SimPolSeg: An Agent-Based Simulation of Political Migration Dynamics and Geographic Polarization*

November 4, 2014

Abstract

In this paper I simulate neighborhood level political migration dynamics following a change in neighborhood racial composition using SimPolSeg, an original agent-based modeling software program. SimPolSeg simulates agent behavior according to the Migration-Polarization (MP) theory of partisan sorting (Anastasopoulos 2014a). Dynamic simulations using SimPolSeg demonstrate how non-white migration and conservative flight lead to racially and ideologically segregated urban neighborhoods.

*This is a draft, please do not cite without permission.
1 Introduction

This paper builds upon the Migration-Polarization (MP) theory of partisan sorting described in Anastasopoulos (2014a). Using SimPolSeg, an original agent-based modeling algorithm, I explore patterns of racial segregation and geographic polarization according to MP Theory processes triggered by non-white migration to urban neighborhoods. Below I describe the MP Theory in more detail and present simulation results produced by SimPolSeg.

1.1 Theory

Since the 1950s, scholarship in several disciplines has found that the introduction of ethnically and racially diverse migrants into urban communities typically results in displacement of whites and segregation along racial and ethnic lines (Boustan 2010; Card, Mas and Rothstein 2008; Cutler, Glaeser and Vigdor 1999; Duncan and Duncan 1957; Fligstein 1981; Jackson 1985; Tolnay 2003). Schelling (1971) provided the first formal explanation for this phenomenon by showing that even if individuals have a weak preference for neighbors that are similar to them along some dimension, segregation can emerge when dynamic choices about where to live are based upon these preferences.

According to the Schelling model, tolerance for neighborhood racial and ethnic diversity determines whether an individual will relocate when changes in demographic composition occur. At higher levels of geography, tolerance determines the rate and extent to which segregation will occur. While tolerance in the Schelling model is pre-
presented as a preference independent of other individual attributes, in reality, factors such as age, marital status, family size, income and political ideology are correlated with it.

If this is true, then sorting induced by increases in diversity should lead to both racial AND ideological segregation (polarization) under certain conditions. This connection between political ideology and Schelling tolerance is the basis of the MP theory which is described in greater detail in Anastasopoulos (2014a).

Figure 1 provides an outline of the MP theory. Since individual political ideology is correlated with Schelling tolerance preferences, when a demographic shift causes an increase in neighborhood diversity, a partisan sorting process is triggered. Less tolerant, ideologically conservative individuals relocate in response to increases in neighborhood diversity, while more tolerant, ideologically liberal individuals remain.

This sorting process is an engine which drives geographic polarization as ideologically conservative “movers” and liberal “stayers” cluster together. Generational replacement and other changes that follow this initial increase in diversity ensure that this pattern of polarization persists in a path-dependent manner even in the absence of subsequent migration events.
2 Agent Based Model Setup and Simulations

2.1 Preferences for Diversity

The model below establishes agent behavioral rules which determine responses to changes in neighborhood diversity in subsequent simulations. In the simulation, neighborhoods \( n \), are geographically bounded spaces within a larger urban space \( A \). Agents that reside in these neighborhoods have the following preferences: (1) they are tolerant of diversity, but prefer neighbors that are racially or ethnically similar and; (2) their degree of tolerance is determined primarily by their political ideology.

\[
\eta_i = \sqrt{I_i - D_i} \quad (1)
\]

\[
D_i = \sqrt{\sum_{k=1}^{K} \sum_{j=1}^{J} \rho_{ji}(E[\alpha_{jk}] - \alpha_{ji})^2} \quad (2)
\]

\[0 < D_i < I_i < 1 \quad (3)\]

Decisions to relocate based upon these premises are operationalized using a simple model shown in the equations above where \( \eta_i \) represents the tolerance of an individual \( i \) for the proportion of minorities in their surrounding area. According to premise (2), tolerance \( \eta_i \) is a function of political ideology, \( I_i \). However, because tolerance also depends upon characteristics of the migrants, I include a “social distance” term, \( D_i \), which reflects differences on a number of attributes between the agent and the minority group(s) surrounding her (Shayo 2008). Higher values of ideology \( I_i \)
correspond to greater ideological liberalism while higher values of $D_i$ correspond to
greater “social distance” between the individual and an out-group. Since political ideology plays a central role in the determination of tolerance, $D_i < I_i$.

Social distance is a measure of characteristics which differ between an individual and an out-group that takes into account the importance an individual gives to each. Mathematically, it is the normalized difference between $E[\alpha_{jk}]$, the average $j^{th}$ attribute of group $k$, and $\alpha_{ji}$, the $j^{th}$ attribute of the individual. The distance on each attribute $\alpha_j$ is weighed by $\rho_{ji} \in (0, 1)$ where $\sum_j \rho_{ji} = 1$, which measures the importance that an individual gives to each attribute (Shayo 2008). For the purposes of this model, these attributes and weights are fixed for each individual\footnote{While it is certainly possible that the weight $\rho_{ji}$ an individual gives to these attributes may realistically change along with demographic changes, these considerations are to be addressed in future research. An example of this type of phenomenon could be an increase in the importance of language differences as individuals have more contact with immigrants that do not speak English.}

The utility that an agent receives for the residing in neighborhood $n$ is a function of tolerance, neighborhood minority population, housing prices, schools and other neighborhood amenities such as commuting distance etc. captured by $f(p_n, \epsilon_n)$:

$$U_i^R(m, \eta_i) = \eta_i m - m^2 = (\sqrt{I_i - D_i})m - m^2 + f(p_n, \epsilon_n)$$

$$0 \leq m \leq 1$$

$$0 < \eta_i < 1$$

$$0 < D_i < I_i < 1$$

Plots of $U_i^R$ v. $m$ with different values of political ideology and social distance in Figure 2 show that the utility an agent derives from neighborhood diversity varies
Figure 2: Utility for neighborhood diversity, $U^D_i$ v. area minority population varying Ideology $I_i$ and Social Distance $D_i$

with agent political ideology and out-group social distance. The shape of $U^R_i$ reflects the idea that agents are tolerant and have ideal levels of diversity represented by $m^*_i$ as shown in Figure 2. The ideal neighborhood minority proportion $m^*_i$ is obtained by simply taking $\partial U^R_i/\partial m$:

$$m^*_i = \eta_i / 2$$

Since $0 < \eta_i < 1$, the ideal proportion minority is always less than 50%, suggesting that agents are tolerant and enjoy diversity, but generally prefer neighbors similar to themselves.
2.2 Moving Decisions

At time $t$, agents make decisions about whether to move from a neighborhood based upon their preferences for diversity. Since this paper is concerned with how individuals respond to changes in diversity, I introduce agent beliefs about whether their neighborhood will “tip” (T) and become entirely majority-minority in the future ($m_{nt+1} = 1$) based upon the current area minority population $m_{nt}$ and the agent’s diversity ideal point $m_i^*$. Racial tipping is a phenomenon that has been extensively documented (Card, Mas and Rothstein 2008; Easterly 2005; Patrick, Fang, and McMillan 2005; Boston, Rigsby, and Zald 1972). Using census tract data from a number of major metropolitan areas, Card, Mas and Rothstein (2008) find that racial tipping points vary between 5-20% of the minority population and tend to be higher where whites are more liberal. Thus, a model which takes fears about tipping into account provides a realistic depiction of how agents might process changes in neighborhood racial composition.

$$P[T|m_{nt}, m_i^*] = P[m_{nt+1} = 1|m_{nt}, m_i^*] = \begin{cases} \sqrt{m_{nt} - m_i^*} & \text{if } m_{nt} > m_i^* \\ 0 & \text{if } m_{nt} \leq m_i^*, 0 < m_{nt}, m_i^* < 1 \end{cases}$$

Beliefs captured by $P[T|m_{nt}, m_i^*]$ reflect agent fears that others similar to themselves will exit the area if diversity exceeds their ideal point, $m_i^*$. When the minority population exceeds the agent’s diversity ideal point, her subjective belief about the likelihood that her neighborhood tips depends upon the actual neighborhood population, $m_{nt}$, and her minority ideal point $m_i^*$. The square-root term reflects the
behavioral assumption that increases in the minority population near the agent’s minority ideal point results in steeper marginal increases in her subjective probability of tipping.

Figure 3 is a plot of subjective tipping points at different levels of neighborhood minority population when $m^*_i = 0.10$.

Incorporating these beliefs into a dynamic model, expected utility for residing in a neighborhood at any given time $t$ is:

$$EU_{nt}^D = P[T|m_{nt}, m^*_i]U^R(T) + (1 - P[T|m_{nt}, m^*_i])U^R(T')$$ (4)

We can now establish a decision rule for moving in response to changes in neighborhood diversity using Equation ??.
about neighborhood tipping, an agent will choose to move from her neighborhood if her utility for residing there \( EU_{at}^R \) can be better satisfied elsewhere.

Thus, assuming that area fixed characteristics are the same, a rational agent will improve her utility by deciding to move in response to changes in diversity when (1) the minority population in her neighborhood exceeds her ideal point \( m_{nt} > m^* \) and; (2) there is at least one other area for which \( m_{n'} \leq m^* \).

Agents also take into account a cost of moving which is represented by \( \delta_{a,a'} \), the Euclidean distance between two neighborhoods as measured by their coordinates on a two dimensional plane:

\[
\delta_{n,n'} = \sqrt{(x_n - x_{n'})^2 + (y_n - y_{n'})^2} \tag{5}
\]

If \( m_{nt} > m^*_i \) and \( \exists n' \in A \) s.t. \( m_{n't} \leq m^*_i \): \[
\begin{align*}
\text{If } EU_{nt}^R < EU_{n't}^R - \delta_{n,n'} & \quad \text{Move from } n \\
\text{Else} & \quad \text{Remain in } n
\end{align*}
\] (6)

Finally, Equation (5) presents the conditions under which an agent will relocate. If these conditions are met, the agent may still be left with several neighborhoods to choose from. Since the agent desires to have the highest utility for remaining in an area, of the given areas that satisfy Equation (5), she will choose the neighborhood \( n' \in A \) which maximizes \( EU_{n'}^R - \delta_{n,n'} \). This implies that she will to move to a neighborhood both nearest her, in terms of spatial distance and closest to her neighborhood diversity ideal point.
3 Simulation

To explore how distributions of urban political ideology would develop if agents behaved according to the model discussed in the previous section, I designed a dynamic agent-based simulation algorithm in R called SimPolSeg. The algorithm generates a set of neighborhoods with two-dimensional spatial coordinates, populates them with minority and non-minority agents that have ideologies and preferences for diversity according to the model in the previous section and then simulates moving behavior over time. Simulation details can be found in the Appendix.

3.1 Initial Values and Demographics

<table>
<thead>
<tr>
<th>Variable</th>
<th>Symbol</th>
<th>Initial Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Neighborhoods</td>
<td>$A$</td>
<td>$N = 20$</td>
</tr>
</tbody>
</table>
| Mean Between Neighborhood Ideology | $I_{t_0}$ | $\mu_0^I = 0.5$  
|                         |        | $\sigma_0^I = 0.1$  
|                         |        | $I_{t_0} = \text{rtnorm}(n=20, \mu=0.5, \sigma=0.1)$ |
| Minoritypop             | $B_{t_0}$ | $\mu_0^B = 5$  
|                         |        | $\sigma_0^B = 5$  
|                         |        | $B_{t_0} = \text{rtnorm}(n=20, \mu=5, \sigma=5)$ |
| Majoritypop             | $W_{t_0}$ | $\mu_0^W = 100$  
|                         |        | $\sigma_0^W = 10$  
|                         |        | $W_{t_0} = \text{rtnorm}(n=20, \mu=100, \sigma=10)$ |

Table 1: Starting Values For Simulation

In the simulation conducted, 20 neighborhoods containing an average of 100 majority agents and 5 minority agents were generated. The mean majority “ideology”
in each neighborhood was set at 0.5 with a standard deviation of 0.1 and the mean minority ideology in each neighborhood was set at 0.8 with a standard deviation of 0.1. Moving behavior depends upon these initial parameters.

Figure 4 contains plots of the 20 areas in two-dimensional space initially and after 151 moving cycles, at which point the proportion of majority agent movers dropped below 1%. Area total population is reflected by point size.

3.2 Simulation Results

Figure 5 plots the percent of the population that moves after each time period. Moving increases dramatically during the first few time periods and then begins a steady decline toward zero after approximately 30 moving cycles.

Figure 6 is a plot of average area ideology over time for each of 20 areas. It is clear from this plot that urban area ideology diverges substantially over time as very liberal and moderate areas cluster together. This phenomenon becomes even more striking when comparing the two-dimensional maps of initial ideological and population distributions in Figure 4 (a) and (d). At the beginning of the simulation, nearly all areas are diverse and ideologically moderate. By the end of the simulation, only a few large population neighborhoods remain diverse and ideologically moderate while surrounding areas are comprised almost entirely of minorities and are ideologically extreme.

Urban ideological polarization as the result of partisan sorting becomes clear when neighborhood segregation and ideological polarization are plotted over time using the interquartile range (75th - 25th%ile) of average neighborhood ideology and
percent minority in Figure 7.
Figure 4: Neighborhoods in 2D Space: Avg. Ideology and Pct. Minority: $t = 0$ and $t = 151$
Figure 5: Movers as Percent of Total Population, $t = 0$ to $t = 151$
Figure 6: Average neighborhood ideology, $t = 0$ to $t = 151$
Figure 7: Polarization and Segregation Between Neighborhood, $t = 0$ to $t = 151$
4 References


Shayo, Moses. 2009. A model of social identity with an application to political econ-


5 Appendix

<table>
<thead>
<tr>
<th>Variable</th>
<th>Symbol</th>
<th>Description</th>
<th>Properties</th>
</tr>
</thead>
<tbody>
<tr>
<td>Areas</td>
<td>$A$</td>
<td>set of simulated neighborhoods.</td>
<td>$A = (a_1, ..., a_N)$</td>
</tr>
<tr>
<td>Agent Ideology</td>
<td>$I_{ai,t_0}$</td>
<td>Ideology of an agent.</td>
<td>$I_{ai,t_0} \in (0, 1)$</td>
</tr>
<tr>
<td>Area Ideology</td>
<td>$I_{a,t_0}$</td>
<td>Mean neighborhood ideology.</td>
<td>$I_{a,t_0} \in (0, 1)$</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>$I_{a,t_0} \sim N(\mu_{I_{a,t_0}}, \sigma_{I_{a,t_0}})$</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>$\hat{\mu}<em>{I</em>{a,t_0}} = \frac{1}{N_a} \sum_{i=1}^{N_a} I_{i,a,t_0}$</td>
</tr>
<tr>
<td>Minority pop</td>
<td>$B_{a,t_0}$</td>
<td>Neighborhood minority population</td>
<td>$B_{a,t_0} \geq 0$</td>
</tr>
<tr>
<td>SdMinority pop</td>
<td>$\sigma^B_{t_0}$</td>
<td>Between-neighborhood sd of the minority group population.</td>
<td>$\sigma^B_{t_0} &gt; 0$</td>
</tr>
<tr>
<td>Majority pop</td>
<td>$W_{a,t_0}$</td>
<td>Neighborhood majority population</td>
<td>$W_{a,t_0} &gt; B_{a,t_0} \forall a \in A$</td>
</tr>
<tr>
<td>SdMajority pop</td>
<td>$\sigma^W_{t_0}$</td>
<td>Between-neighborhood sd of the majority group population.</td>
<td>$\sigma^W_{t_0} &gt; 0$</td>
</tr>
</tbody>
</table>

Table 2: Initial Parameter Values of PolSegSim Algorithm

Since the simulation is designed to reflect changes in real populations over time, two population
dynamics are added to the model discussed above: (1) population growth - after each time period,
there is a 1% increase in the majority group agent population and a $1/t^2$ “migration shock” increase
in area minority population; (2) generational replacement - ideological preferences among new
agents reflect those of the area after one round of moving.

The PolSegSim algorithm allows users to input five variables which determine initial charac-
teristics of the PolSegSim universe. Table [2] describes these inputs which include: 1) the number
of areas populated by agents; 2) average majority population of the areas; 3) majority population
standard deviation between areas; 4) average minority population of the areas and finally; 5) mi-
nority standard deviation between areas. In the simulation, minority and majority populations are
homogeneous within groups. The majority group, for example, could be thought of as being all
white and the minority group all black.
Neighborhoods are first assigned two-dimensional spatial coordinates from a random uniform
distribution and are then populated with a number of majority and minority group agents generated
from a random normal distribution with means and standard deviations according to inputs 2, 3, 4
and 5 above. Initial ideology of each neighborhood $I_{a,t_0}$ is a draw from a truncated random normal
distribution with $N(\mu_{i,t_0} = 0.5, \sigma_{i,t_0} = 0.1)$. Minority and majority agent ideology within each
area $I_{at_0}$ are, in turn, draws from another truncated random normal distribution $N(\mu_{a,t_0} = I_{a,t_0}, \sigma_{a,t_0} = 0.1)$ with a mean equal to the randomly assigned area ideology. Average minority
agent ideology within each area is assumed to be three standard deviations higher (+0.3) than
average majority agent ideology. “Social distance” between groups are fixed for each agent at
$D_i = 0.1$.

During each moving cycle, agents simultaneously relocate to areas when the expected utility of
residing in their current area is less than their expected utility of moving: $EU_{R}^a < EU_{R}^a - \delta_{a,a'}$. This generally occurs when the area minority population at any given time exceeds their diversity
ideal point $m_{at} > m^*_i$. Once they have decided to move, they choose to move to an area which
maximizes $EU_{R}^a - \delta_{a,a'}$. This will be a candidate area which is both spatially closest to them as
calculated by Euclidean coordinate distance and has a minority proportion closest to their diversity
ideal point.

As mentioned above, to simulate real population dynamics, after each moving cycle a number of
agents are added equivalent to $1\%$ of the majority population and $1/t^2$ of the minority population.
The $1/t^2$ increase in the minority population reflects migration shocks in which areas that originally
had higher minority populations receive the greatest initial share of minority migration that de-
creases over time. New minority and majority agent ideologies are draws from a truncated random
normal distribution whose mean and standard deviation are the mean and standard deviation of
area ideology after a cycle of moving but before the new agents are added.

Thus, for example, at time $t$ agents decide to move. After they move, new average area ideology

\[^2\text{the distribution is truncated because } 0 < I < 1\]
\[^3\text{The only situation in which agents would not move when } m_{at} > m^*_i \text{ is if they happen to reside in a very isolated area where the Euclidean distance from their current location is greater than the utility gain they expect to gain in any new area, } \delta_{a,a'} > EU_{a'}^R - EU_{a}^R\]
is computed $I_{a,t+\frac{1}{2}}$. New majority and minority agents are added to an area whose ideologies are a draw from a random normal distribution with a mean equivalent to $I_{a,t+\frac{1}{2}} = \frac{1}{N_{a,t+\frac{1}{2}}} \sum_{i=1}^{N_{a,t+\frac{1}{2}}} I_{a,i,t+\frac{1}{2}}$ and standard deviation equivalent to $\sigma_{a,t+\frac{1}{2}}$. Thus, new agents entering the area have ideologies in line with current area ideology after moving.

After each moving cycle, the simulation continues to run until the percent of the population that moves is below 1%. In the simulation conducted below, this occurred after 151 moving cycles.