Chapter 11

Immigration and Integration: Online Appendix

The Schelling Segregation model is a classic example of social simulation. After his publication, other scholars have used agent based modelling to understand yet other social phenomena than residential segregation (Granovetter, 1978; Railsback & Grimm, 2019; Squazzoni, 2012). This social simulation literature has identified the relevance of sociological concepts such as *thresholds* and *tipping points*. This Appendix is a brief introduction to the art of social simulation, using the Schelling Segregation model as an example, and reviewing more recent work in this field.

If you want to develop an Agent Based Model, you can consider going through the following seven steps. Table 11A summarizes these steps, and reveals –in a stylized way- how they are incorporated in the Schelling model.

Table 11A How to build an agent based model, with application to the Schelling model

Step	*To Do*	*Example: Schelling Model*
1	Define Agents	• Blacks and Whites • Equal in size
2	Define Social System	• Neighborhoods
3	Define Micro Rules	• Agents evaluate racial composition of their own neighborhood. • They move to another neighborhood if they are not satisfied with the racial composition • Threshold rule
4	Design the Simulation	• Start with random distribution. • Sequential steps. • Checkerboard
5	Evaluate collective outcomes	• Hypersegregation

Step	To Do	Example: Schelling Model
6	Evaluate individual-collective dynamics	• Simple aggregation does not work • Self-reinforcing processes • Predictably unpredictable • Tipping points
7	Check assumptions and start again	• Threshold rules • Group size

Step 1. Define Agents

The first step in ABM is to define *who* is in the model. In the language of ABM, the people who are in are called the *agents*. These agents are hypothetical (and not real) individuals, and they belong to a certain group, which need to be specified. To develop an ABM, as in any theory (Chapter 2) it is important to keep things simple. The Schelling model assumes that agents are either Black or White. Additionally, the model assumes that the two groups are equal in size. These are highly simplified assumptions, as in reality there are many different ethnic groups, and they differ in size, too. As a start, however, one needs to begin somewhere, and ideally, this is as simple as possible.

Step 2. Define the Social System

In the next step, you need to define the social context: what is the environment, the social system? Again, you need to keep things simple. The Schelling model says that the social system consists of neighborhoods, in which agents live and move between. Thus, Black and White agents live in a certain neighborhood, there are multiple neighborhoods, and there is movement between neighborhoods.

Step 3. Define Micro Rules

Now that you have defined the social context, you need to specify behavioral assumptions or micro rules. Again, keep it simple and focus only on those elements which are important. The Schelling model assumes that agents are aware of the racial composition of neighborhoods; that agents decide to stay in or leave their neighborhood; that this decision depends only on the racial composition of their neighborhood. When the share of people of the same race as that of the agent in the neighborhood falls below a certain *threshold*, then people move. And they move to another neighborhood in which the proportion same-race people is above the threshold.

In the Schelling model, all agents have *identical* thresholds with respect to neighborhood preferences. There are no differences between individuals in what they find a satisfactory racial composition of their neighborhood. Schelling assumed that *all* members of *both* groups have a *modest* preference to live in a neighborhood with people of their own race. In one version of the segregation model, Schelling assumed that people had a threshold of more than 33%. In other words, people want to live in a neighborhood with more than one third of the neighbors being of the same race. If the neighborhood has more than one third of same-race members, agents stay. If it is less than one third, the agent moves to a neighborhood where there are more than 33% of same-race members.

Step 4. Design the Simulation

In the next step, you design the simulation. To get things started, you need to make a decision about the initial conditions, i.e. the conditions in a certain social context at the beginning

-before the simulation starts. In this case: how are the agents distributed across the neighborhoods? In Schelling model, Black and White agents are *randomly* distributed. Thus, before the simulations start, there is no residential segregation beyond random deviations. Then, you start the simulation. In the Schelling model, this is done by letting agents move between neighborhoods sequentially. This means that at a certain point in time, only one agent can move to another neighborhood, and then another agent can move, and so forth. You can use a checkerboard to run these simulations, as we have seen.

Step 5. Evaluate Collective Outcomes
In step 5, you evaluate the collective outcomes: what is the end result of the simulation? We have discussed the outcomes of the Schelling results in Chapter 11.

Step 6. Evaluate Individual-Collective Dynamics
After the simulations are finished, and the collective outcomes have been interpreted, it is important to evaluate the insights from the simulations on the dynamics between individuals and their social context. What can we learn from these simulations about the emergence of social phenomena, like hypersegregation? I highlight four insights that deserve special attention.

a. Collective outcomes are not (always) simply aggregated micro-motives
The surprising outcome of the simple simulation game is that when we assume that the actions of individuals depend on what others have done, micro-level preferences cannot be simply aggregated to the collective level. In the racial segregation model, members of both groups are quite tolerant, but nevertheless a high level of segregation is the outcome in the end. Such surprising, collective outcomes that are not simply aggregations of individuals occur under conditions of social interdependency. They result in non-normal distributions and they are often counterintuitive.

When there is interdependency between the actions of people, we need to take into account this dynamic, i.e. the process of interaction. When the actions of one individual do not affect that of other individuals –when there is no interdependency– we can simply aggregate micro motives and behavior to predict collective outcomes. In the Schelling model, however, the system of interaction is dynamic as agents respond to each other (or more precisely: to changing neighborhood conditions), and therefore we need to understand how the system evolves over time (Squazzoni, 2012).

b. Self-reinforcing dynamics
When we have a situation in which there exists interdependency between the actions of individuals, as in the Schelling Segregation Model, we need to establish the kind of feedback that takes place. As we have discussed (Chapter 2), a feedback relation exists between two variables *X* and *Y* when the influence is going in both directions, i.e., *X* affects *Y* and *Y* affects *X*. This is the case here as well. Specifically, the propensity of certain groups (i.e., Whites versus Blacks) to migrate in- and out of a certain neighborhood (*X*) is related to the racial composition of that neighborhood (*Y*). In the Schelling model, the relation between the two is a *positive feedback* (Miller & Page, 2007). This means that *X* has a positive influence on *Y*, and *Y* has a positive influence on *X*.

To see this, consider the following example. A certain neighborhood contains more White than Black people, say 55% against 45%. This means that this place is attractive for White persons to settle, more so, than it is for Blacks. Hence, it is very likely that this neighborhood will see more White persons coming than leaving, whereas for Blacks the opposite will happen. Due to these selective in- and out-migrations, the composition of this neighborhood changes: it becomes even more White over time (e.g., 65%). And for that reason, it will become even more attractive for White persons to move into that area, whereas Blacks will move to places elsewhere. This is a self-reinforcing process that operates at the level of neighborhoods.

c. Predictably unpredictable

Another insight from the Schelling model, and from ABM more generally, is that the *precise* collective outcomes for each *specific neighborhood* are hard to predict. Although the model shows that in the end of the process there will be high levels of racial segregation across neighborhoods, it is not easy to foresee exactly *which* neighborhood will become predominantly White, and which one will not.

The reason for this is that, in the beginning of the simulations, when Blacks and Whites were randomly scattered across the board, the specific choices made by the initial moves matter a great deal. Remember that, in the beginning of the game, Blacks and Whites who were unsatisfied with the racial composition in their neighborhood, had the opportunity to choose to which neighborhood they would like to migrate, say N1, N2, and N3 all of them satisfying their preferences.

Suppose that in the beginning of a certain simulation, at *t1*, a White person W1 moves into neighborhood N1, then this very single movement could set into motion the self-reinforcing process discussed above. It could result in other White people also moving to that neighborhood N1, and in the end of the process that neighborhood could have become largely White. But what would have happened if, right in the beginning, that same White person W1 had decided to move to neighborhood N2 instead? In that case, other White people would not have migrated to neighborhood N2 rather than N1. It would have resulted in neighborhood N2 becoming largely White instead of neighborhood N1.

This brings us to another important insight of the segregation simulations, namely that although we can predict that in the end there will be high levels of racial segregation *overall* (i.e., across all neighborhoods), we cannot predict precisely *which* neighborhoods will become White and which ones will become Black. The self-reinforcing process amplifies differences that emerge right in the beginning, and this means that the collective outcomes are *very sensitive to the initial condition and to the behavior of the very first actors who migrate*: they effectively determine which neighborhoods become White, and which ones become Black. This may remind you of the conclusions from the Salganik-Watts study on the success of music songs (Chapter 5): if people could see how many others had downloaded a certain song, then they would follow the crowd and listen to the most popular songs. That's another example of a self-reinforcing process. Hence, the choices of the very first people largely determines what the rest will do.

It is interesting and fun to check this principle with the Schelling model. The only thing you have to do, is to run the racial segregation model again, but then start with another initial condition, or let the first actor move to another location than before. When playing the game again and again, with varying starting conditions or first moves, you will see that what become White and Black neighborhoods changes all the time.

d. Tipping points

Another insight from agent based modelling relates to so-called *tipping points*. These refer to situations in which a *small* change in the value of a certain variable has a *big* effect on the collective outcomes –either locally or more globally. The level at which such an abrupt change occurs is the "tipping point" (Lamberson & Page, 2012).

In the Schelling segregation model, tipping points are related to the ethnic/racial composition of the neighborhood. Remember that agents acted according to the threshold rule: if they have 33% of same-race neighbors or less than that, they move to a neighborhood that satisfies this 33% threshold rule. This means that when a certain neighborhood changes in racial composition and exceeds this threshold, it can result in a major change in neighborhood composition. To illustrate, let's say we have a certain neighborhood N1, which consists of 63% Black and 37% White. Both groups are satisfied with this composition. However, imagine that over time more and more Blacks enter N1, for whatever reason. The percentage of Whites in N1 decreases to 36%, 35%, and 34%. This will not affect the decision of Whites to move. But then, when there are 33% Whites, all of them will suddenly start to migrate to another neighborhood. Given the preference for having more than 33% of the own-race, a small change in neighborhood composition (from 34% to 33% Whites) has major consequences for the neighborhood.

Step 7: Check assumptions and start again

The Schelling model, developed in the late 1970s, has become a classic example of social simulation, suggesting that hypersegregation may emerge even when people have only mild in-group preferences. It has stimulated social scientists to use agent based models, not only to explain segregation, but to understand a wide variety of social phenomena in which social interactions play a key role (Granovetter, 1978; Squazzoni, 2012). The Schelling model of segregation has also been elaborated, refined and empirically tested in numerous studies. This is step 7 in our sequence, and it consists of two sorts of actions.

To begin, you can check whether some of the *key assumptions of the ABM are in line with reality*, and appear not too unrealistic. Remember that the Schelling model assumed that *all* agents, i.e., *all* Whites and *all* Blacks, are *equally* unsatisfied if their neighborhood has 33% or fewer same-race neighbors. But is this a realistic assumption? In one early study on ethnicity and residential preferences in Detroit in 1976 (Farley, Schuman, Bianchi, Colasanto, & Hatchett, 1978), it was found that Whites had much stronger residential in-group preferences than Blacks, thereby undermining the assumption that both groups have *equal* in-group preferences. Using flash cards showing different variants of neighborhood compositions of 15 White/Black households, it was found that Whites generally favored a nearly all white neighborhood, whereas Blacks tended to prefer a 50/50 Black-White neighborhood. No less than 31% of Blacks indicated that they would *not* be willing to move into an all-Black neighborhood. Only 5% of Blacks said that they would not move into a neighborhood in which they represent only 20% of the neighborhood, i.e. a majority-white neighborhood. Thus, the large-majority of Blacks had no problems with living in such predominantly White neighborhoods. Follow-up studies have confirmed these results, suggesting a robust empirical pattern of *group-asymmetry of residential in-group preferences*, i.e., ethnic majority members are less willing to live in areas with many ethnic minorities whereas ethnic minority members are willing to live in areas with many ethnic majority members (Clark, 1991, 1992).

Another insight from these studies is that such preferences tend to differ between persons within the same ethnic group. The Schelling model assumed that *all* Blacks and *all* Whites have a threshold of 33% same-race. But what does empirics show?

Let's focus on Whites. The 1976 Detroit study revealed that 7% of White respondents have *extremely strong in-group* preferences. Specifically, this group would migrate to another neighborhood in case they would live in a neighborhood with one or more Black households out of 15 households in total (Farley et al., 1978). A total of 24% of White respondents would move away when there would be 3 (or more) Black households out of 15 households in the neighborhood. In case there would be 5 Black households out of 15, then 41% of the

White respondents would migrate, and this goes up to 64% in case of Blacks make up 9 out of 15 households. This pattern has been found more generally, and it tells us that there is not a common threshold. Instead, people differ in their thresholds: some have strong in-group preferences and migrate to another neighborhood already if only one family from another races settles in their neighborhood. Others do not care so much. In summary, there is evidence suggesting *within-group heterogeneity in thresholds*.

The second kind of action you could take in step 7, next to checking the realism of some of the key assumptions, is to redo the simulations, but then using different assumptions. The core idea behind redoing the simulations, is that you could discover what would happen at the collective level, if you would use different assumptions. Would hypersegregation still emerge, if you would start with the more realistic assumption of group-asymmetry of residential ingroup preferences? What do you find if you would model within-group heterogeneity in thresholds? What would happen if you would redo the simulation not with two ethnic/racial groups but with three or four? What if we would change the size of the ethnic groups? The answers to these questions could fall into two categories:

a) *I get the same results!* If the results of your simulations –using different assumptions- are very similar to the original setup you can conclude that the Schelling model is *robust* to the changes you made to the model. For example, if you changed the number of ethnic/racial groups in the population from 2 to 4, and you still get Black hypersegregation, then you could conclude that apparently the number of groups is irrelevant for the phenomenon to emerge.
b) *I get different results!* If the simulations show very different results, for example that hypersegregation does not occur when there are many ethnic groups, then the phenomenon of hypersegregation is *sensitive* to how many groups there are in the population. This suggests that, somehow, you have found certain conditions under which Black hypersegregation emerge and when it does not. It may arise when there are only two groups, but not so, when there are more than two.

In contemporary research on residential segregation that follows in the footsteps of the Schelling model step 7 is still a matter of deep controversy and debate among scholars. Researchers are checking the realism of key assumptions, and they redo simulations using different (and often more realistic) assumptions (Benard & Willer, 2007; Clark & Fossett, 2008; Fossett & Waren, 2005; Fossett, 2006; Macy & Van de Rijt, 2006; Stoica & Flache, 2014; Zhang, 2004). A major improvement in this field, since the work of Sakoda and Schelling, is that scientists today no longer need to rely on checkerboards to run their simulations. Instead, the simulations are programmed with computer software (such as NetLogo), and computer power allows them to run simulations with millions of actors. Such computer-based agent based models are core in the study of what is nowadays called 'computational social science' and the study of 'complexity' (Macy & Willer, 2002; Mitchell, 2009).

References

Benard, S., & Willer, R. (2007). A Wealth and Status-Based Model of Residential Segregation. *Journal of Mathematical Sociology, 31*(2), 149–174.
Clark, W. A. V. (1991). Residential Preferences and Neighborhood Racial Segregation: A Test of the Schelling Segregation Model. *Demography, 28*(1), 1–19.
Clark, W. A. V. (1992). Residential Preferences and Residential Choices in a Multiethnic Context. *Demography, 29*(3), 451–466.

Clark, W. A. V., & Fossett, M. (2008). Understanding the Social Context of the Schelling Segregation Model. *Proceedings of the National Academy of Sciences of the United States of America, 105*(11), 4109–4114.

Farley, R., Schuman, H., Bianchi, S., Colasanto, D., & Hatchett, S. (1978). "Chocolate City, Vanilla Suburbs:" Will the Trend Toward Racially Separate Communities Continue? *Social Science Research, 7*(4), 319–344.

Fossett, M. (2006). Ethnic Preferences, Social Distance Dynamics, and Residential Segregation: Theoretical Explorations using Simulation Analysis. *Journal of Mathematical Sociology, 30*(3-4), 185–273.

Fossett, M., & Waren, W. (2005). Overlooked Implications of Ethnic Preferences for Residential Segregation in Agent-Based Models. *Urban Studies, 42*(11), 1893–1917.

Granovetter, M. (1978). Threshold Models for Collective Behavior. *American Journal of Sociology, 83*(6), 1420–1443.

Lamberson, P., & Page, S. E. (2012). Tipping Points. *Quarterly Journal of Political Science, 7*(2), 175–208.

Macy, M. W., & Willer, R. (2002). From Factors to Actors: Computational Sociology and Agent-Based Modeling. *Annual Review of Sociology, 28*(1), 143–166.

Macy, M. W., & Van de Rijt, A. (2006). Ethnic Preferences and Residential Segregation: Theoretical Explorations Beyond Detroit. *Journal of Mathematical Sociology, 30*(3–4), 275–288.

Miller, J. H., & Page, S. E. (2007). *Complex Adaptive Systems: An Introduction to Computational Models of Social Life*. Princeton, NJ: Princeton University Press.

Mitchell, M. (2009). *Complexity: A Guided Tour*. New York, NY: Oxford University Press.

Railsback, S. F., & Grimm, V. (2019). *Agent-Based and Individual-Based Modeling: A Practical Introduction*. Princeton, NJ: Princeton University Press.

Squazzoni, F. (2012). *Agent-Based Computational Sociology*. Chichester, United Kingdom: John Wiley & Sons.

Stoica, V. I., & Flache, A. (2014). From Schelling to Schools: A Comparison of a Model of Residential Segregation with a Model of School Segregation. *Journal of Artificial Societies and Social Simulation, 17*(1), 5.

Zhang, J. (2004). A Dynamic Model of Residential Segregation. *Journal of Mathematical Sociology, 28*(3), 147–170.