighbors preference over racial composition of the neighborhood. This in turn affect other agents decision to move an creates an undesirable equilibrium outcome. With a multi-neighborhood setting spatial interaction become a dominant trait in the dynamics that has not been entirely explored at the empirical level.

Whereas most of the empirical attention has centered in the single neighborhood case Card et al. (2008) and Easterly (2009), the most recent literature introduces a multi-neighborhood setting by considering an explicit outside option based on all other neighborhoods within the metropolitan area, Blair (2016).

In this paper we will explore further the idea of tipping in a multi-neighborhood as in the theoretical model proposed by Zhang (2011) by first revisiting the data used by Card et al. (2008) for the case of Chicago and second formulating an econometric model that is able to estimate a tipping point and control for spatial interaction between adjacent tracts.

We propose an econometric model that incorporates a threshold effect on a standard spatial autoregressive model, this allows ut to control for spatial dynamics between contiguous neighbors while estimating an average tipping point for a metropolitan area using census tract information from 1970 to 2000. We estimate the model using maximum likelihood by deriving the concentrated maximum likelihood in terms of the threshold parameter (the tipping point) and the spatial autoregressive parameter.

Before estimating the model we look a the same data used by Card et al. (2008) for the case of Chicago and find that the outcome data that determine the low tipping behavior 5% in 1970-1980 and (12%) for afterwards $(1980, 1990, 2000)$ is driven by a strong growth in low density neighborhoods with low or non-existent minorities and sometimes spatially disconnected to highly minority neighborhoods. If we remove this low density neighborhoods or control for spatial interaction the tipping point becomes statistically weak and increases to around 75% to 95%. Furthermore, we also find that indicators of racial distance measured with respect the racial composition of neighborhood sharing a border are an important factor to determine the racial dynamics of the Chicago metropolitan area. This is also consistent with the dynamics of tipping behavior identified in racial dynamics for Chicago neighborhoods when we look at maps and ten year transition probabilities. We look at transitions from highly white neighborhoods to highly minority neighborhoods, that is tipping behavior, and find that this phenomena is strongly related to racial composition of the surrounding neighborhoods.

The rest of the document proceeds as follows: In section 2 we recall the main findings of Card et al. (2008) with respect to Chicago, in particular we include in the analysis the population density variable. Furthermore we look at the index of dissimilarity to understand the spatial dynamics of the two metropolitan areas. In section 3 we introduce the spatial autoregressive model with threshold effects (SAR-T) that we will use for estimation. In section 4 we present the empirical application and estimation of the tipping points. Finally we conclude in section 5.

2 Spatial dynamics in the Chicago metropolitan area

Card, Mas and Rothstein (2008) find evidence of a discontinuity in the relationship between lagged fraction minority and change in fraction minority in US neighborhoods. They interpret this finding as evidence in favor of a particular variety of“tipping point” models of segregation. In such models, the dynamics of the fraction minority are determined by the lagged fraction minority together with other characteristics of the community (e.g. average income). In figure 1, we replicates the case for Chicago as presented in the original paper. The vertical line identifies the point of discontinuity around 4.5% the minority share. This tipping point is consistent with the empirical results obtained form two different methods a threshold based method and a fixed point method that is applied in their paper. However, what is not evident from the Card et al. (2008) but is pointed out by Easterly (2008) it that the tipping point is identified by the high growth rate of low density neighborhoods. In figure 2, we make a distinction based on census tracks that had fewer than 5,000 inhabitants per square mile by a red-x marker as opposed to all other tracks. The figure shows that those track driving the so call discontinuity are those that have a low population density and are predominantly white.

We now look a the visual neighborhood dynamics: we look at the growth rates of white population (figure 3), the population density (figure 4), and the change in the racial composition of neighborhoods (figure 5) over each ten year period from 1970 to 2000. From figures 4 and 3 it is clear that the majority of neighborhoods with a strong growth of the white population is in the suburbs of the metropolitan areas where population density in the previous decade is low.

To look at the racial composition of neighborhoods, following Badel and Martinek(2011), we classify tracks for each year that we have information on the Neighborhood Change Database (NCDB), that is 1970, 1980, 1990 and 2000.

The classification is based on the minority share of the population and the share of the non-Hispanic white population, at the level of the neighborhood (census track) and with respect to the metropolitan area shares. For a given minority share of the population for the metropolitan area, if the neighborhood minority share is above (below) the metropolitan area minority share then the neighborhood is classified as High Minority (Highly White). Using the previous classification for all census tracks, that is either Highly White (HW_t) tracks in a particular year t of High Minority (HM_t), we look at the same classification after ten years, that is (HW_{t+10}) and (HW_{t+10}). We also define a similar classification to identify track with a very low population density and we denote these tracks as vacant tracks (V_t)
We can name each kind of change using popular terminology:

- White Resegregation, track that stay Highly White: $HW_t \Rightarrow HW_{t+10}$.
- Minority Resegregation, track that stay Highly Minority: $HM_t \Rightarrow HM_{t+10}$.
- Tipping White to Minority, track that switched from HW to HM:$HW_t \Rightarrow HM_{t+10}$.
- Tipping Minority to White, track that switched from HM to HW:$HM_t \Rightarrow HW_{t+10}$.
- White Sub-urbanization, track that switched from vacant to HW:$V_t \Rightarrow HW_{t+10}$.
- White Depopulation, track that switched from HW to vacant:$HW_t \Rightarrow M_{t+10}$.
- Minority Depopulation, track that switched from HM to vacant:$HM_t \Rightarrow V_{t+10}$.

When we look at the evolution of the tracks (in figure 5) the important phenomena that we are able to identify is that in the particular case of tipping from predominantly white neighborhoods to predominantly minority neighborhoods it is important to account for the evolution of the neighboring tracks, that is there is a spatial continuity aspect that could determine whether there is a tipping or not. This idea of tipping is more related to a multi-neighborhood set up rather than exclusively accounting for the minority share in each single neighborhood.

To provide further evidence we build unconditional transition probability matrices based on the classification of tracks according our initial three state-space classification: HW, HM and V (Table 1). In addition we build transition probability matrices that are conditional on the racial configuration of the neighboring tracts, that is on the number of tracts that share a border that are HW, HM and V (figure 2). The unconditional probabilities consistently show a larger probability of neighborhood tipping from white to minority (11% to 34%) than from minority to white (1.4% to 5.3%). On the other hand there is not clear consensus with respect to sub-urbanization and race from the transition probabilities. Vacant tracts are not necessarily more likely to become predominantly white or minority. This phenomenon seems clear in the long run thirty years but not when we look at decennial transition probabilities.

Form the conditional transition probability matrices we find significant evidence that tipping across racial lines seems to be determined by the racial configuration of neighboring tracks. For example, the decennial probabilities indicate that the probability of tipping from white to minority almost doubles when highly white tracks have more than one highly minority neighbor. However, for highly minority neighborhoods sharing having more than one predominantly white tracts sharing a border does not really have a significant effect on their racial dynamics.

Both the condition transition probabilities (table 2) and the change in the racial configuration of the tracts (figure 5) indicate that the racial distance between neighborhoods (defined as census tracts) that share a physical border are important determinants of the racial dynamics in the city of Chicago. This evidence is more in line with the multi-neighborhood explanation of Schelling or the unified Schelling model proposed by Zhang (2011) rather than looking at the empirical validity of the single neighborhood tipping model explored by Card et al. (2008) and Easterly (2008). Using the information of the racial configuration of neighboring census tracts be build a indicator of racial distance for each census track. The racial distance between neighboring census tracts is the difference between the minority share, for highly white tracts, of a given track minus the maximum minority share across all neighboring tracks that share a border with the census tract of interest. For highly minority tracts the difference is with respect to the share of non-Hispanic whites. For both predominantly white or minority neighborhoods the indicator should be around zero if the census track is surrounded by other tracks with more or less the same racial configuration. If we look at the growth rate for the non-Hispanic white population between 1970 and 1980, as Card et al. (2008) did, but we remove all census tracts with a low population density and instead of looking at the minority share of the own neighborhood we consider our indicator or racial distance. We find that for highly white neighborhoods (blue dots) the higher the racial distance between its neighbors (that is the lowest value of the racial distance indicator) then there is a stronger decrease in the non-Hispanic white population in that tract (figure 6). In contrast, although highly white neighborhoods with population density above 5000 per sq. mile and a minimal or zero racial distance with respect to their neighboring census tracts, are also decreasing in terms some race population but at a much slower rate (we have already mentioned that during the same period the only predominantly white neighborhoods that had positive population growth rates are low population density tracts in 1970). For the non-Hispanic white population in predominantly minority neighborhoods the dynamics is more or less the same but the change in the growth rates is smaller.

3 A spatial autoregressive model with threshold effects on the covariates (SAR-T)

3.1 Tipping point estimation with a threshold model

The main stylized fact that you want to capture in the Card, Mas and Rothstein (2008) is the discontinuity of outflows of non-hispanic white population in urban areas. Thus giving a possible explanation to the dynamics of segregation within in US cities. The main driving force of the outflow is the percentage of minorities within each census track (unit of analysis, i). According to the economic model there is a critical level of minorities that the white population is willing to "tolerate" within their surroundings. The percentage of minorities may change because of exogenous shocks (income). This critical level can be better interpreted as a threshold point above which there is a strong incentive to leave a particular neighborhood.

In the econometric specification the outflow is measured as $\Delta w_{i,t}$ the ten year change in the neighborhoods white population taken as a share of the initial (total) population. According to the economic model this change is explained fundamentally by non-linear relationship between the based year (lagged) minority share and the threshold (tipping point), $f(m_{i,t-10}, m^*)$, as well as a set of other lagged explanatory variables (controls) $x_{i,t-10}$. This relationship is comprised of a smooth function of $(m_{i,t-10})$ except at the tipping point.

$$\Delta w_{i,t} = f(m_{i,t-10} - m^*) + d\mathbf{1}_{m_{i,t-10}>m^*} + \gamma \mathbf{x}_{i,t-10} + \varepsilon_{i,t}$$

Note that the smooth function tries to capture the dynamics both before and after the tipping point with an additional level effect after the tipping point. According to the data we should expect a large (negative) level structural change just after tipping point that levels out strongly (toward zero) after the minority share in the neighborhood is above 60% (see figure 1).
The identification and estimation of the tipping point m^* as well as the other parameters in the model is done using two approaches. The first related to the approaches used in structural change in time series and in particular to the literature associated with threshold models (Hansen, 1999,2000; Hansen et al. 2004; Chan 1993). The authors however point out that this approach works well for large cities but not smaller ones where the effect of outliers can be severe on the identification of the threshold (m^*). The second approach exploits the stylized facts on the data, in particular the smooth varying relationships between $\Delta w_{i,t}$ and $m_{i,t-10}$ before and after the tipping point. They look for a "fixed point" defined as the minority share at which the neighborhood white population grows at the average rate of the city.
The threshold approach is the most common and the most well documented. For estimation of the tipping point and the introduction of spatial effects in the sequel we consider a different set up that is fully parametric instead of the semi-parametric specification in expression 3.1. Instead of considering a smooth non-linear function between the based year (lagged) minority share and the threshold (tipping point), $f(m_{i,t-10}, m^*)$, we consider a simple linear functional

$$\Delta w_{i,t} = \alpha_0 + \alpha_m \mathbf{1}_{m_{i,t-10}>m^*} + \beta_0 m_{i,t-1} + \beta_m m_{i,t-1} \mathbf{1}_{m_{i,t-1}>m^*} + + \gamma \mathbf{x}_{i,t-10} + \varepsilon_{i,t}.$$

The parameter set of such model $(\alpha_0, \alpha_w, \beta_0, \beta_w, \gamma, m^*)$ are estimated sequentially using LS and 2SLS or GMM. We perform estimation by maximum likelihood, because this makes the transition simpler to a setup with spatial effects.

3.2 Introducing spatial effect in the model

Section 2 shows there might be a missing part to the story that cast doubt the tipping point phenomena. We provide evidence from the data that migration from predominantly white neighborhoods to low population density or vacant tracts and/or a combination of outside options for incumbent white residents might be misleading with respect to the evidence for small tipping points. This evidence is consistent with a recent paper by Blair (2016).

This other part is a so-called spatial effect, where borderline neighborhoods can have an effect on the dynamics. For example, a white neighborhood may be swallowed up by the surrounding minority neighborhoods or a cluster of white neighborhoods might delay the tipping point phenomena. The empirical challenge is to come up with a econometric specification and estimation method that allows us to control for spatial effects.

$$\Delta w_{i,t} = (\alpha_0 + \alpha_m \mathbf{1}_{m_{i,t-1}>m^*}) + (\beta_0 + \beta_m \mathbf{1}_{m_{i,t-1}>m^*}) m_{i,t-1} + \\ \rho W \Delta w_{i,t-1} + \gamma \mathbf{x}_{i,t} + \varepsilon_{i,t}.$$

where W is an nxn symmetric, non-negative scaled spatial weight matrix and $\varepsilon_{i,t} \sim N(0, \sigma^2)$. The advantage of this setup is that the parameters $(\alpha_0, \alpha_m, \beta_0, \beta_m, \gamma, m^*, \rho)$ will tell us the relative importance of the lagged effect (inclusive of the structural change due to the percentage of white population) versus the spatial effect.

3.3 Estimation strategy for the SAR-T

We follow the considerations Card, Mas and Rothstein (2008). Since we are using census track data n is very large but T is quite small (1970, 1980, 1990, 2000), so considering our data set as a panel might put too much strain on the estimation strategy. Therefore, we take each effective year (1980, 1990, 2000, because of the predetermined lagged variables) as a separate data set. For this reason the indexing of the variables should not give the interpretation of a panel (henceforth for simplicity we eliminate the time index).
The model can be written as a spatial autoregresive model (SAR) in vector form (LeSage and Pace, 2009),

$$\Delta \mathbf{w} = \rho W \Delta \mathbf{w} + \mathbf{X}\gamma + \mathbf{Z}\delta + \varepsilon \tag{1}$$

where $\mathbf{Z} = [\mathbf{1} \quad \mathbf{1}_{m_i>m^*} \quad \Delta \mathbf{w} \quad \Delta \mathbf{w}_{m_i>m^*}]$ and $\delta = [\alpha_0 \quad \alpha_m \quad \beta_0 \quad \beta_m]'$. If the population value of the parameters ρ and m^* is known to be ρ^* and $m^\dagger$, respectively, we can rearrange the model as follows,

$$
\begin{aligned}
\Delta\mathbf{w} - \rho W \Delta\mathbf{w} = \mathbf{X}\gamma + \mathbf{Z}\delta + \varepsilon \\
(\mathbf{I}_n - \rho^* W)\Delta\mathbf{w} = \mathbf{X}\gamma + \mathbf{Z}(m^\dagger)\delta + \varepsilon \\
\Delta\mathbf{w} = (\mathbf{I}_n - \rho^* W)^{-1}(\mathbf{X}\gamma + \mathbf{Z}(m^\dagger)\delta) + (\mathbf{I}_n - \rho^* W)^{-1}\varepsilon \\
\varepsilon \sim N(\mathbf{0}, \sigma^2 \mathbf{I}_n)
\end{aligned}
$$

where can denote $\Delta\mathbf{w}(\rho^*) = (\mathbf{I}_n - \rho^* W)\Delta\mathbf{w}$ as our new dependent variable. Note that $\Delta\mathbf{w}(\rho^*)$ and $\mathbf{Z}(m^\dagger)$ are functions of the known parameters ρ^* and $m^\dagger$.

Estimation of spatial models via least squares can lead to inconsistent estimates of the regression parameters for models with spatially lagged dependent variables, inconsistent estimation of the spatial parameters and inconsistent estimation of standard errors. In contrast maximium likelihood is consistent for these models (LeSage and Pace, 2009). The estimation strategy is to concentrate the full (log Gaussian) likelihood with respect to the parameters γ, δ and σ^2, and reduce the maximum likelihood to a bi-variate optimization problem in the parameters ρ and m^*. In what follows we will sequential concentrate the parameters of the model,

$$\Delta\mathbf{w}(\rho^*) = \mathbf{X}\gamma + \mathbf{Z}(m^\dagger)\delta + \varepsilon$$

The first step involves the use of the Frisch-Waugh theorem (partitioned regression) in order to concentrate out γ. For this particular model the estimating equations, for δ and γ are as follows,

$$\begin{pmatrix} \mathbf{X}'\mathbf{X} & \mathbf{X}'\mathbf{Z}(m^\dagger) \\ \mathbf{Z}(m^\dagger)'\mathbf{X} & \mathbf{Z}(m^\dagger)'\mathbf{Z}(m^\dagger) \end{pmatrix} \begin{pmatrix} \gamma \\ \delta \end{pmatrix} = \begin{pmatrix} \mathbf{X}'\Delta\mathbf{w}(\rho^*) \\ \mathbf{Z}(m^\dagger)'\Delta\mathbf{w}(\rho^*) \end{pmatrix}$$

The first of these equations can be re-written as follows, in order to obtain an expression for $\hat{\gamma}$,

$$\mathbf{X}'\mathbf{X}\gamma + \mathbf{X}'\mathbf{Z}(m^\dagger)\delta = \mathbf{X}'\Delta\mathbf{w}(\rho^*) \tag{2}$$

$$\mathbf{X}'\mathbf{X}\gamma = \mathbf{X}'\Delta\mathbf{w}(\rho^*) - \mathbf{X}'\mathbf{Z}(m^\dagger)\delta \tag{3}$$

$$\hat{\gamma} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'(\Delta\mathbf{w}(\rho^*) - \mathbf{Z}(m^\dagger)\delta) \tag{4}$$

We now have an expression for $\hat{\gamma}$ as a function of δ (an unknown) and the known parameters ρ^* and $m^\dagger$. Now we introduce expression 4 into the second estimating equation,

$$\mathbf{Z}(m^\dagger)'\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\Delta\mathbf{w}(\rho^*) - \mathbf{Z}(m^\dagger)'\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Z}(m^\dagger)\delta + \mathbf{Z}(m^\dagger)'\mathbf{Z}(m^\dagger)\delta = \mathbf{Z}(m^\dagger)'\Delta\mathbf{w}(\rho^*)$$
$$\mathbf{Z}(m^\dagger)'\underbrace{(\mathbf{I}_n - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}')}_{\mathbf{M}_X}\mathbf{Z}(m^\dagger)\delta = \mathbf{Z}(m^\dagger)'\underbrace{(\mathbf{I}_n - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}')}_{\mathbf{M}_X}\Delta\mathbf{w}(\rho^*)$$
$$\hat{\delta} = (\mathbf{Z}(m^\dagger)'\mathbf{M}_X\mathbf{Z}(m^\dagger))^{-1}\mathbf{Z}(m^\dagger)'\mathbf{M}_X\Delta\mathbf{w}(\rho^*) \quad (5)$$

We now have an expression for $\hat{\delta}$ as a function of the known parameters ρ^* and $m^\dagger$ and the known projection matrix $\mathbf{M}_X$. From 4 and 5 we can now obtain an unbiased estimate for the residual vector,

$$\hat{\varepsilon} = \Delta\mathbf{w}(\rho^*) - \mathbf{X}\hat{\gamma} - \mathbf{Z}(m^\dagger)\hat{\delta} \quad (6)$$

From now on we will make explicit (in the notation) that the residual vector $(\hat{\varepsilon}(\rho^*, m^\dagger))$ is a function of the known parameters ρ^* and $m^\dagger$. Furthermore, since we have a Gaussian likelihood the estimate for the noise variance parameter is $\hat{\sigma}^2 = \frac{\hat{\varepsilon}(\rho^*, m^\dagger)'\hat{\varepsilon}(\rho^*, m^\dagger)}{n}$.
The full log-likelihood function for the SAR model (1) takes the form,

$$lnL = -\frac{n}{2}ln(2\pi\sigma^2) + ln \mid \mathbf{I}_n - \rho^* W \mid - \frac{\varepsilon'\varepsilon}{2\sigma^2} \quad (7)$$

where $\varepsilon = \Delta\mathbf{w} - \rho W\Delta\mathbf{w} - \mathbf{X}\gamma - \mathbf{Z}\delta$. The parameter driving the spatial dependence is generally consider to take values between zero and one, $\rho \in [0, 1)$. Maximizing the full log-likelihood for the SAR model (lnL) involves setting the first derivatives, with respecto to the parameter of interest $(\gamma, \delta, \sigma^2, \rho$ and $m^*)$, equal to zero and simultaneously solving these first order conditions for all of them. In contrast, a simpler and equivalent (Davidson and MacKinnon, 1993) procedure involves concentrating the log-likelihood with respect to γ, δ and σ^2; using the known closed form solutions for these parameters. The concentrated parameters will be functions of ρ, m^* and the sample data $(4, 5, 6$). The resulting concentrated log-likelihood lnL_c will only be a function of ρ and m^*. Finally, since in this particular case the tipping point, measures the critical point of the minority share, by definition, $m^* \in [0, 1)$. Therefore, we will need to numerically maximize the concentrated log-likelihood over the unit square (figure 9).
The concentrated log-likelihood function is,

$$lnL_c = -\frac{n}{2}(1 + ln(2\pi) + ln(\hat{\varepsilon}(\rho, m^*)'\hat{\varepsilon}(\rho, m^*))) + ln \mid \mathbf{I}_n - \rho W \mid \quad (8)$$

where $\hat{\varepsilon}(\rho, m^*)$ is given by expression 6. Various methods can be used to obtain the maximum likelihood estimates of $\hat{\rho}, \hat{w}^*$ (LeSage and Pace, 2009; Pace and Barry, 1997). Another advantage of using the concentrated log-likelihood is that simple adjustments to the output of the optimization problem can be used to produce a computationally efficient variance-covariance matrix, used in the sequel for inference on the estimated parameters.

4 Empirical results

We estimate the threshold model and spatial autoregressive threshold model and report the estimation results in tables 3 and 4. We find consistent results across decades or the whole sample (thirty year period). Estimation is performed, unless otherwise stated, using the information of the available number of tracks (1802) for the Chicago metropolitan area. Consistent with Card et al (2008) and Blair(2016) there is a increase in the estimated tipping point in each decade 4.4% (1970-1980), 14.3% (1980-1990), 23.1% (1990-2000)[1]. In addition, we find that the estimated tipping point using maximum likelihood for the threshold model is not affected by functional form used fit the trend before or after the tipping point, that is whether we only consider a level shift or a location-scale change[2].

The estimation first set of estimation results (the first column of table 3) is consistent with Card et al.(2008) result and figure 1. There is a strong decrease in the growth of the non-Hispanic white population after the tipping point and the tipping point is around 4.4%. Introducing covariates does not affect the tipping point, there is a slight increase in explained variance and they have the expected sign; that is an increase in average household earnings (in the base year, 1970) of the neighborhood has an increase in the non-Hispanic white population, an increase in the average value of the homes (in the base year, 1970) decreases the rate of growth of the non-Hispanic white population. We include as covariate the racial distance between own neighborhood and the surrounding neighbors (see figure 6). Since a negative distance indicates a higher racial difference for predominantly non-Hispanic

[1] We only report the estimation results for the period 1970-1980 and 1970-2000, however the complete results are available upon request.

[2] We only report the estimation results using a specification that only considerers a level change (a change in the mean before or after the tipping point), however the complete results are available upon request.

white neighborhoods, an increase on the indicator has a large positive effect on the growth rate of the non-Hispanic white population. The third column shows that the results are also consistent for the thirty year period (1970-2000). In terms of the estimation performance figure 7 presente the log likelihood function as a function of the minority share in 1970, and we find a strong argument for the significance of the tipping point at around 5%. In column two the estimation if performed only on the census tracks with population density above 5000 per sq. mile. In this case the tipping point is estimated to be at around 76% and the log likelihood function indicates a local maxima (figure 8) at 5% but a global one some where above 70%, which is less precise but anyway significant. This result confirms the discussion in section 2 and places a shadow doubt on the statistical robustness of Card et al. (2008) result. In other words the strong vertical outliers of low density census tract introduce this strong discontinuity in the data, however this does not really provide conclusive evidence that the decision of non-Hispanic white household is racially drive. There is a racial aspect that we are able to capture through the introduction of the racial distance control variable which is more consistent with the idea of a multi-neighborhood model rather than a single neighborhood one.

Is it possible to empirically test the validity of the multi-neighborhood explanation and compare it to the standard empirical tipping point identification approach? the answer is yes and it is precisely the reason why we specify an spatial autoregressive threshold model (section 3). The estimation results of the model are presented in table 4. One important aspect is that we do not require to select a particular population density cut-off point, that is the results are not sensitive to our previous choice of only keeping for estimation purposes census tracks with population density above 5000 per sq. mile. The estimated spatial lag (0.85) is strongly significant and indicates that neighborhood that share borders (spatial contiguity) are more likely to have similar growth rates for the non-Hispanic white population. Given the dynamics observed in figures (4, 3,5), the spatial effect (average growth rate of the neighboring tracts) becomes a quite important control variable in particular with respect to the low density tracts that are contiguous and in the outer region of the metropolitan area. One of the advantages of the estimation strategy through concentrated maximum likelihood is that visually (figure 9) we can see the effect on the estimation of the tipping point. If we impose that the spatial autocorrelation coefficient is zero, the lower part of the surface is exactly the result we get from the threshold model and that Card et al. (2008) gets, a tipping point of around 5%. However, as the spatial effect becomes larger in magnitude an reaches an area where the log likelihood is maximal, then the tipping point es much larger but at the same time there is a large uncertainty, the surface becomes flat along the tipping point (w^*). In the model that includes spatial effect the control variables, average income of the household, average value of the home and the racial distance are still significant and have the expected sign.

5 Conclusions

Schelling's work on dynamic models of segregation goes beyond the analysis of the single neighborhood, that is his most famous tipping point model. The tipping model generated important interest both theoretically and empirically. However, the validity of the former or the extent to which the tipping hypothesis is an adequate depiction of racial residential segregation in large US cities, like Chicago is not entirely solved. The reason is that the single neighborhood approach ignores the dynamic around the neighborhood or the metropolitan area at large.

Using theoretical results that support multi-neighborhood model consistent with the tipping model, we develop an econometric framework that is able to test for tipping points controlling for spatial interaction. We find in the data evidence along two lines: the tipping point results presented by Card et al. (2008) ignore the effect of population density, if we control for population density the tipping point phenomena, associated to very low migration of minority households to predominantly non-Hispanic white, is not as extreme as initially estimated. On the other hand, the change from predominantly non-Hispanic white neighborhoods to predominantly minority neighborhoods can be explained by racial distance measured with respect to the racial composition of adjacent neighborhoods (census tracts). We do not claim that racial distance is the only determinant but it seems to be more robust than the tipping phenomena associated with a very small increase of minorities and a large flight of non-Hispanic white households from a single neighborhood.

From an econometric point of view our contribution is to provide an spatial autoregressive model that is able to look for structural changes along one dimension of the data and ist relationship with the dependent variable, a threshold model.

Our empirical application focuses exclusively on Chicago and therefore we are not able to say that this is an adequate explanation for all metropolitan areas in the US. We leave the application to other metropolitan areas as a future research agenda.

References

Badel, A., Martinek, C.J.: Black-White segregation in the eight district: a look at the dynamics, *The Regional Economist*, Federal Reserve Bank of St. Louis, July, (2011).

Blair, P.: The Effect of Outside Options on Neighborhood Tipping Points, Working paper, Department of Economics, Clemson University, (2016).

Card, D., Mas A., Rothstein, J.: Tipping and the dynamics of seggregation, *The Quarterly Journal of Economics*, February, (2008).

Caner, C., Hansen, B.: Instrumental Variable Estimation of a Threshold Model, *Economic Theory*, 20, 813-843, (2004).

Chan, K.S.: Consistency and limiting distribution of the least square estimator of a threshold autoregressive model, *Annals of Statistics*, 21, 520-533, (1993).

Davidson, R., MacKinnon, J.G.: Estimation and Inference in Econometrics, Oxford University Press, (1993).

Easterly, W.: Empirics of strategic interdependence: The case of the racial tipping point, *The B.E. Journal of Macroeconomics: contributions*, 9(1), 25, 1-33 (2009).

Hansen, B.: Threshold effects in non-dynamic panels: Estimation, testing and inference, *Journal of Econometrics*, 93, 345-368 (1999).

Hansen, B.: Sample splitting and threshold estimation, *Econometrica*, 68(3), 575-603 (2000).

Iceland, J., Weinberg, D.H., Steinmetz, E.: Racial and Ethnic Residential Segregation in the United States: 1980-2000.U.S. Census Bureau, Series CENSR-3. Washington, DC: U.S. Government Printing Office, (2002); http://www.census.gov/prod/2002pubs/censr-3.pdf.

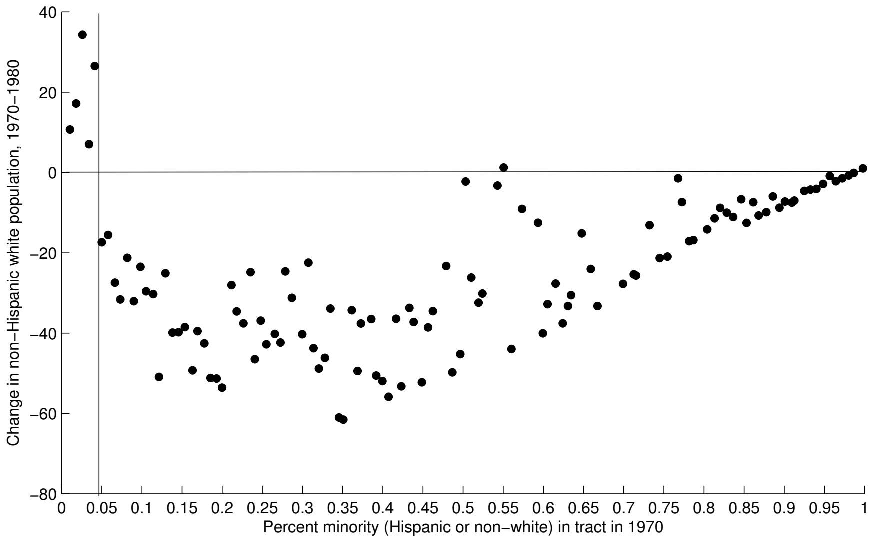
LeSage, J., Pace, K.: Introduction to Spatial Econometrics, Chapman Hall/CRC, (2009).

PACE, R. K., BARRY R.: Quick computation of regression with a spatial autoregressive dependent variable, *Geographical Analysis*, (1997).

ZHANG, J.: Tipping and residential segregation: a unified schelling model, *Journal of Regional Science*, 51(1), 167-193, (2011).

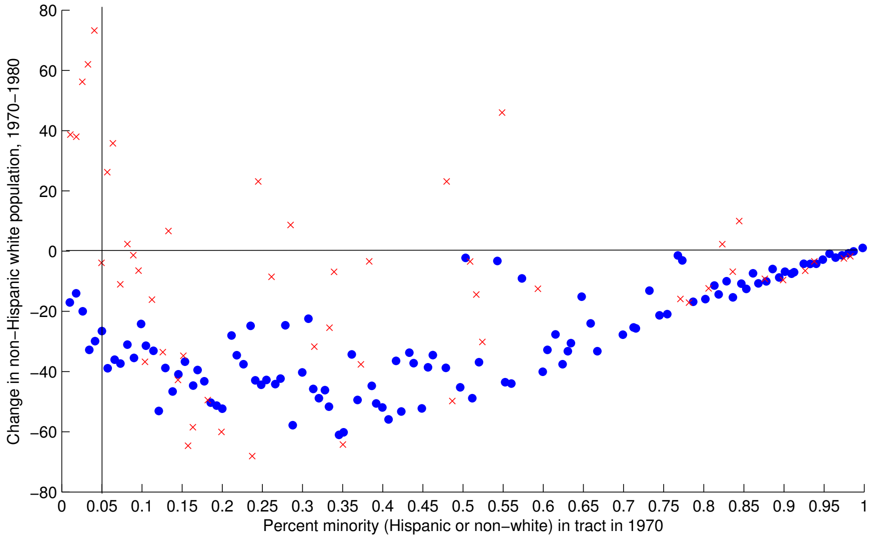
6 Figures

Figure 1: Neighborhood Change in Chicago, 1970-1980



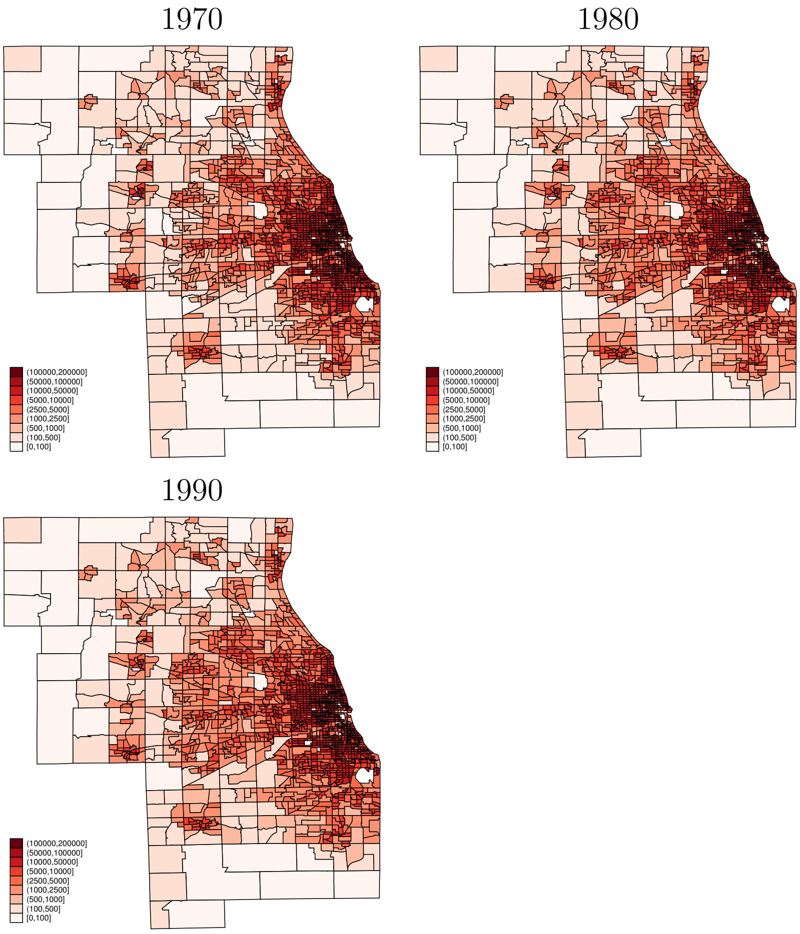
Notes: Dots show mean of the change in the tract-level non-Hispanic white population between 1970 and 1980 as a percentage of the total tract population in 1970, grouping tracts into cells of width 1% by the 1970 minority (Hispanic and/or nonwhite) share. The vertical line show the estimation of the tipping point using the threshold method mentioned in section 3.1.

Figure 2: Neighborhood Change in Chicago, 1970-1980



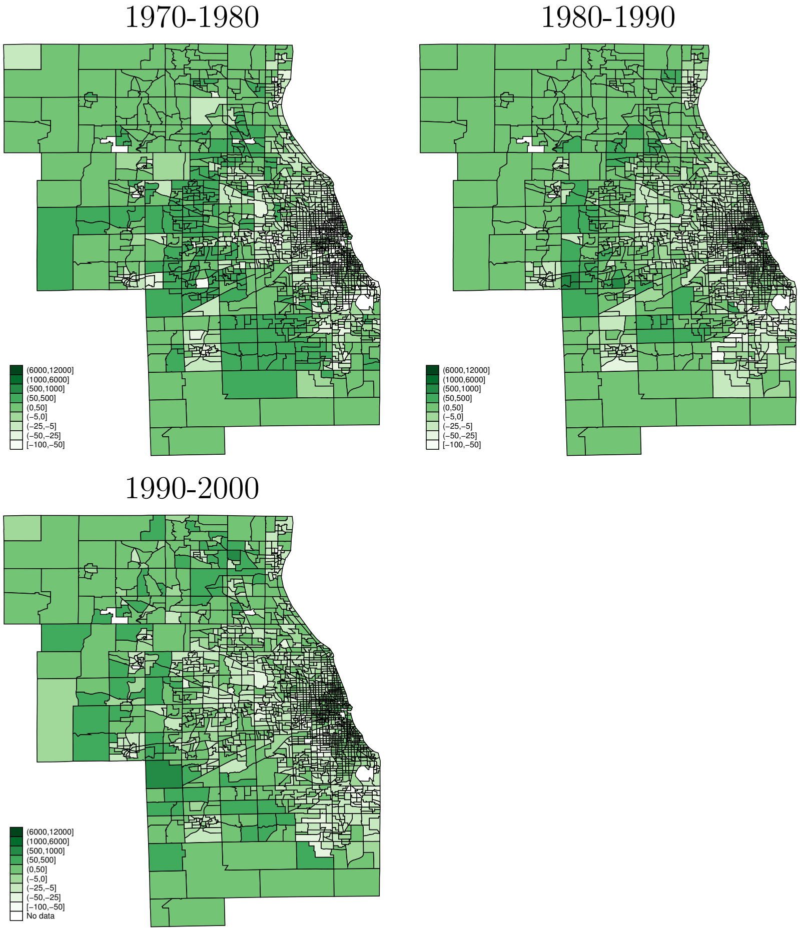
Notes: Data points show mean of the change in the tract-level non-Hispanic white population between 1970 and 1980 as a percentage of the total tract population in 1970, grouping tracts into cells of width 1% by the 1970 minority (Hispanic and/or nonwhite) share. The sample provides a distinction of tracts using population density. Tracks with a population density below (above) 5000 per sq. mile are market with a red-x (blue dot).

Figure 3: Population Density (per sq. mile) for Chicago



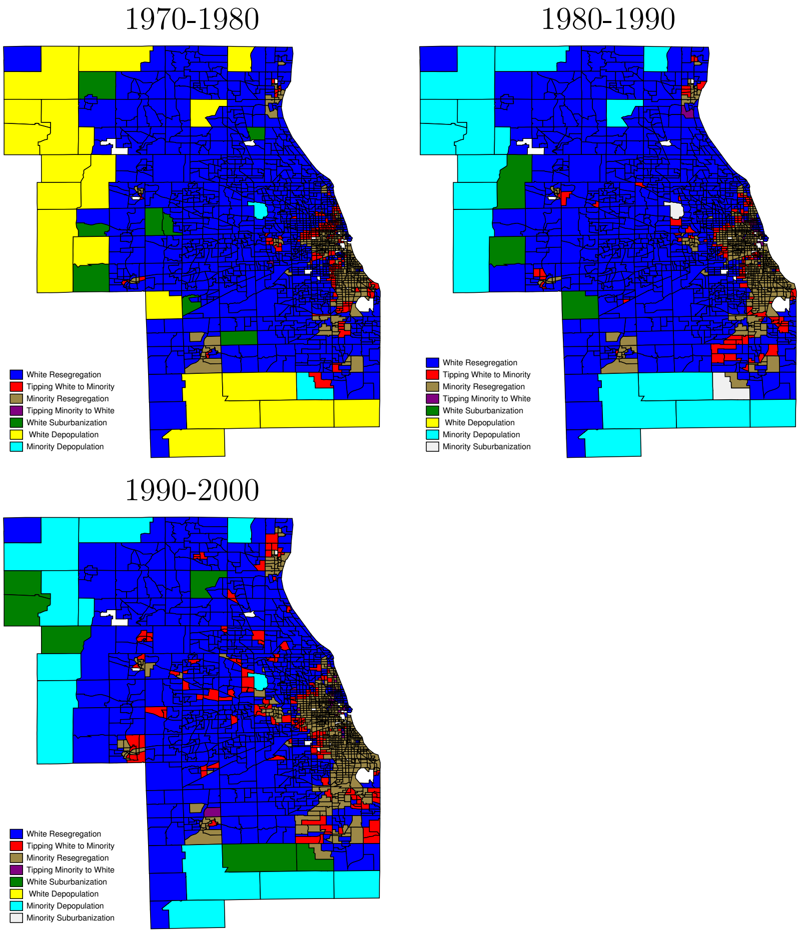
Note: Data for the figures comes from the NCDB database. Sample selection follows Card et al. (2008).

Figure 4: Decennial growth rate of the white population



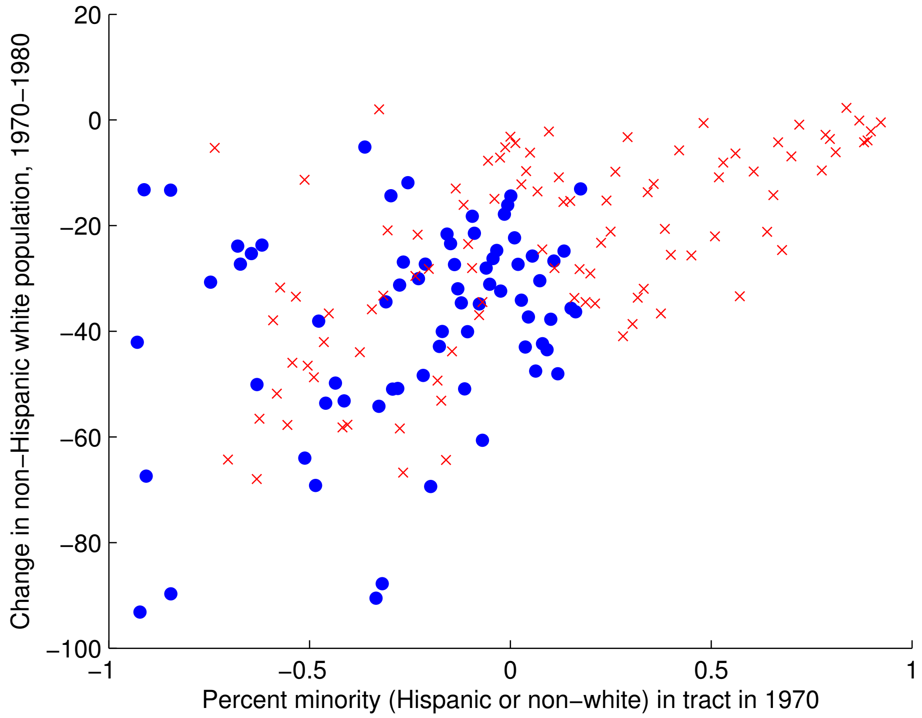
Note: Data for the figures comes from the NCDB database. Sample selection follows Card et al. (2008).

Figure 5: Classification Census Tracks for Chicago



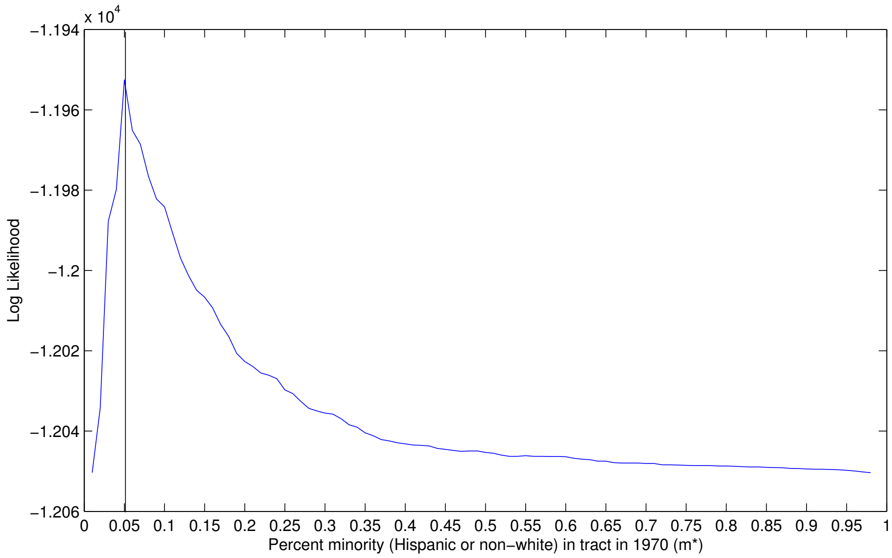
Note: Data for the figures comes from the NCDB database. Sample selection follows Card et al. (2008).

Figure 6: Neighborhood Change in Chicago, 1970-1980



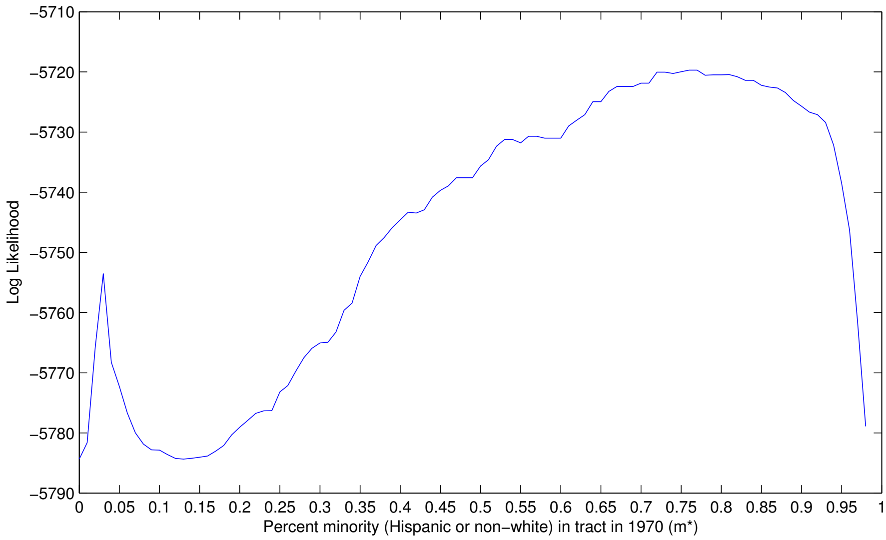
Notes: Data points show mean of the change in the tract-level non-Hispanic white population between 1970 and 1980 as a percentage of the total tract population in 1970, grouping tracts into cells of width 1% by the 1970 racial distance between neighboring track. The sample provides a distinction of tracts using the classification of highly white (minority) neighborhoods market with a blue dot (red-x). The sample selected exclude tracks with a population density below 5000 per sq. mile.

Figure 7: Concentrated maximum likelihood for threshold model, Chicago 1970-1980



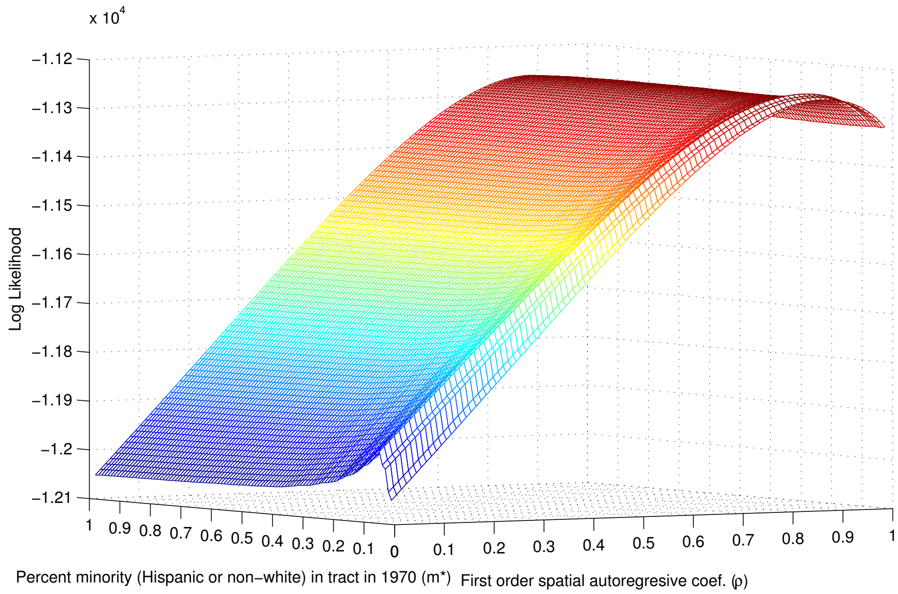
Notes: The Log-Likelihood is maximized at the tipping point identifies by Card et al. (2008).

Figure 8: Concentrated maximum likelihood for threshold model with limited sample, Chicago 1970-1980



Notes: The Log-Likelihood is maximized quite far from the tipping point identifies by Card et al. (2008) once all low population density (bellow 5000 per sq. mile) tract are removed from the sample.

Figure 9: Concentrated maximum likelihood for spatial-threshold model, Chicago 1970-1980



Notes: The Log-Likelihood indicates that the estimation of the tipping point is not reliable (flat along that dimension). On the other hand it seems relatively clear the value of the first order spatial autoregressive coefficient ρ that maximizes the log likelihood.

7 Tables

Table 1: Unconditional transition probabilities

	HW_{t+10}	HM_{t+10}	V_{t+10}
1970-2000			
HW_t	0.639309	0.337653	0.00072
HM_t	0.05123	0.831967	0.010246
V_t	0.542169	0.156627	0.301205
1970-1980			
HW_t	0.815695	0.159107	0.00288
HM_t	0.014344	0.872951	0.006148
V_t	0.192771	0.518072	0.289157
1980-1990			
HW_t	0.861017	0.118644	0
HM_t	0.047346	0.878049	0.004304
V_t	0.342466	0.232877	0.424658
1990-2000			
HW_t	0.822624	0.148416	0.000905
HM_t	0.053109	0.919689	0.002591
V_t	0.16	0.34	0.5

Note: HW_t: highly white tract at time t. HM_t: highly white tract at time t. V_t: vacant tract at time $t.HW_{t+10/30}$: highly white tract after 10 or 30 years. $HM_{t+10/30}$: highly white tract tract after 10 or 30 years. $V_{t+10/30}$: vacant tract after 10 or 30 years.

Table 2: Conditional transition probabilities.

	HW_{t+1}	HM_{t+1}	V_{t+1}
1970-2000			
$HW_t(OWN)$	0.74605	0.251693	0.002257
$HW_t(1MN)$	0.454082	0.545918	0
$HW_t(PMN)$	0.336585	0.658537	0.004878
$HM_t(OMN)$	0	0.995261	0.004739
$HM_t(PWN)$	0.122905	0.871508	0.005587
$HW_t(VAN)$	0.837838	0.07027	0.091892
1970-1980			
$HW_t(OWN)$	0.920993	0.075621	0.003386
$HW_t(1MN)$	0.770408	0.229592	0
$HW_t(PMN)$	0.473171	0.526829	0
$HM_t(OMN)$	0	1	0
$HM_t(PWN)$	0.03352	0.96648	0
$HW_t(VAN)$	0.821622	0.010811	0.167568
1980-1990			
$HW_t(OWN)$	0.954368	0.043025	0.002608
$HW_t(1MN)$	0.801282	0.198718	0
$HW_t(PMN)$	0.603015	0.396985	0
$HM_t(OMN)$	0.005141	0.989717	0.005141
$HM_t(PWN)$	0.13587	0.853261	0.01087
$HW_t(VAN)$	0.819549	0.007519	0.172932
1990-2000			
$HW_t(OWN)$	0.928465	0.070045	0.00149
$HW_t(1MN)$	0.802632	0.197368	0
$HW_t(PMN)$	0.618605	0.381395	0
$HM_t(OMN)$	0.021882	0.97593	0.002188
$HM_t(PWN)$	0.106061	0.888889	0.005051
$HW_t(VAN)$	0.759615	0.076923	0.163462

Note: $HW_t(OWN)$: highly white tract at time t with only highly white tracts sharing a border. $HW_t(1MN)$: highly white tract at time t with 1 highly minority neighbor. $HW_t(PMN)$: highly white tract at time t with with more than one highly minority neighbor. $HM_t(OWN)$: highly minority tract at time t with only highly minority tracts sharing a border. $HM_t(PMN)$: highly minority tract at time t with with more than one highly white neighbor.$HW_t(VAC)$: highly white tract at time t sharing a border with a vacant track. V_t: vacant tract at time t.$HW_{t+10/30}$: highly white tract after 10 or 30 years. $HM_{t+10/30}$: highly white tract tract after 10 or 30 years. $V_{t+10/30}$: vacant tract after 10 or 30 years.

Table 3: Threshold model

Sample	All	¿5000	All
Variable	1970-1980	1970-1980	1970-2000
α_0	19.33	-3.66	30.45
	(18.63)	(6.72)	(84.33)
α_m	-41.30***	22.68	-92.94***
	(2.80)	(1.98)	(8.58)
$\gamma_{earnings}$	13.34	-14.21***	59.18***
	(8.76)	(0.49)	(6.97)
γ_{value}	-7.44***	6.91***	-33.22***
	(1.34)	(0.55)	(4.1)
γ_{racial}	38.23***	13.72***	76.39***
	(7.18)	(3.28)	(20.76)
m*	0.04***	0.76***	0.04***
	(0.001)	(0.015)	(0.001)
R^2	0.15	0.14	0.05
R^2_x	0.17	0.26	0.11

The columns of the table contain the results for: the full sample, that is all tracks or only tracks with population density above 5000 per sq. mile. We introduce control variables that are based on the average characteristics of the neighborhood: household earnings ($\gamma_{earnings}$) and home value (γ_{value}). In addition a control for the racial distance (γ_{racial}) of the neighborhood and the surrounding neighborhoods. The estimated tipping points are not affected by the introduction of the control variables, R^2 are provided for the model without control variables and with control variables $R^2_{\mathbf{x}}$.Standard errors are shown in parentheses. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

Table 4: Spatial autoregressive model with threshold effects.

Sample	All	All
Variable	1970-1980	1970-1980
α_0	3.92 (18.06)	
α_m	-12.93*** (2.05)	
$\gamma_{earnings}$	10.21*** (1.32)	6.07*** (1.28)
γ_{value}	-5.87*** (0.88)	-3.48*** (0.74)
γ_{racial}	21.07*** (4.17)	16.07*** (4.12)
m*	0.95 (0.004)	
ρ	0.85*** (0.01)	0.89*** (0.01)
R^2	0.72	0.71

The columns of the table contain the results for: the full sample, that is all tracks or only tracks with population density above 5000 per sq. mile. We introduce control variables that are based on the average characteristics of the neighborhood: household earnings ($\gamma_{earnings}$) and home value (γ_{value}). In addition a control for the racial distance (γ_{racial}) of the neighborhood and the surrounding neighborhoods. Standard errors are shown in parentheses. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.