Socio-dynamic discrete choice: Theory and application

Dugundji, E.R.

Publication date
2013

Citation for published version (APA):
This chapter explores a multi-agent based model of binary choice behavior with interdependence of decision makers’ choices. Related early agent based modeling approaches to social influence on networks are Hedstrom [2004], Rolfe [2004], Rahmandad and Sterman [2004], Stauffer [2001], Flache and Hegselmann [2001], Deffuant, Ambland, Weisbuch and Faure [2002], Urbig [2003], Stauffer, Sousa and Schulze [2004], among others in a large stream of interdisciplinary research on opinion dynamics addressed by both social scientists and physical scientists alike. The research presented in this chapter notably differs from these works in its focus on issues that arise in statistical estimation for empirical application of such agent based models, and the interplay between the econometric estimation on one hand and the agent based model on the other. Studying the long-run behavior of more than 120,000 multi-agent based simulation runs reveals that the initial estimation process can be highly sensitive to small variations in network instantiations. We show that this is an artifact of two issues in estimation, and highlight particular attention that is due at low network density and at high network density. This finding is an important warning with respect to empirical application of agent based models.

In addition to this contribution, we present a number of sub-results which we believe are important to highlight as good practice with agent based modelling. One key feature of agent based modelling is internal verification, or otherwise said, how can the researcher be confident that the agent based model is performing the actions that it is expected to do? What is the evidence that the programming implementation of the abstract or conceptual model is correct? There are many subtleties of agent based modelling code that arise with scheduling and simulation of probabilistic events over time steps involving ordering of agents and multiple random seeds for different draws. Short of complete re-implementation of the conceptual model with a different modelling toolbox or with a different language by a different team, there are all too often few guarantees of internal verification with a complex model. To address this fact, we begin our modelling endeavor with a very simplified model studied previously by others, building up our agent based model step by step, and adding different layers of complexity one at a time. In our case, this means adding different kinds of heterogeneity to the agent based model.

In section 6.1, we begin by benchmarking the agent based model with established analytical results where agent heterogeneity is not
explicitly treated. Subsequently in the rest of the chapter, we incrementally add different layers of complexity, first relaxing the condition of global connectivity to allow for local connectivity (and thus heterogeneous information), and then bringing in additional heterogeneity in the agent characteristics (such as gender) and agent-specific attributes of trips (such as travel purpose, travel time, travel cost). In section 6.2, two abstract classes of networks are considered to introduce agent heterogeneity via an explicit local interaction structure. In sections 6.3 and 6.4, the model is applied in an example of intercity travel demand using empirical data to introduce individual agent heterogeneity beyond that induced by the local interaction structure. Discrete choice estimation results controlling overall mechanisms related to individual heterogeneous preferences are embedded in the agent based model to be able to observe the simulated evolution of choice behavior over time with sociodynamic feedback due to network effects.

6.1 Multi-Agent Based Social Simulation

Due to the focus on heterogeneous agents’ evaluation of alternatives, the discrete choice theory framework lends itself quite naturally to combination with social simulation of multi-agent systems. In a recent tutorial on agent-based modeling, Macal and North (2010, p. 151) write:

“Agent based modeling and simulation is (an) approach to modeling complex systems composed of interacting, autonomous ‘agents.’ Agents have behaviors, often described by simple rules, and interactions with other agents, which in turn influence their behaviors. By modeling agents individually, the full effects of the diversity that exists among agents in their attributes and behaviors can be observed as it gives rise to the behavior of the system as a whole. By modeling systems from the ‘ground up’ – agent-by-agent and interaction-by-interaction – self-organization can be observed in such models. Patterns, structures, and behaviors emerge that were not explicitly programmed into the models, but arise through the agent interactions. The emphasis on modeling the heterogeneity of agents across a population and the emergence of self-organization are two of the distinguishing features of agent based simulation as compared to other simulation techniques…”

In this dissertation, the agents are our sample population of decision makers. Their behavior is their choice between a number of discrete choice alternatives. The simple rule that the agents follow to update
their decisions is a stochastic rule given by individual choice probabilities $P_n(i|C_n)$. The emergent system behavior that we will be interested to track is the proportions of agents in the population that have made each choice. The types of patterns that we will be interested to observe is whether or not a run-away effect occurs with agents flocking to one of the choices, whether there are multiplicity of such steady states, what the stability of the steady states is, and if there is evidence of stochastically transitioning between stable steady states.

Once we have conceptually specified the agents, their behaviors and the interactions with other agents, a computational engine is necessary to “make the model run”, i.e. to simulate the repeated updating of choices made by each agent. Although it is possible to program the agents, their behaviors and the interactions with other agents from scratch with a programming language such as Java, Python, C or C++, there are dedicated agent based modeling “toolkits” which simplify this task and provide handy modules such as schedulers. In their excellent tutorial, Macal and North (1) describe the many useful services that agent based modeling toolkits can provide: project specification (via a library oriented approach, an integrated development environment, or both); agent specification (including specialized support features for adaptation and learning, optimization, social networks, geographical information systems, etc); input data specification and storage in a variety of formats; model execution (including visual display of single runs during runtime and interactive probing, as well as batch execution of multiple runs on one computer or divided over a cluster); results storage and analysis; and model packaging and distribution. For the purposes of this project we used the Repast modeling platform (http://repast.sourceforge.net) developed at University of Chicago Argonne National Laboratory by North, Macal and other collaborators.

The multi-agent based simulations for the runs in this dissertation are carried out in two phases: the initialization phase and the iteration (simulation) phase. We will describe each of these phases briefly.

### 6.1.1 Initialization

During the initialization phase, the survey data is read into memory and the agents and their interactions are specified. Concretely, this involves reading in two flat text files: one file contains a “node list” indicating agent-by-agent for all agents which other agents a given agent is connected to for a particular run as defined by agent characteristics in the survey data; the other file contains an ordinary table of other information specific to each agent. Both files are organized with one row per agent, but the node list file has in general varying row width, since the number of agents that each agent is connected to will in general vary across agents, except for the special case with
a fully connected network where everyone is connected to everyone. The table of other information per agent is rectangular. The reason for the separation of the input information into two files is primarily technical, and facilitates the convenience of modifying independently the initial conditions in terms of the initial mode choices and the hypothesized network structure of the agent interactions.

With regard to the implementation of the agent interactions, a brief note is worth mentioning at this point, regarding so-called “self loops.” This refers to counting an agent’s own current state (i.e. which alternative the agent chose most recently), in the aggregate effect influencing the agent’s evaluation of utilities for a possible change of choice. Later in this chapter, in the case study application in sections 6.3 and 6.4 we specifically test the effect of including the agent’s own choice or not in the aggregate influence for a sweep of network densities. For simplicity here, the weight from an agent’s own choice is counted in the same way together with all of the other reference agents for that agent. In a practical application for policy purposes however, it could be highly relevant to also estimate a separate coefficient for the influence from an agent’s own past choice(s), depending on the particular choice context, such as in Dugundji, Poorthuis and van Meeteren (2011). This is to be expected to yield additionally complexity in the envelope of possible steady state outcomes, depending on the magnitude of such influence.

6.1.2 Iteration

During the simulation phase, each agent one-by-one executes their interactions with their connections to find out what choices their reference agents have most recently made, and then updates their own choice according to the probabilistic logit rule. The action of one single agent executing interactions, computing the choice distribution seen by that particular agent, and then updating their own choice is one iteration in our case study. Whether or not a given agent actually changes their previously made choice during the updating stage of the iteration is determined by computing the individual choice probabilities for each alternative according to the choice distribution seen by the given agent of their reference agents, and dividing a line from zero to one into bins according to these choice probabilities. The agent then makes its choice by drawing a uniform random number from zero to one and seeing in which bin this number falls. Each agent is chosen in randomized order without replacement by the scheduler until all agents have executed their interactions and have had a chance to update their choice. The scheduler then starts all over again and again choosing agents without replacement, and iterates as many times as necessary until the fixed maximum number of iterations is reached. This entire process is called a single simulation run.
The randomized order sequence used by the scheduler for a given run is set by a pseudo-random number generator seed. We will then want to repeat a particular simulation run with different seeds some number of times in order to check if there are any systematic consequences of the decision making order. By setting the agent decision making order sequence with the seed however, we have the possibility to repeat an exact particular run if necessary. This can be useful for example, to zoom in on the details of the time steps, for probing agents during run-time, to make modifications in inputs, and/or to export different output for analysis. Concretely, the iteration/simulation phase involves accessing two flat text files: one file contains the necessary stochastic dynamic parameters such as the total number of runs to carry out, the seed for the agent decision making order for each run, and the total number of iterations per run; the other file contains the necessary configuration information for the logit rule such as the values of the estimated coefficients.

6.1.3 Benchmark Model

For the purpose of verifying the programming implementation of the agent based model in this section, we now calculate some analytical benchmark results for the long term system behavior of \( \phi \). For the case \( f(\phi) = \phi \), we have simply the derivative \( f'(\phi) = 1 \). The condition for local stability at an equilibrium state reduces to:

\[
\left. \frac{d}{d\phi} \left( \frac{d\phi}{dt} \right) \right|_{\phi=\tanh(\frac{1}{2} \beta f(\phi))} \leq 0 : \ (1 - \phi^2) \cdot \beta \leq 2 \tag{6.1}
\]

For the case \( \beta = 0.03 \), we have effectively a repeated coin toss. In the long term, \( f(\phi) \approx 0 \) and for \( \beta \) close to zero, we have also \( \beta f(x) \approx 0 \). This corresponds to a situation where the curve \( y = \tanh(\frac{1}{2}) \beta f(\phi) \) crosses the line \( y = \phi \) at \( \phi = 0 \) "from above" at the left panel in Figure 6.1. The solution at \( \phi = 0 \) is stable; moving a little bit away from the solution at \( \phi = 0 \) along the tanh curve returns you back to the solution at \( \phi = 0 \). Indeed we also see the inequality in equation (6.1) is satisfied:

\[
(1 - \phi^2) \cdot \beta = (1 - 0^2) \cdot 0.03 = 0.03 \leq 2 \tag{6.2}
\]

For the case \( \beta = 5 \), the dynamics of the system converge to one of two stable equilibria. We find bimodal behavior. In the long term, we have effectively \( f(\phi) \approx +1 \) or \( f(\phi) \approx -1 \). In the former case we effectively recover the "certain" choice in favor of alternative i:

\[
P_{\text{in}} = \frac{e^{\beta f(\phi)}}{e^{\beta f(\phi)} + 1} = \frac{e^{5 \cdot (0.9856)}}{e^{5 \cdot (0.9856)} + 1} = \frac{138.1}{139.1} = 0.9928 \tag{6.3}
\]
Figure 6.1: Plotted graphs of \( y = \varphi \) and \( y = \tanh(1/2) \varphi \) versus \( \varphi \) showing stable equilibrium at \( \varphi = 0 \) for \( \beta = 0.03 \), and unstable equilibrium at \( \varphi = 0 \) for \( \beta = 5 \); both the y-axis and \( \varphi \) are shown in the range from -1 to 1.

In the latter case we effectively recover the "certain" choice against alternative i (ie. in favor of alternative j):

\[
P_{in} = \frac{e^{\beta f(\varphi)}}{e^{\beta f(\varphi)} + 1} = \frac{e^{5\cdot(-0.9856)}}{e^{5\cdot(-0.9856)} + 1} = \frac{0.0072}{1.0072} = 0.0072 \quad (6.4)
\]

This corresponds to the situations where the curve \( y = \tanh(1/2) \beta f(\varphi) \) crosses the line \( y = \varphi \) "from above" respectively at \( \varphi = 0.9856 \approx +1 \) and at \( \varphi = -0.9856 \approx -1 \) in the right panel in Figure 6.1. These solutions are stable. Indeed the inequality in equation (6.1) is satisfied in both cases:

\[
(1 - \varphi^2) \cdot \beta = (1 - 0.9856^2) \cdot 5 = 0.143 \leq 2
\]
\[
(1 - \varphi^2) \cdot \beta = (1 - (-0.9856)^2) \cdot 5 = 0.143 \leq 2 \quad (6.5)
\]

The solution at \( \varphi = 0 \) is unstable, where the curve \( y = \tanh(1/2) \beta f(\varphi) \) crosses the line \( y = \varphi \) "from below." Indeed we see the inequality in equation (6.1) is violated:

\[
(1 - \varphi^2) \cdot \beta = (1 - 0^2) \cdot 5 = 5 > 2 \quad (6.6)
\]

Using the Repast (http://repast.sourceforge.net) multi-agent based modeling platform we created a computational version of the
binary logit model with social interactions and replicated the benchmark analytical results. Figure 6.2 shows a schematic of the working of a single simulation run for the fully connected network and global social influence. Figure 6.3 shows the parameter sweep in the response surface analysis for a set of simulations.

Figure 6.2: The working of a single simulation run for a fully connected network with global social influence

Figure 6.3: The parameter sweep (response surface analysis) for a set of simulations with a fully connected network and global social influence
Example time series for \( N = 100 \) agents with \( f(x) = x \) are shown in Figure 6.4. There are two classes of behavior. With a low value of the parameter \( \beta \), we obtain one class of behavior with a stable equilibrium at \( x = 0 \), as predicted analytically in the left panel of Figure 6.1, with the condition for local stability met in equation (6.2). Conversely, using a high value of the parameter \( \beta \), we obtain another class of behavior with an unstable equilibrium at \( x = 0 \), and with the system being driven away over time with equal probability towards a steady state either the extreme \( x = +1 \) or the extreme \( x = -1 \), as predicted analysis in the right panel of Figure 6.1, with the conditions for local stability met in equation (6.5) and violated in equation (6.6).

Cumulative histograms for 500 runs (i.e. 500 distinct random seeds defining the decision making order) are shown in Figure 6.5, plotting the value of \( x \) after 2000 iterations, reproducing the analytical results.

Figure 6.4: Example time series of \( f(x) = x \) for individual simulation runs with 100 agents in a fully connected network for low certainty \( (\beta = 0.03) \), and for high certainty \( (\beta = 5) \) with two distinct random seeds; the y-axis is shown in the range from \(-1\) to \(1\); the time axis is shown from time step \( t = 0 \) to 2000.

6.1.4 Road Map

In the following sections, we turn towards introducing agent-level heterogeneity in our model. Table 6.1 presents a road map for the remainder of this chapter which we will unfold step by step. Recall in Chapter 3 we have considered the case of global information. At this
Figure 6.5: Histograms of $f(x) = x$ after 2000 iterations over 500 simulation runs (i.e. 500 distinct random seeds) with 100 agents in a fully connected network for low certainty ($\beta = 0.03$) and for high certainty ($\beta = 5$); the bins of the histogram encompass a range from $-1$ to $1$.

Point, we will relax the limiting case of the fully connected network to allow for heterogeneous local information. Until now, we have considered $x_i = n_i/N$ and $x_j = n_j/N$ to be the global proportions of agents who have made each choice. In what follows in the remainder of this chapter, $x_{ln}$ and $x_{jn}$ are the local proportions of agents in a given agent’s ‘reference’ group who have made each choice.

We will define what we mean by a reference group in section 6.2, in terms of a network structure among the agents. Whereas in section 6.1 every agent sees the same global choice proportions, in the rest of this chapter the choice proportions that influence a given agent’s choice are allowed to be heterogeneous. Different agents in general will see different choice proportions depending on their specific reference group, except in the special case of a fully connected network where we recover the results in Chapter 3. The abstract classes of network models with local social influence which we will discuss starting in section 6.2 have the model in section 6.1 as a limit when the link probability is unity.

The empirical model which we will discuss in sections 6.3 and 6.4 goes a step further and subsequently considers a global bias towards one alternative over another, heterogeneous agent characteristics (e.g. gender, travel purpose) as well as agent-specific attributes of the choice alternatives (e.g. travel time, travel cost), all of which additionally influence decision making in a heterogeneous way. While agent characteristics vary across agents, the agent-specific attributes of the choice alternative in general vary across both the alternatives and the agents, e.g. travel time, travel cost are in general different for travel by car versus public transit, and also different for the trip
### Table 6.1: Response surface analysis in step-by-step development of multi-agent based simulation model

<table>
<thead>
<tr>
<th>Section 6.1</th>
<th>Section 6.2</th>
<th>Section 6.3</th>
<th>Section 6.4</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Network</strong></td>
<td><strong>Parameter</strong></td>
<td><strong>Value</strong></td>
<td><strong>Value</strong></td>
</tr>
<tr>
<td>Network class</td>
<td>Erdos-Renyi</td>
<td>n/a</td>
<td>n/a</td>
</tr>
<tr>
<td>Link probability</td>
<td>0.1</td>
<td>0.9</td>
<td>0.9</td>
</tr>
<tr>
<td>Endogenous</td>
<td>27</td>
<td>100</td>
<td>100</td>
</tr>
</tbody>
</table>

**Socio-Dynamic Binary Logit: Application**

**Table 6.1:** Results of step-by-step development of multi-agent based simulation model.
made by agent 1 from origin A to destination B versus the trip made by agent 2 from origin C to destination D. In the special case that no bias, no agent characteristics and no agent-specific attributes are considered, we again recover the model in section 6.1 as a limit.

The preliminary assumptions that we have made until now have been useful for the purpose of model verification in the incremental process of building up the agent based model, step-by-step. If we had immediately considered the case of full heterogeneity in sections 6.3 and 6.4, the dimensionality of the system of equations to be solved as a benchmark would theoretically be as large as the number of agents in the system without simplifications. This said, for the special limiting case of a network with link probability approaching unity, we can indeed derive some additional approximate mean field analytical results as a guidepost which we will return to at the end of this chapter.

6.2 Abstract Classes of Heterogeneous Local Information

Aoki’s model described in Chapter 3 assumes uniform, global and perfect information access. The very fact that certain influences are transferred via social interactions, and thus via social networks implies heterogeneous local information. Therefore, in the following we extend the model to explicitly model interaction networks. See Figure 6.6.

In our model each decision making entity \( n \) is assigned a set of "reference" decision making entities influencing its choice. In section 6.2, we will discuss different two ways of making this assignment during the initialization phase of the multi-agent based model. Then at each time step during the iteration phase, the decision making entities look at the choices their particular reference entities made in the previous round, plus their own choice, and calculate localized values of the difference in systematic utility between the alternatives:

\[ V_{in} - V_{jn} \equiv \beta f(x_{in} - x_{jn}) = \beta f(x_n) \]  \hspace{1cm} (6.7)

The critical difference between equation (6.7) here and equation (3.3) on page 27 considered in the initial presentation of the field effects model at the outset of Chapter 3 is that subscript \( n \) now becomes important in determining \( x_n = x_{in} - x_{jn} \). The "reference" relationships introduced here define a graph or network. Let us denote this graph by \( G = (N, L) \), where \( N \) is the set of nodes (vertices) and \( L \) is the set of links (edges) between them. In our case, each decision making entity is a node, i.e., \( N = \{1, \ldots, n, \ldots, N\} \) and a decision maker’s "reference" entities are defined by its links \( L_n \). We assume that edges are symmetric.
Figure 6.6: The working of a single simulation run for a network with heterogeneous local social influence

It is hypothesized that different network structures yield different system behavior. In practice however, it can be difficult to reveal the exact details of the relevant network(s) of reference entities influencing the choice of each decision making entity. Moreover, the actual reference entities for a given decision making entity may not be among those in the data sample. One way to test the above hypothesis theoretically even without reliable empirical information about the social influence network is by studying abstract classes of networks in the hope of identifying classes of networks that yield similar results.

6.2.1 Erdős-Rényi Networks

An early abstract model of social interaction is due to Erdős and Rényi [1959]. Their random network consists of a number of nodes and set of random edges between them, such that the probability of the existence of a given link is uniform across all possible edges. The actual number of the links is determined by the density $p$ of the network, which is usually perceived as a parameter of the Erdős-Rényi graph. Here network density $p$ is defined as the ratio of the number of actual existing links to the number of all theoretically possible links in a fully connected network with the given number of nodes. Otherwise said, $p$ is the "link probability," the probability that a link exists.
One advantage of studying random networks is that they are perhaps the simplest possible networks that are general enough to describe a wide range of graphs, from unconnected nodes to a fully connected network (i.e., a graph that contains all possible links, as in our initial implementation of the Aoki model in section 6.1). In addition, they accomplish this without introducing any explicit bias into the structure of the network. Moreover, results are known about important properties such as at approximately what value of $p$ will the network become connected (i.e., when each node is "reachable" along the edges from any other node), or otherwise said, when a so-called "giant component" will emerge. Finally, an important feature of random networks which is observed in real-life social networks is the so-called "small world" property: the average path length $l$ (the average number of "hops" between an arbitrary pair of nodes) is less than or of the order $\ln(N)$, where $N$ is the number of nodes. Figure 6.7 shows the response surface analysis for a set of simulations using the Erdős-Rényi network for assignment of reference groups with heterogeneous local social influence.

![Response Surface Analysis](image)

Figure 6.7: The parameter sweep (response surface analysis) for a set of simulations with heterogeneous local social influence on an Erdős-Rényi random network.

With the discrete choice model on a fully connected network, both analytical results and computational replications (top panel in Figure 6.4, and left panel in Figure 6.5) confirm that for $f(x) = x$, values of the certainty parameter close to zero will yield $\beta f(x) \approx 0$ and we get effectively a "fair coin toss" between the alternatives. The probability of choosing alternative $i$ in the binary case is one-half. This is due to the fact that the relative "certainty" of others’ choices becomes low...
and thus there is accordingly low importance in the valuation of the utility. Therefore, there is no reason to believe that the structure of the underlying interaction network would play any significant role in determining the outcome. In order to understand the effect of network structure, it is much more interesting to take a high certainty case, to see where the bifurcation arises as the network structure is varied. The following experiments were therefore carried out with a high value of the certainty parameter, $\beta = 5$. For cases where the network density $p$ approaches unity, we should recover the results from Chapter 3, another verification check in the modelling process.

As a base consistency check, testing our model with $p = 1.0$ should yield a model equivalent to the one discussed in the section 6.1. Simulations confirm this expectation. With low connectivity, however, we get a different picture. Figure 6.8 shows results for density values ranging from 0 to 0.025 ($p = 0, 0.005, 0.010, 0.015, 0.020, \text{ and } 0.025$). It is clear that low densities yield behavior similar to that of low certainty, while higher $p$ values display tendencies toward the high certainty outcome of the global information model.

The density values above were chosen to embrace the critical point when a giant component emerges, i.e., when, in practical terms, the graph becomes connected. It is known (Molloy and Reed, 1998) that this occurs around $p = 1/N + \delta$, where $N$ is the number of nodes and $\delta > 0$ is a small value. In case of 100 agents this formula gives $p = 0.01$ as the critical point. Indeed, simulation results show "random outcomes" for sub-critical densities. Also, a significant change in behavior occurs around a density of 0.02. The range in between yields more ambiguous outcomes, which may warrant further study. Nonetheless, the overall picture suggests that it doesn’t really matter what structure the interaction network has; as long as it is connected, it starts yielding outcomes similar to the fully connected graph.

6.2.2 Watts-Strogatz Networks

In the previous subsection we found that random networks yield behavior similar to that based on global aggregate information, given connectedness, or more precisely, the existence of a giant component. This result may be counter-intuitive as it appears relatively easy to craft examples where it is, at least, unlikely that one alternative will eventually reach total dominance. Therefore, we re-visit the unbiased nature of the Erdős-Rényi graph. An important property that is observed in real-life social networks, but not embodied in a random network, is "clustering": i.e. two friends of a certain person are more likely to be mutual friends themselves than an arbitrary pair of individuals.

The abstract network model to overcome this deficiency is the Watts-Strogatz model, which starts from an ordered network, or lat-
Figure 6.8: Histograms of $f(x) = x$ after 2000 iterations over 500 runs with 100 agents in an Erdős-Rényi network for a sweep of network densities from $p = 0$ to $p = 0.025$ by increment $p = 0.005$, with high certainty ($\beta = 5$); the bins of the histogram encompass a range from $-1$ to $1$. 
tice, which contrary to the random network, has high clustering, but long average path length, or in effect, no small world property. The dimensionality of the lattice is a parameter, although only 1D and 2D models are commonly discussed in this context. The extent to which the neighbors are connected is also a parameter of the Watts-Strogatz model. To avoid artificial boundary effects, toric lattices are considered which are “wrapped around,” that is, nodes on the boundary of the system link to nodes at the opposite boundary. “Shortcuts” are then introduced into these systems to create the Watts-Strogatz model, by randomly rewiring a few links. The controlling parameter is the probability of rewiring $w$, which gives the probability that each original link in the system is replaced by a random connection. Only a very few shortcuts, i.e., a fairly low $w$, is needed to achieve the small world property. Figure 6.9 shows the response surface analysis for a set of simulations using the Watts-Strogatz network for assignment of reference groups with heterogeneous local social influence.

Figure 6.9: The parameter sweep (response surface analysis) for a set of simulations with heterogeneous local social influence on a Watts-Strogatz network

Turning our attention back to our discrete choice model, the first thing to consider is the density of the social influence network, now modeled as a Watts-Strogatz graph. In the following we fix the link probability at $p = 0.04$ (1D with extent 2) a network density that is sufficiently high to be in the two-equilibrium regime of the random network model. Indeed, experiments with $w = 1.0$, which renders the Watts-Strogatz model equivalent to that of Erdős and Rényi, confirm this. Varying $w$ from 0 to 0.3 shows, however, that this is not always the case. See Figure 6.10. When rewiring is unlikely, the system stays
in the “fair coin toss” regime. Then, as the average path length $l$ falls below the small world threshold, the outcome converges to the two certain choices behavior.

Figure 6.10: The effect of rewiring $w$ and decreasing average path length $l$ in histograms of $f(x) = x$ after 2000 iterations over 500 runs with 100 agents in a Watts-Strogatz network for constant network density ($p = 0.04$) and high certainty ($\beta = 5$); the bins of the histogram encompass a range from $-1$ to $1$; the "small world threshold" is average path length $\ln(100) \approx 4.6$

6.2.3 Preliminary Conclusions

Our initial results with both Erdős-Rényi and Watts-Strogatz graphs suggest that when a network representing the interactions between a decision making entity and the aggregate behavior of other (local) reference entities has the small world property, the system behaves in the long-run as Aoki’s original model with global mean field infor-
mation. If we are only interested in long-run behavior and not how long the system takes to transition there, testing for the small world property may be an empirically advantageous alternative to collecting data on the precise details of a social network. But what happens if agents have an inherent heterogeneity beyond that induced by local interactions? Is the same still true?

6.3 An empirical example

Having docked our agent based model against analytical mean field results in section 6.1 and having explored the behavior of some abstract classes of networks in section 6.2, in the next step of our model development process we now turn our attention to an empirical application of intercity transportation mode choice behavior. Here we include individual level heterogeneity in two ways.

We use revealed preference survey data collected by the Hague Consulting Group for the Netherlands Railways to assess factors which influence the binary choice between car versus rail for intercity travel. We also test the role of the social influence, modeled as an Erdős-Rényi graph over a full sweep of link probabilities from zero to one. In the limit that the link probability approaches zero, we have a classical binary logit model without social interaction. In the limit that the link probability approaches one, we recover a fully connected network discussed in section 6.1. For the special case of very high link probabilities, we can therefore establish some additional approximate theoretical benchmark results to re-verify our agent based model.

At the outset of the dissertation in section 2, we discussed the notion of a "systematic" utility $V_{in}$ that a given decision making entity $n$ is presumed to associate with a particular alternative $i$. For the purpose of the step-by-step development of our agent based model, we have considered until now the interaction effect as the only term in the systematic utility. In typical transportation applications, the systematic utility is commonly assumed to be defined by a function of observable characteristics $S_n$ of the decision making entity and observable attributes $z_{in}$ of the choice alternative for a given decision making entity:

$$V_{in} = h_i + V(S_n, z_{in})$$

(6.8)

The term $h_i$ is a so-called "alternative specific constant" (ASC), included as good practice to explicitly account for any underlying bias for one alternative over another alternative. In other words, $h_i$ reflects the mean of $\epsilon_{jn} - \epsilon_{in}$, that is, the difference in the utility of alternative $i$ from that of $j$ when all else is equal. Since it is the difference that is relevant, for a general multinomial case with $J$ alternatives we can define a set of at most $J - 1$ alternative specific constants. For our
binary case, we can thus define only one constant, which we will call $h$.

In sections 6.3 and 6.4, we will consider the term $V_{in} - V_{jn}$ in equation (6.9) to have the general form:

$$V_{in} - V_{jn} \equiv \beta f(x_{in} - x_{jn}) = \beta f(x_n) = \beta x_n + h + \gamma' S_n + \zeta_i' z_{in} - \zeta_j' z_{jn}$$

(6.9)

where \(\gamma = [\gamma_1, \gamma_2, \ldots]'\), \(\zeta_i = [\zeta_{i1}, \zeta_{i2}, \ldots]'\) and \(\zeta_j = [\zeta_{j1}, \zeta_{j2}, \ldots]'\) are vectors of unknown utility parameters respectively corresponding to the relevant observable agent characteristics $S_n$, and observable agent-specific attributes $z_{in}$ of the choice alternative such that:

$$\gamma' S_n = \gamma_1 S_{n1} + \gamma_2 S_{n2} + \ldots$$

$$\zeta_i' z_{in} = \zeta_{i1} z_{in1} + \zeta_{i2} z_{in2} + \ldots$$

$$\zeta_j' z_{jn} = \zeta_{j1} z_{jn1} + \zeta_{j2} z_{jn2} + \ldots$$

(6.10)

In general the utility parameters $\zeta_i$ and $\zeta_j$ may take alternative specific values, however in this chapter we will consider only "generic" values of the utility parameters $\zeta \equiv \bar{\zeta}_i = \bar{\zeta}_j$ and define $z_n \equiv z_{in} - z_{jn}$ so that we have the further simplification:

$$V_{in} - V_{jn} = \beta f(x_n) = \beta x_n + h + \gamma' S_n + \zeta' z_n$$

(6.11)

The following is a description based on an (unpublished) Hague Consulting Group report which appeared in Ben-Akiva and Morikawa [1990] of the revealed preference intercity travel survey data that we use in this chapter:

"The city of Nijmegen, in the eastern part of the Netherlands near the border with West Germany, was selected as the data collection site. This city has rail connections with the major cities in the western metropolitan area called the Randstad which contains Amsterdam, Rotterdam and The Hague. Traveling from Nijmegen to the Randstad takes approximately two hours by both rail and car.

The sample consisted of residents of Nijmegen who: made a trip in the previous three months to Amsterdam, Rotterdam or The Hague; did not use a yearly rail pass, or other types of pass which would eliminate the marginal cost of the trip; had the possibility of using a car (ie. possessed a driver’s license and had a car available in the household); and had the possibility of using rail (ie. did not have very heavy baggage, were not handicapped, and did not need to visit multiple destinations).

Qualifying residents of Nijmegen were identified in a random telephone survey and requested to participate in a
home interview. 235 interviews were conducted out of the 365 people who were reached by telephone and satisfied the above criteria. ... The home interview first asked about a recent intercity trip from Nijmegen to the Randstad. The respondents were requested to report the characteristics of that trip and those of a trip to the same destination but with the other (non-chosen) mode."

The intercity travel survey data thus have 235 observations, corresponding to 235 distinct heterogeneous agents in our model. Descriptions of the survey variables available for use in our modeling endeavor and their summary statistics are given in Table 6.2. There are no reported missing values.

As in section 4.2, in our agent based model each decision making entity $n$ is again assigned a set of "reference" decision making entities influencing its choice. See Figure 6.11. At each time step, the decision making entities look at the choices their particular reference entities made in the previous round, plus their own choice, and calculate localized values of the difference in systematic utility between the alternatives:

$$V_{in} - V_{jn} = \beta x_n + h + \gamma_1 \cdot \text{gender}_n + \gamma_2 \cdot \text{business}_n + \gamma_3 \cdot \text{shoprec}_n + \zeta_1 \cdot (t_{tcar} - t_{trail})_n$$
$$+ \zeta_2 \cdot (t_{tcar} - t_{trail})_n + \zeta_3 \cdot (o_{tcar} - o_{trail})_n \tag{6.12}$$

Figure 6.11: The working of a single simulation run for a network with heterogeneous local social influence, node characteristics and choice attributes per node
Table 6.2: Variables in Hague Consulting Group intercity travel survey data.

<table>
<thead>
<tr>
<th>NAME</th>
<th>TYPE</th>
<th>DESCRIPTION</th>
<th>MEAN</th>
<th>STD DEV</th>
<th>MIN</th>
<th>MAX</th>
</tr>
</thead>
<tbody>
<tr>
<td>choice</td>
<td>$y_n$</td>
<td>1 if rail was used on the recent intercity trip; 0 if car</td>
<td>0.366</td>
<td>0.483</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>gender</td>
<td>$S_n$</td>
<td>1 if respondent is female; 0 if male</td>
<td>0.443</td>
<td>0.498</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>business</td>
<td>$S_n$</td>
<td>1 if trip purpose is business; 0 otherwise</td>
<td>0.247</td>
<td>0.432</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>shoprec</td>
<td>$S_n$</td>
<td>1 if trip purpose is shopping or recreation; 0 otherwise</td>
<td>0.400</td>
<td>0.491</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>ttcar</td>
<td>$z_{in, i = car}$</td>
<td>In-vehicle travel time for car alternative</td>
<td>98.5</td>
<td>21.7</td>
<td>45</td>
<td>180</td>
</tr>
<tr>
<td>tccar</td>
<td>$z_{in, i = car}$</td>
<td>Travel cost in NLG for the car alternative</td>
<td>31.45</td>
<td>23.74</td>
<td>1</td>
<td>135</td>
</tr>
<tr>
<td>ovtcar</td>
<td>$z_{in, i = car}$</td>
<td>Out-of-vehicle time to walk from parking place to destination</td>
<td>5.2</td>
<td>6.4</td>
<td>0</td>
<td>50</td>
</tr>
<tr>
<td>ttrail</td>
<td>$z_{jn, j = rail}$</td>
<td>In-vehicle travel time for the rail alternative</td>
<td>97.1</td>
<td>21.1</td>
<td>15</td>
<td>180</td>
</tr>
<tr>
<td>tcrail</td>
<td>$z_{jn, j = rail}$</td>
<td>Travel cost in NLG for the rail alternative</td>
<td>35.31</td>
<td>11.83</td>
<td>1</td>
<td>125</td>
</tr>
<tr>
<td>ovtrail</td>
<td>$z_{jn, j = rail}$</td>
<td>Out-of-vehicle time for the rail alternative (access plus egress)</td>
<td>32.7</td>
<td>15.0</td>
<td>5</td>
<td>90</td>
</tr>
</tbody>
</table>

The trip purpose indicators are derived from a 3-level categorical variable: business; shopping/recreation; other. All travel times are given in minutes; all travel costs are given in Dutch guilders (NLG).
We are interested in how the dynamics of the discrete choices of these heterogeneous individuals depend on the structure of the underlying social influence network. Figure 6.12 shows the response surface analysis for a set of simulations using the Erdős-Rényi network for assignment of reference groups with heterogeneous local social influence, agent characteristics and agent-specific choice attributes. We vary the network density \( p \) on the parameter range \((0, 1)\) ranging from a non-connected to a fully connected graph, excluding endpoints. In each case, we repeatedly situate the agents in 20 distinct instantiations of an Erdős-Rényi graph per network density value. As in section 6.2 we might expect a transition in long-run sociodynamic behavior to occur at network density \( p = 1/N + \delta \), where \( N \) is the number of nodes and \( \delta > 0 \) is a small value. In case of 235 agents, this formula gives \( p \approx 0.005 \) as a lower bound for a critical point. However given that previously we saw a significant change in behavior only at \( p = 0.02 \) with 100 agents, we might also similarly expect interesting changes in behavior to become noticeable only with values of \( p \) significantly higher than the lower bound. Keeping this in mind, we select the following 30 values of \( p \) to sample: 0.005 to 0.100 at increment 0.005 (20 network density values), 0.200 to 0.600 at increment 0.100 (5 network density values), 0.700 to 0.900 at increment 0.050 (5 network density values).
For these 600 networks (30 network density values times 20 network instantiations per density value), we repeatedly compute the local interaction field variable $x_n$ for each of the 235 agents in the sample in two ways, one with counting the agent’s own choice in its reference group (that is, with so-called “self loops”), and one without counting the agent’s own choice in its reference group (that is, without self loops). The econometric considerations regarding these two approaches to the local interaction field variable $x_n$ are discussed in section 6.4. From a theoretical perspective, the model with self loops is interesting because in the limiting case of network density $p = 1$ (a fully connected network), we recover Aoki’s original model if there were no other explanatory variables in the utility function. Likewise from a theoretical perspective, the model without self loops is interesting because in the limiting case of network density $p = 0$ (a non-connected network), we have pure random behavior if there were no other explanatory variables in the utility function. From a multi-agent based simulation perspective, the model with self loops at low network density might logically provide inertia in the behavior, damping down the volatility of switching from one choice to another. At high network density, when the number of agents is large, there is not likely to be discernible difference in the multi-agent based simulation behavior between the model with self loops and without self loops.

After some preliminary model specification testing, we proceed to repeatedly estimate sets of the utility parameters $\beta, \gamma_1, \gamma_2, \gamma_3, \zeta_1, \zeta_2, \zeta_3$ in equation (6.12) via maximum likelihood estimation for each of the 1200 network scenarios described above, with two different binary logit utility specifications: (i) with the alternative specific constant $h$ freely estimated, and (ii) with the alternative specific constant constrained to zero. Using the distinct sets of coefficients for each of the 2400 estimated models, we then run 50 multi-agent based simulations with distinct pseudo-random number sequences for 2000 iterations per each run, per each model. The value of $x$ representing the difference between the aggregate mode shares $x_i$ and $x_j$ in the sample at the last time step of each run is counted in histograms, one per each network instantiation. We group these histograms by network density in each of the four experimental settings (with or without self loops and with and without an alternative specific constant). Figure 6.13 shows aggregated histograms from selected groups from each experimental setting with corresponding network density values.

According to our results from section 6.2, we might potentially expect to see behavioral transitions occurring somewhere in the low network densities. However, the most striking result is that only with the model with the alternative specific constant and with self loops do we ever get the signature 2-peak histogram which we know from the analytical benchmark, and abstract experimentation with the Erdös-
Figure 6.13: Selected aggregated histograms for a sweep of network density $p = 0.005$ to $p = 0.9$ for four model specifications with and without an alternative specific constant and with and without self loops; there are 235 heterogeneous agents in an Erdős-Rényi network with 20 network instantiations per density value, 50 multi-agent based simulation runs per network instantiation, and 2000 iterations per simulation run; the values of the set of utility parameters are re-estimated per each network instantiation; the bins of the histogram encompass a range from $-1$ to $1$. 
Rényi and Watts-Strogatz graphs. Thus we can conclude that adding additional agent-specific heterogeneity in our model beyond the heterogeneity automatically induced by the localized interactions does indeed seem to matter.

A closer analysis of the histograms reveals a significant insight: not all of the network instantiations for the models with the alternative specific constant and with self loops give the signature 2-peak histogram. In fact at higher network densities we see both the single sharply-peaked distribution and the signature 2-peak histogram co-existing as outcomes with the alternative specific constant and with self loops, giving a hint at some kind of instability in the model. See Figure 6.14. This insight motivates a further analysis of the estimated coefficient values which we will undertake in section 6.4.

![Histograms](image)

**Figure 6.14:** Selected detailed histograms for four Erdős-Rényi network instantiations with different network generator random seeds at network density $p = 0.9$, for the model specification with an alternative specific constant and with self loops; there are 235 heterogeneous agents, 50 multi-agent based simulation runs per network instantiation, and 2000 iterations per run; the value of the model coefficients are re-estimated per each network instantiation; the bins of the histogram encompass a range from $-1$ to $1$.

For the case shown in Figure 6.14 with $p = 0.9$, we may suppose that the network density is close enough to approaching unity that the mean field analytical results in section 6.1 may be relevant as
an approximate guidepost. Using the mean values of variables given in Table 6.2 and the estimated utility parameters $\beta, h, \gamma_1, \gamma_2, \gamma_3, \zeta_1, \zeta_2, \zeta_3$ for each of the network instantiations with random seeds shown in Figure 6.14, we plot the left-hand-side and the right-hand-side of equation on a graph, and find their intersection for $f(xn) =Vin-Vjn = \beta xn + h + \gamma' Sn + \zeta' zn$ as given in equation (6.12). In Figure 6.15, we can see that the effect of adding the alternative specific constant, the agent characteristics and the agent-specific attributes of choice alternatives to the model is to shift the tanh curve horizontally so that the curve no longer crosses the line $y = \varphi$ at $\varphi = 0$.

Figure 6.15: Plotted graphs of $y = \varphi$ and $y = \tanh(1/2)[\beta x_n + h + \gamma S_n + \zeta z_n]$ versus $\varphi$ for values of estimated utility parameters $\beta, h, \gamma_1, \gamma_2, \gamma_3, \zeta_1, \zeta_2, \zeta_3$ for the specific network instantiations with the same network generator random seeds shown in Figure 6.14; stable equilibria are seen where the tanh curves the line $y = \varphi$ “from above,” and unstable equilibrium are seen where the tanh curves the line $y = \varphi$ “from below”; both the $y$-axis and $\varphi$ are shown in the range from $-1$ to $1$.
A larger value of the certainty parameter $\beta$ is accordingly necessary to achieve the signature bimodal behavior. We see this is true for both a shift of the tanh curve to the right as shown in the left panel of Figure 6.16 where we have:

$$\frac{1}{\beta} (h + \gamma' S + \zeta' z) = -0.226 < 0$$  \hspace{1cm} (6.13)

and a shift of the tanh curve to the left as shown in the right panel of Figure 6.16 where we have:

$$\frac{1}{\beta} (h + \gamma' S + \zeta' z) = 0.132 > 0$$  \hspace{1cm} (6.14)

In the former the transition from unimodal behavior to bimodal behavior occurs at approximately $\beta \approx 3.5$ and in the latter it occurs at approximately $\beta \approx 3$, depending on the absolute magnitude of the shift.

Figure 6.16: Plotted graphs of $y = \varphi$ and $y = \tanh(\frac{1}{2}[\beta \varphi + h + \gamma S + \zeta z])$ versus $\varphi$ for values of estimated utility parameters $\beta$, $h$, $\gamma_1$, $\gamma_2$, $\gamma_3$, $\zeta_1$, $\zeta_2$, $\zeta_3$ for the specific network instantiations with network generator random seeds 1 and 9 at top left and bottom left panels of Figure 6.15, with a sweep of the certainty parameter $\beta$ embracing the transition from unimodal to bimodal behavior; both the $y$-axis and $\varphi$ are shown in the range from $-1$ to $1$
6.4 SOME ISSUES IN ESTIMATION

We plot the sets of estimated coefficient values for the same 30 network density values swept in section 6.3. Figure 6.17 shows selected detail from the estimated coefficient values for the four model specifications with and without an alternative specific constant and with and without self loops (20 network instantiations per density value).

Analyzing the plots, the clue to our puzzling behavior in section 6.3 becomes obvious in light of our analytical results. Far from being constant across all estimated models, we see instead systematic variation in the estimated coefficient values. From section 6.2 we know that the coefficient on the interaction variable must be sufficiently large and positive in order to trigger the signature 2-peak histogram long-run behavior. What we see is that for many of the models, this coefficient on the local interaction variable is in fact negative. In such case we can never expect to see the signature 2-peak histogram.

To test the hypothesis of whether the results in section 6.3 may have more to do with the particular value of the estimated coefficient on the local interaction variable rather than the density of the network per se, we construct the following experiment. We choose a set of estimated coefficients with a fairly high value of the coefficient on the interaction variable. See Table 6.3. In this case, the coefficients correspond to the network instantiation with network generator random seed = 1 for network density \( p = 0.6 \) for the model with an alternative specific constant and with self loops. It is our first observation of sharp bimodal behavior as we sweep the link probability from \( p = 0.005 \) to \( p = 0.9 \), but we could have used any arbitrarily chosen set of coefficients with sufficiently high coefficient on local interaction variable to yield the signature 2-peak histogram for long-run behavior. Now suppose there is an exogenous shock to the system whereby the network density \( p \) changes, but the estimated coefficients are not re-calibrated. Figure 6.18 shows the response surface analysis for a set of simulations using the Erdős-Rényi network for assignment of reference groups with heterogeneous local social influence, agent characteristics and agent-specific choice attributes, with fixed utility parameters.

We plot groups of histograms resulting from a parameter sweep of \( p \) from 0.005 to 0.100 at increment 0.005, with 20 network instantiations per density. Figure 6.19 shows selected groups of aggregated histograms from this series confirming that here we do indeed consistently get the originally predicted transition in long-run sociodynamic behavior for all 20 network instantiations at network density \( p = 1/N + \delta \), with \( N \) the number of nodes and \( \delta > 0 \) a small value.

There are two subtle issues to understand about the estimation. One issue has to do with correlation of explanatory variables. The other issue has to do with an explanatory variable or linear combi-
Figure 6.17: Estimated coefficient values for the four model specifications with and without an alternative specific constant, and with and without self loops; there are 20 network instantiations per density value for a sweep of network density $p = 0.005$ to 0.9.
nation thereof being (almost) a perfect predictor for the dependent variable. What we must recognize about the local interaction variable is that in networks with a large number of agents and at high network density, the "local" interaction variable will become effectively constant across agents in the network, whereby the variable will become highly correlated with the value of unity included in the model when estimating an alternative specific constant. This leads to a violation in the estimation process. In fact the local interaction variable will be perfectly correlated with unity at network density $p = 1$ (a fully connected network) when the model includes self loops.

In Figure 6.17, we can visually track the increasing correlation as the network density increases, between the coefficient on the local interaction variable the alternative specific constant (ASC) for rail in the models with the ASC. To see this more clearly, it is illustrative to recast the estimated coefficient values for the local interaction variable, in terms of an ensemble of static plots of the estimated coefficient value for the local interaction variable versus network density, by network instantiation. See Figure 6.20. In Figure 6.21 we also show the correlation value versus network density for the same ensemble of plots per network instantiation.

When the model does not include an alternative specific constant, the local interaction variable takes on this function at high densities. Since the alternative specific constant happened to be positive for this case study in the baseline model without any interaction, simple cal-

Table 6.3: Example binary logit model estimation results with local field variable

<table>
<thead>
<tr>
<th>Estimated Parameters</th>
<th>Value</th>
<th>Std Err</th>
<th>T-Stat</th>
</tr>
</thead>
<tbody>
<tr>
<td>Share of agent’s reference group choosing each mode, defined generically</td>
<td>9.63</td>
<td>3.64</td>
<td>2.65</td>
</tr>
<tr>
<td>Rail bias (alternative specific constant)</td>
<td>3.14</td>
<td>1.10</td>
<td>2.85</td>
</tr>
<tr>
<td>Female bias for rail</td>
<td>0.972</td>
<td>0.351</td>
<td>2.77</td>
</tr>
<tr>
<td>Business trip bias for rail</td>
<td>1.05</td>
<td>0.438</td>
<td>2.39</td>
</tr>
<tr>
<td>Shopping/recreation trip bias for rail</td>
<td>-0.951</td>
<td>0.420</td>
<td>2.27</td>
</tr>
<tr>
<td>In-vehicle travel time</td>
<td>-0.0121</td>
<td>5.98e-3</td>
<td>2.03</td>
</tr>
<tr>
<td>Travel cost</td>
<td>-0.0253</td>
<td>7.27e-3</td>
<td>3.48</td>
</tr>
<tr>
<td>Out-of-vehicle travel time</td>
<td>-0.0589</td>
<td>0.0124</td>
<td>4.75</td>
</tr>
<tr>
<td>Null log likelihood ($L_0$)</td>
<td>-162.9</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Final log likelihood</td>
<td>-111.9</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Likelihood ratio test</td>
<td>102.0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rho-squared ($\rho^2$)</td>
<td>0.313</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
calculation can show that the coefficient on the local interaction variable will be negative for this case study in models without an alternative specific constant for high network density, since the "local" interaction variable itself will be negative (the sample mode share for rail is less than the sample mode share for car). This is the very simple reason for why we never saw the signature 2-peak histogram in the movies in section 6.3 for the models without an alternative specific constant at high network density.

At low network densities in models with self loops, we have a problem in that the local interaction variable will be almost a perfect predictor for the dependent variable, particularly if we do not have time series data. At high network density in models without self loops and with an alternative specific constant, we are confronted with a double effect: not only is the local interaction variable almost perfectly correlated with unity whereby we have a violation in estimation due to the correlation between explanatory variables, but additionally a precise linear combination of the local interaction variable and unity is actually highly correlated with the choice variable itself, leading to the second violation in estimation. This linear combination is a perfect predictor for the model without self loops and with an alternative specific constant when network density $p = 1$. 
Figure 6.19: Selected aggregated histograms with constrained coefficients for a sweep of network density $p = 0.005$ to $p = 0.1$ for the model specification with an alternative specific constant and with self loops; there are 235 heterogeneous agents in a Erdős-Rényi network with 20 network instantiations per density value, 50 multi-agent based simulation runs per network instantiation, and 2000 iterations per simulation run; the values of the set of utility parameters are held fixed across network instantiations; the bins of the histogram encompass a range from $-1$ to $1$. 
Figure 6.20: Ensembles of plots for 20 network instantiations depicting the envelope of values of estimated local interaction coefficient versus network density, for models with an alternative specific constant, with and without self loops.

Figure 6.21: Ensembles of plots for 20 network instantiations depicting the envelope of values for the correlation between local interaction coefficient and alternative specific constant, versus network density, for models with an alternative specific constant, with and without self loops.
Figure 6.22: Ensembles of plots for 20 network instantiations depicting the envelope of values for rho-squared (model fit) versus network density
To see this more clearly, we plot the rho-squared values for the model fit versus network density for the ensemble of 20 network instantiations for all four classes of models with and without self loops and with and without an alternative specific constant. See Figure 6.22.

6.5 CONCLUSIONS AND RECOMMENDATIONS FOR FURTHER RESEARCH

In this chapter, we have explored a multi-agent based model of discrete choices with interdependence of decision makers’ choices. By applying the model to an example of intercity travel demand using empirical data, we introduced individual agent heterogeneity beyond that induced by the local interaction structure. We found that the model’s characteristic phase transition is dependent on network density and clustering in examples with Erdős-Rényi graphs and Watts-Strogatz graphs, and on the importance of the estimated value of the coefficient for the local interaction variable relative to other coefficients in the binary model. For a specific set of coefficients the appearance of a phase transition is robust against different instantiations of a random network at given network density, but the estimation process to determine the set of coefficients can be highly sensitive to the small variations in the different instantiations, particularly in models including an alternative specific constant.

Special care must be taken in estimation of empirical models:

- With networks with very low densities when the model includes self loops; and
- With networks with very high densities when the model includes an alternative specific constant (ASC), especially in a model without self loops.

In general, preference goes to models with an ASC in order to ensure the error terms in the utility function have zero mean and the estimated coefficients are unbiased. Whether self loops are implemented or not in an empirical model depends on the rationale of the system, and ideally on availability of panel data over multiple time periods.

In addition to this central contribution, there are a number of sub-results which we believe are important to highlight as good practice with agent based modelling. One key feature of agent based modelling is internal verification, or otherwise said, how can the researcher be confident that the agent based model is performing the actions that it is expected to do? What is the evidence that the programming implementation of the abstract or conceptual model is correct? To address this fact, we begin our modelling endeavor with a very simplified model studied previously by others, building up our agent based model step by step, and adding different layers of com-
plexity one at a time. In our case, this means adding different kinds of heterogeneity to the agent based model.

In section 6.1, the dynamics of the model are driven by choices made by the agents. For this simple model, we are able to calculate an analytical solution. We show that the agent based model yields the same results as the analytical benchmark. This gives us confidence that the subtleties of scheduling, event simulation and sequences of random draws in our model behave exactly as expected. We can then proceed to add additional complexity.

In section 6.2, there is additional heterogeneity due to the network structure and the fact that agents have local information, rather than global information. We experiment with two different abstract classes of networks to see the effect of density and of clustering. Our results with Erdős-Rényi and Watts-Strogatz graphs suggest that when a network representing the interactions between a decision making entity and the aggregate behavior of other (local) reference entities has the small world property, the system behaves in the long-run as the original analytical model with global mean field information. If we are only interested in long-run behavior and not how long the system takes to transition there, testing for the small world property may be an empirically advantageous alternative to collecting data on the precise details of a social network.

In section 6.3, we finally add heterogeneity due to individual characteristics of agents (gender) as well as agent-specific attributes of choice alternatives (travel purpose, travel time, travel cost). We find that adding this additional layer of heterogeneity beyond the heterogeneity induced by the local information of the behavior of other agents in a given agent’s reference group does indeed matter. The additional layer of heterogeneity dampens the social effect of the flocking behavior. The parameter beta must be significantly higher than the case in section 6.2 in order to recover the signature two-peak histograms.

Whereas in section 6.3, we re-estimated the coefficients driving the social dynamics for each network instantiation at each experimental level of level network density, in section 6.4, we now consider the case of an exogenous shock to the social system whereby the network density is varied, but the estimated coefficients are not re-calibrated. Choosing a set of estimated coefficients with significantly high positive beta, we do recover the signature two-peak histograms for all network instantiations at a given level of network density. We conclude therefore that the absence of this signature behavior in many of the runs in section 6.3 has to do with the initial values of the estimated coefficients, and in particular the sensitivity of the estimation process at low network density and high network density depending on whether or not self loops are considered.
The results presented in this chapter form the basis of further research, investigating, for example, what class of networks may be needed to address interactions between identifiable decision making entities. Technically, these may also be well-modeled by interaction graphs. This is an interesting question in particular, as the current results suggest that more sparse networks are more dependent on the actual reference structure. Another important direction concerns discrete choices on abstract network models with more realistic degree distributions (in this case, the distribution of the number of "references" the decision making entities have). Both the Erdős-Rényi graph and Watts-Strogatz graph yield a Poisson degree distribution, which is unlike various real-life cases.