Segregation in Social Networks of Heterogeneous Agents Acting under Incomplete Information

D. Kai Zhang
Imperial College London
United Kingdom
dz819@imperial.ac.uk

Alexander Carver
Imperial College London
United Kingdom
alexander.carver15@imperial.ac.uk

ABSTRACT

We propose an agent-based network formation model under uncertainty with the objective of relaxing the common assumption of complete information, calling attention to the role beliefs may play in segregation. We demonstrate that our model is capable of generating a set of networks that encompasses those of a complete information model. Further, we show that by allowing agents to hold biased beliefs toward each other based on social group membership, individual utility-maximising decisions may lead to group segregation at the cost of social welfare. We accompany our theoretical results with a simulation-based investigation of the relationship between beliefs and segregation and show that biased beliefs are an important driver of segregation under incomplete information.

KEYWORDS

Social Networks; Segregation; Incomplete Information; Heterogeneous Agents

1 INTRODUCTION

Members of the same group frequently exhibit similar outward appearances – a phenomenon referred to as homophily – which in its extreme form encompasses segregation. We provide an approach to studying this feature of social networks under uncertainty, building on an agent-based network formation model first proposed by Jackson and Wolinsky [14].

A number of extensions have been made to study segregation from a network formation perspective. These extensions have in common that agents belong to different types, where a type denotes shared attributes. Further, agents may exhibit heterogeneity in either benefits or costs when connecting to other agents. Specifically, connecting with someone of the same type is usually assumed to either offer greater benefits or come at a lower cost than connecting when types differ. These models show that segregation equilibria exist when agents exhibit even small preferences for homophily.

1.1 Our Contribution

The key shortcoming in most of the existing literature (Section 2) is the strong assumption that agents have complete information about

each other's types or, equivalently, utilities depend only on public attributes. This paper departs from this assumption by introducing a distinction between social groups, which subsume all public attributes, and private types, encompassing private attributes. Agents are assumed to prefer connecting to other agents of the same type which, however, is unobservable in advance. Agents are thus indifferent over public attributes. Instead, agents heuristically use social group memberships to gauge someone's private type. This, in turn, allows for "errors" in people's judgement of each other's types.

We formally present this model in Section 3. Section 4 provides an analysis of the networks that can form under this model and demonstrates that segregation between social groups can simply occur when people are sufficiently biased, i.e., exhibit beliefs that are detached from the true distribution of private types. Finally we implement simulations (Section 5) to provide experimental comparisons with complete information networks and incomplete information networks with rational beliefs, i.e., beliefs that coincide with the true type distribution. The results in Section 6 illustrate that even mildly biased beliefs are an important driver of segregation under incomplete information. Section 7 concludes.

2 RELATED LITERATURE

Since the landmark Schelling segregation models [20, 21], numerous studies have analysed the impact group membership can have in homophilous social networks. This has been examined, amongst others, in the context of migration and assimilation decisions [2, 4, 19, 25], occupation [9, 18, 26], and even marriage [22].

Jackson and Wolinsky [14] contributed a seminal model of the process by which such networks can form based on individuals maximising utility. The network is represented as the tuple (N, g) consisting of agents N and an undirected graph g in which nodes are agents and edges are connections. The graph g can then be simplified as the set of pairs ij, $i, j \in N$, between which there is an edge. Given a network, agent i has a utility function $u_i(g)$ given by:

$$u_i(g) = \sum_{j \neq i} \delta^{d_{ij}(g)} - \sum_{j:\, ij \in g} c \qquad (1)$$

where the benefit for i to be connected to j in graph g depends on a factor $\delta \in (0, 1)$ which decays with the number of edges on the shortest connecting path, $d_{ij}(g)$. Agent i, further, pays a cost $c > 0$ for every edge they maintain.

In this network, agents form and sever edges until the network reaches a pairwise stable equilibrium (PSE), defined as:

(i) $\forall ij \in g, u_i(g) \geq u_i(g \setminus ij)$ and $u_j(g) \geq u_j(g \setminus ij)$, and

(ii) $\forall ij \notin g$, if $u_i(g \cup ij) > u_i(g)$ then $u_j(g \cup ij) < u_j(g)$ $\qquad (2)$

which specifies that (i) no agent should be better off severing an existing edge and (ii) no pair of agents could improve utility individually by forming a new edge. This framework offers the ability to associate different network patterns with social welfare, which can be defined as the sum of individual utilities, i.e., $U(N, g) = \sum_{i \in N} u_i(g)$.

Note that this is a homogeneous agent model. Benefits and costs merely depend on agent locations within the network and not on identities or types. A result of this assumption is that networks must either be empty or fully connected (i.e., an edge exists between any two agents) in equilibrium [14]. Networks with separate groups of individuals (or "components") that coexist cannot be generated, making them unsuitable for the study of group segregation.

Jackson and Rogers [11] address this limitation and introduce heterogeneity by assuming that agents belong to K different islands with inter-island edges being more expensive than intra-island edges. The utility function becomes:[1]

$$u_i(g) = \sum_{j \neq i} \delta^{d_{ij}(g)} - \sum_{j: \, ij \, \in \, g} c_{ij} \qquad (3)$$

where $c_{ij} = c_L > 0$ if i and j belong to the same island and $c_{ij} = c_H > c_L$ if not.

The inclusion of different agent types (by way of island occupation) and heterogeneous costs is the key assumption that enables equilibria in which agents are segregated by type. Significantly, their model implies that segregation is, in fact, optimal for social welfare – a conclusion also reached in other works [1, 7].

These works, however, share the strong assumption of complete information, which for instance means that agents already know each other perfectly when deciding to form a new connection. This paper belongs to a body of literature that relaxes this assumption. Within this literature, the focus, however, often lies on the impact of incomplete information on games played in an *existing* network [6, 15] which therefore differs from the focus here on network *formation* under incomplete information.

The approach of Song and van der Schaar [23] and Zhang and van der Schaar [27] is closest to ours. Both consider a network formation model in which agents possess heterogeneous types which are unknown in advance. The former, however, assumes agents with homogeneous beliefs while this work permits belief heterogeneity. Their work also mainly focuses on topological implications for the network. The latter work extends the former but assumes rational beliefs tied to the true distribution of types. We instead draw inspiration from works showing that humans frequently exercise heuristic thinking in deviation of rational expectations [5, 10, 17, 24] and allow beliefs to be biased.

3 THE INCOMPLETE INFORMATION MODEL

This section introduces the network formation model with heterogeneous agents who possess incomplete information about each other's private types. The key addition in our model is the concept of *beliefs* over these types based on public social group membership which impact the expected cost of forming an edge.

---

[1]Jackson and Rogers [11] introduce a free parameter D to truncate benefits. We do not consider truncation and implicitly set $D = \infty$ here for presentation.

3.1 Design

Here we introduce the idea that agents belong to distinct *social groups* which share public characteristics but also have a private *type* which can be revealed by forming an edge.

By way of illustration, consider a school in which students publicly belong to distinct social groups – their school classes – and wish to make like-minded friends, i.e., friends of the same type. Due to inter-class rivalries, students use class membership to gauge how easy it will be to become friends with each other and strike up friendships based on these expectations. The actual ease with which a friendship is then maintained, however, depends on the two students' private types which are revealed once they are friends.

Agents. Let every agent $i \in N = \{1, ..., n\}$ have a public social group $s_i \in \mathcal{S} = \{\sigma_1, ..., \sigma_S\}$ and a private type $t_i \in \mathcal{T} = \{\theta_1, ..., \theta_T\}$, so that then we can define $N^{\sigma\theta} = \{i \in N : s_i = \sigma, t_i = \theta\}$, $N^{\sigma+} = \{i \in N : s_i = \sigma\}$, and $N^{+\theta} = \{i \in N : t_i = \theta\}$. Agents know their own type and the distribution of types in a given social group. They learn and memorise the private types of other agents with whom they form an edge. Let the $|N| \times |N|$ matrix M represent the memory of all agents, which we define as:

$$M_{ij} = \begin{cases} 1 & \text{if agent } i \text{ knows the type of agent } j \\ 0 & \text{otherwise} \end{cases} \qquad (4)$$

so that $M = I_{|N|}$ (where $I_{|N|}$ is the identity matrix) corresponds to agents who only know their own type and $M = J_{|N|}$ (where $J_{|N|}$ is a matrix of ones) implies complete information.

Networks. We define networks as the tuple (N, M, g), consisting of agents N with memory M represented as nodes in an undirected graph g where, as before, $ij \in g, i, j \in N$, denotes the presence of an edge from i to j. Agents can bilaterally agree to form an edge or unilaterally decide to sever an existing edge.

Actual Utility. Agents aim to maximise their *actual utility* derived from their location in graph g, defined for $i \in N$ as follows:

$$u_i(g) = \sum_{j \neq i} \delta^{d_{ij}(g)} - \sum_{j: \, ij \, \in \, g} c_{ij} \qquad (5)$$

where the benefit for i to be connected to j, $\delta \in (0, 1)$, decays with the number of edges on the shortest connecting path, $d_{ij}(g)$, where $d_{ij}(g) = \infty$ if $ij \notin g$. Further, we assume the cost of a direct connection, c_{ij}, depends on the agents' types, giving:

$$c_{ij} = c(t_i, t_j) \qquad (6)$$

Beliefs. Given incomplete information over types, we define agent i's belief that agent j is of type θ given j's group membership s_j as follows:

$$\begin{aligned} &\text{if } M_{ij} = 0 : \pi_i(\theta|s_j, M) \in [0, 1] \\ &\text{if } M_{ij} = 1 : \pi_i(\theta|s_j, M) = \mathbb{1}(t_i = t_j) \end{aligned} \qquad (7)$$

where $\mathbb{1}(\cdot)$ is an indicator function. Intuitively, beliefs are uncertain if i and j are unacquainted ($M_{ij} = 0$) but reflect certainty if they have discovered each other's types ($M_{ij} = 1$). For ease of notation, we subsequently suppress M in the belief function.

Expected Utility. Since agents at least know the types of all they currently have an edge with, only edge *addition* occurs under uncertainty. The expected utility for i from a new edge to j added to graph g given memory M can be expressed as:

$$E_i[u_i(g \cup ij)|M]$$
$$= \delta - \sum_{\theta \in \mathcal{T}} \pi_i(\theta|s_j) \, c(t_i, \theta) + \sum_{k \notin \{i,j\}} \delta^{d_{ik}(g \cup ij)} - \sum_{\substack{k:ik \in g \cup ij \\ k \neq j}} c_{ik} \quad (8)$$

where the first two terms are the benefit of an edge with j less its expected cost. This expected cost is an average of the costs to connect with each type $\theta \in \mathcal{T}$ weighted by agent i's belief that j is of type θ given social group s_j. The last two terms are the utility from the remainder of the network if the edge were added ($g \cup ij$) less the cost of the other edges i maintains.

We can use the above to define the expected *incremental change* in utility from adding an edge to agent j by subtracting the utility i has prior to the change, $u_i(g)$, from the expected utility following the change, $E_i[u_i(g \cup ij)|M]$, giving:

$$E_i[\Delta u_i(g \cup ij)|M] = E_i[u_i(g \cup ij)|M] - u_i(g) \quad (9)$$

where $E_i[\Delta u_i(g \cup ij)|M] \geq 0$ would imply agent i's consent to an edge with j. For ease of notation, we subsequently suppress M in the expectation operator.

Stability. The PSE definition can be refined now to reflect that agents do not know each other's types in advance:

$$\begin{array}{ll} \text{(i)} \; \forall ij \in g, \; u_i(g) \geq u_i(g \setminus ij) \; \text{and} \; u_j(g) \geq u_j(g \setminus ij), \; \text{and} \\ \text{(ii)} \; \forall ij \notin g, \; \text{if} \; E_i[\Delta u_i(g \cup ij)] \geq 0 \; \text{then} \; E_j[\Delta u_j(g \cup ij)] < 0 \end{array} \quad (10)$$

which implies that (i) no agent should be better off severing an existing edge and (ii) no pair of agents could weakly improve *expected* utility individually by forming a new edge. We assume condition (ii) only requires non-negative utilities to cover cases where both agents are indifferent.

Network Formation. Following Jackson and Watts [13], the network is formed by repeatedly checking for a PSE (Equation 10). In a given period, if a PSE condition is violated, then there exist agents $i, j \in N$ who can either bilaterally form an edge or unilaterally sever an edge to improve utility, moving the process into the next period. We can thus define a sequence of τ PSE-violating agent pairs $P(\tau) = \{\{i, j\}, \{k, l\}...\}$ which form the network over τ periods.

Further Assumptions. Having set up the general model, it is useful to follow the insider-outsider cost assumption of Galeotti et al. [8] or Jackson and Rogers [11] who simplify costs for tractability:

$$c(t_i, t_j) = \begin{cases} c_L & \text{if } t_i = t_j \\ c_H & \text{otherwise} \end{cases} \quad (11)$$

where c_L is the cost of an *intra-type* edge and c_H the cost of an *inter-type* edge and $0 < c_L < c_H$. Expected costs then reduce as follows:

$$\sum_{\theta \in \mathcal{T}} \pi_i(\theta|s_j) \, c(t_i, \theta) = \pi_i(t_i|s_j) \, c_L + (1 - \pi_i(t_i|s_j)) \, c_H \quad (12)$$

Complete Information. The model nests the complete information case by letting $M = J_{|N|}$. Beliefs are then updated to $\pi_i(t_i|s_j) \in \{0, 1\}$ for all $i, j \in N$ prior to the start of the formation process. Expected and actual utilities then coincide as do the PSE conditions (Equations 2 and 10).

4 THEORETICAL ANALYSIS

In this section, we provide theoretical intuition for the operation of the model. We characterise the impact of beliefs on network formation (Section 4.1), show that our model can generate all complete information network graphs under incomplete information (Section 4.2), and demonstrate that our model encompasses individually utility-maximising decisions that result in socially sub-optimal segregation in equilibrium (Section 4.3).

4.1 The Role of Beliefs

The relationship between beliefs and networks is, in general, complex when costs are entirely unrestricted. Given our aim of studying social networks, we assume that agents of the same type principally want to connect and thus focus on the low intra-type cost case.

Section 4.1.1 helps characterise the general willingness of agents to risk forming edges and discover each other's types. Section 4.1.2 provides additional such characterisation for the empty network ($g = \emptyset$) as the initial starting point. Section 4.1.3 demonstrates that segregation is possible in our model. Specifically, beliefs can support insular groups of agents which in previous works were either absent [14] or relied on complete information [11]. These sections collectively demonstrate that in many cases beliefs are sufficient for edges to initially form or not form, irrespective of potential gains in actual utility.

4.1.1 Beliefs and Type Discovery. Proposition 4.1 shows that agents will always form an edge (which may be severed later) regardless of their beliefs if inter-type costs are low. In contrast, agents need to be sufficiently optimistic about a potential connection if inter-type costs are high. Using this proposition, we state Corollary 4.2 which implies that there exist beliefs that are able to persuade agents to form an edge (which may be severed later) for arbitrarily high inter-type costs.

PROPOSITION 4.1. *Assume* $c_L \leq \delta - \delta^2$ *and suppose* $\exists i, j \in N$ *with:*

(i) $c_H > \delta - \delta^2$ *and* $min\{\pi_i(t_i|s_j), \pi_j(t_j|s_i)\} \geq \frac{c_H - (\delta - \delta^2)}{c_H - c_L}$, *or*

(ii) $c_H \leq \delta - \delta^2$ *and* $\pi_i(t_i|s_j), \pi_j(t_j|s_i) \in [0, 1]$,

then in a PSE we must have $M_{ij} = 1$.

PROOF. To prove this by contradiction, suppose there is a PSE with graph g and $M_{ij} = 0$. Without loss of generality, the least amount of utility agent i expects to gain from forming an edge with j corresponds to the case when i and j are as closely connected as possible without a direct connection, i.e., $d_{ij}(g) = 2$. The expected gain is therefore non-negative if:

$$\delta - [\pi_i(t_i|s_j) \, c_L + (1 - \pi_i(t_i|s_j)) \, c_H] - \delta^2 \geq 0$$

where the first two terms are the benefit of forming an edge less the expected cost of doing so, and where the last term is the lost

benefit of being closely but indirectly connected. This expression can be rewritten as:

$$\pi_i(t_i|s_j) \geq \frac{c_H - (\delta - \delta^2)}{c_H - c_L}$$

which means agents i and j would be compelled to form an edge, so that $M_{ij} = 1$, resulting in a contradiction. Statements 1 and 2 then follow directly. □

COROLLARY 4.2. *Let $c_L \leq \delta - \delta^2 < c_H$, then $\forall i, j \in N$ there exist $\pi_i(t_i|s_j), \pi_j(t_j|s_i) \in [0, 1]$, so that $M_{ij} = 1$ in a PSE.*

PROOF. Take any $i, j \in N$. If $M_{ij} = 1$ initially, then the statement holds trivially. Next assume $M_{ij} = 0$ initially, then Proposition 4.1 implies $M_{ij} = 1$ in a PSE for $\pi_i(t_i|s_j) \in [\frac{c_H - (\delta - \delta^2)}{c_H - c_L}, 1] \subset [0, 1]$. □

4.1.2 Beliefs and the Empty Network. Proposition 4.3 and Corollary 4.4 provide more specific characterisation for networks that emerge when agents start off in an empty network. This is particularly useful for simulations, as the proposition provides parameterisation that ensures agents will start forming edges to discover types, while the corollary provides some additional predictions about the extent to which discovery occurs. Note that they do not imply that the empty network is not a PSE, as agents may decide to sever their edges once they know each other's types.

PROPOSITION 4.3. *Consider the empty network $g = \emptyset$ with $M = I_{|N|}$ and let:*

  (i) *$c_H > \delta$ and $\exists i, j \in N, i \neq j, min\{\pi_i(t_i|s_j), \pi_j(t_j|s_i)\} \geq \frac{c_H - \delta}{c_H - c_L}$, or*

  (ii) *$c_H \leq \delta$,*

then this network is not a PSE.

PROOF. To prove this by contradiction, suppose $g = \emptyset$ with $M = I_{|N|}$ was a PSE. Take any two agents i and j, then the expected utility of forming an edge for agent i is always non-negative if $\delta - \pi_i(t_i|s_j) c_L - (1 - \pi_i(t_i|s_j)) c_H \geq 0$, or equivalently:

$$\pi_i(t_i|s_j) \geq \frac{c_H - \delta}{c_H - c_L}$$

An edge will be formed if agent j's expected utility is also non-negative, so that $M_{ij} = 1$ or $M \neq I_{|N|}$, resulting in a contradiction. The proposition thus follows by symmetry. □

COROLLARY 4.4. *Consider the empty network $g = \emptyset$ with $M = I_{|N|}$ and let $c_L \leq \delta - \delta^2 < c_H$. Then after converging to a PSE, it must be that:*

  (i) *$\forall i, j \in N, i \neq j, \pi_i(t_i|s_j) < \frac{c_H - \delta}{c_H - c_L} \rightarrow M = I_{|N|}$,*

  (ii) *$\exists i, j \in N, i \neq j, min\{\pi_i(t_i|s_j), \pi_j(t_j|s_i)\} \geq \frac{c_H - \delta}{c_H - c_L} \rightarrow M_{ij} = 1$, and*

  (iii) *$\forall i, j \in N, min\{\pi_i(t_i|s_j), \pi_j(t_j|s_i)\} \geq \frac{c_H - (\delta - \delta^2)}{c_H - c_L} \rightarrow M = J_{|N|}$.*

PROOF. To show part (i) by contradiction, assume there are $i, j \in N$ with $M_{ij} = 1$, then the expected utility for i prior to forming the edge with j must have been $\delta - \pi_i(t_i|s_j) c_L - (1 - \pi_i(t_i|s_j)) c_H \geq 0$, or equivalently, $\pi_i(t_i|s_j) \geq \frac{c_H - \delta}{c_H - c_L}$, resulting in a contradiction. Part (ii) follows from Proposition 4.3 and part (iii) from Proposition 4.1. □

4.1.3 Beliefs and Coexisting Components. Proposition 4.5 shows that the model supports PSEs in which insular groups of agents coexist. Such groups can be represented as *components* of the graph, $C(g) \subset N$, where all agents of the same component are (indirectly) connected, i.e., $\forall i, j \in C(g), d_{ij}(g) < \infty$. The proposition nests the special case of a singleton agent refusing to connect to any component. Its interpretation is intuitive. An agent i will refuse to branch out of their group, even to the best connected individual of another group, if they are sufficiently pessimistic.

PROPOSITION 4.5. *Consider $i, j \in N$ from separate components $C_i(g) \cap C_j(g) = \emptyset$ and $M_{ij} = 0$ initially. Let $c_H > \delta + (|C_j(g)| - 1) \delta^2$. In a PSE, $M_{ij} = 0$, if:*

$$\pi_i(t_i|s_j) < \frac{c_H - (\delta + (|C_j(g)| - 1) \delta^2)}{c_H - c_L}$$

PROOF. The highest net utility agent i can receive from forming an edge with j occurs when j has edges to all individuals in his component $C_j(g)$, so that i could become indirectly connected to everyone in $C_j(g)$. The expected incremental gain (Equation 9) is negative if $\delta + (|C_j(g)| - 1) \delta^2 - \pi_i(t_i|s_j) c_L - (1 - \pi_i(t_i|s_j)) c_H < 0$ which rearranged gives the proposition.

□

4.2 Generating Complete Information Network Graphs

In Section 3.1, we noted that our model directly nests the complete information case by assuming $M = J_{|N|}$ (effectively resulting in the model of Jackson and Rogers [11]). In this section, we draw a connection to a proposition due to Song and van der Schaar [23] which can be used to show that the network graphs generated by our model with incomplete information ($M \neq J_{|N|}$ suffices) represents a superset of those generated with complete information. This suggests that the introduction of uncertainty does not come at the cost of reduced variety in the set of possible networks graphs.

First, denote $G^C(N)$ as the set of network graphs that form from the empty network with agents under complete information while $G_S^C(N)$ denotes the set of such graphs that are also PSE. Let $G^{IC}(N)$ and $G_S^{IC}(N)$ be the incomplete information equivalents. We can then state:[2]

PROPOSITION 4.6 (SONG AND VAN DER SCHAAR [23]). *If for all $i, j \in N$ with $M_{ij} = 0$, $E_i[\Delta u_i(g \cup ij)|M] \geq 0$, then $G^C(N) \subset G^{IC}(N)$ and $G_S^C(N) \subset G_S^{IC}(N)$.*

which implies that if agents are sufficiently optimistic – such that the expected incremental utility is always non-negative – then the set of all complete information network graphs and the set of complete information PSE network graphs are, respectively, subsets of those under incomplete information.

4.3 Application to Segregation in Social Networks

Proposition 4.7 shows that a socially sub-optimal equilibrium with segregated social groups is possible in our model. This stands in

---

[2]For further details and a sketch of their proof, see supplemental materials.

contrast to previous works on heterogeneous agent models [7, 11] which concluded that such equilibria were optimal for social welfare. Section 4.2 showed that these equilibria continue to be possible in our model, so that the introduction of incomplete information provides a more flexible way of modelling segregation.

PROPOSITION 4.7. *Assume the network is initialised as $g = \emptyset$, $M = I_{|N|}$ and let all agents be of the same private type but belong to S separate social groups, $N^1, N^2, ..., N^S \subset N$ and let $c_L \leq \delta - \delta^2$. Assume that beliefs are such that:*

(i) for $i, j \in N$ with $s_i = s_j$, $\pi_i(t_i|s_j) \geq \frac{c_H - (\delta - \delta^2)}{c_H - c_L}$, and

(ii) for $i \in N^{s_i}, j \in N^{s_j}, s_i \neq s_j$, $\pi_i(t_i|s_j) < \frac{c_H - (\delta + (|N^{s_j}| - 1) \, \delta^2)}{c_H - c_L}$,

then social groups will become perfectly segregated in a PSE that is not socially optimal.

PROOF. Start with the empty network and take any agent $i \in N$. Given condition (i), Proposition 4.1 implies that $M_{ij} = 1$ for any $j \in N^{s_i}$. Since all agents have the same private type and $c_L \leq \delta - \delta^2$, all edges that are formed are also kept. By symmetry, this applies equally to $k \in N^{s_k}, s_k \neq s_i$. Given condition (ii), Proposition 4.5 implies i refuses an edge with k no matter how well-connected k is within N^{s_k}. N^{s_i} and N^{s_k} are therefore segregated.

Given $c_L \leq \delta - \delta^2$, the efficient network must be fully connected amongst agents of the same type, as the incremental utility of forming an edge will be at least $\delta - \delta^2 - c_L \geq 0$. Yet we have just shown that N^{s_i} and N^{s_k} are segregated, resulting in a socially sub-optimal outcome, giving the proposition. □

5 SIMULATION SETUP

We numerically investigate the role of beliefs in generating segregative PSE networks under incomplete information. Section 5.1 introduces setup specifics for the network and its components, while Section 5.2 details the metrics that we are measuring.[3]

5.1 Network Setup

Beliefs. We instantiate beliefs (Equation 7). As before, beliefs are set to $\pi_i(t_i|s_j) = \mathbb{1}(t_i = t_j)$ when $M_{ij} = 1$. For $M_{ij} = 0$, we let agents judge each other's types based on social group membership. Specifically, agents have base beliefs anchored by rational expectations but are optimistically biased when judging people of their own social group and pessimistically biased for any other social group. Formally expressed for $i, j \in N$ with $M_{ij} = 0$, we then have:

$$\pi_i(t_i|s_j) = \begin{cases} \frac{|N^{s_j t_j}|}{|N^{s_j+}|} + \left(1 - \frac{|N^{s_j t_j}|}{|N^{s_j+}|}\right) \gamma_i & \text{if } s_i = s_j \\ \frac{|N^{s_j t_j}|}{|N^{s_j+}|} - \frac{|N^{s_j t_j}|}{|N^{s_j+}|} \gamma_i & \text{otherwise} \end{cases} \quad (13)$$

with $N^{s_j t_j} = \{i \in N : s_i = s_j, t_i = t_j\}$ and $N^{s_j+} = \{i \in N : s_i = s_j\}$. The ratio of their cardinalities is the true probability that an agent belongs to type t_j in group s_j and therefore equals rational expectations. The bias enters via $\gamma_i \in [0, 1]$ which moves beliefs upward for agents of the same group (optimism) and downward otherwise (pessimism). The adjustment is scaled to ensure beliefs remain in the unit interval. We let γ_i be drawn from a positively

---

skewed beta distribution $\gamma_i \sim Beta(\alpha, \beta)$, with $\alpha = 1, \beta > 1$, where the shape parameter β controls the degree of bias. In the base case parameterisation we choose $\beta = 7$ (Figure 1 for illustration) which implies that c.52% of agents do not deviate more than 10% and c.79% do not deviate more than 20% from rational beliefs.
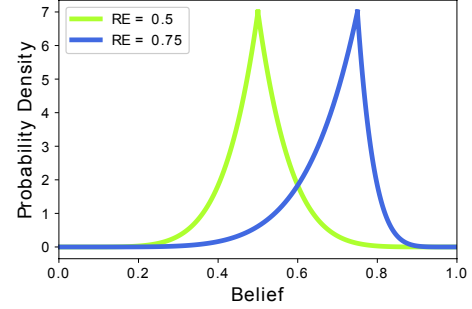

Figure 1: Optimistic beliefs are drawn from the right-hand side of the distribution peak, pessimistic beliefs from its left-hand side. Belief distribution shifts right if the rational expectations (RE) anchor increases.

Initial Conditions. We initialise with the empty network ($g = \emptyset$) and consider formation: (i) under incomplete information with biased beliefs ($M = I_{|N|}, \gamma_i \sim Beta(1, \beta)$), (ii) under incomplete information with rational expectations ($M = I_{|N|}, \gamma_i = 0$), and (iii) under complete information ($M = J_{|N|}$).

Network Formation Algorithm. We implement Algorithm 1 based on the network formation process described in Section 3.1. In a given period, we randomly select agent pairs until one violates a PSE condition (process moves into next period) or until no pairs remain (PSE reached). The same random seed is used between the three cases to ensure consistency in the random selection of agent pairs. If the process exceeds 5000 periods, we terminate for non-convergence.

Algorithm 1 Network formation algorithm.

1: **procedure** NETWORKFORMATION(N,M)
2:     Initialise $g = \emptyset$ and agent beliefs given their memory M
3:     **for** each period $\tau = 1, ..., 5000$ **do**
4:         $allPairs$ = (randomised list of all agent pairs)
5:         **for** each pair $\{i, j\} \in allPairs$ **do**
6:             **if** the pair has an edge $ij \in g$ **then**
7:                 Calculate $u_i(g), u_i(g \setminus ij), u_j(g), u_j(g \setminus ij)$
8:                 **if** $\exists k \in \{i, j\}, u_k(g) < u_k(g \setminus ij)$ **then**
9:                     Set $g = g \setminus ij$ and go to line 3
10:             **else**
11:                 Calculate $E_i[\Delta u_i(g \cup ij)], E_j[\Delta u_j(g \cup ij)]$
12:                 **if** $E_i[\Delta u_i(g \cup ij)], E_j[\Delta u_j(g \cup ij)] \geq 0$ **then**
13:                     Set $g = g \cup ij, M_{ij} = 1$ and go to line 3
14:     **return** g (converged)
15:     **return** g (not converged)

Parameterisation. We choose a network size for which convergence can be reasonably expected in a real-life setting, considering that agents need to have interacted a potentially large number of times. We, further, focus on the network dynamics by varying one cost at a time to probe their independent effects. In addition, we hold the decay factor, δ, constant in all cases, as results mainly depend on the relative magnitude of benefits and costs. Table 1 provides an overview of the shared parameters of all simulations that follow. Specimen networks are illustrated in Figure 2.

Table 1: Simulation parameters.

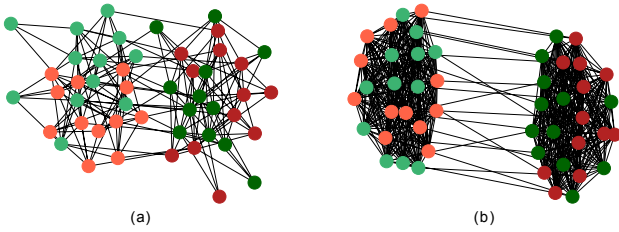| Parameter	Value
Number of agents $|N|$	48
Decay factor δ	0.7
Intra-type cost c_L (base case)	0.2
Inter-type cost c_H (base case)	1.0

Figure 2: Specimen networks with (a) biased beliefs and (b) complete information. Colours (green vs red) indicate social groups, tones (light vs dark) denote private types.

5.2 Metrics

Segregation Indices. We make use of the generalised Freeman's segregation index [3] which is given by:

$$S_F = 1 - \frac{p|N|(|N|-1)}{(\Sigma_\theta |N^{+\theta}|)^2 - \Sigma_\theta |N^{+\theta}|^2} \qquad (14)$$

where $N^{+\theta} = \{i \in N : t_i = \theta\}$. Further, p denotes the proportion of inter-group edges, defined as:

$$p = \frac{\Sigma_{g,h;g \neq h} m_{gh1}}{m_{++1}} \qquad (15)$$

where $m_{gh1} = |\{ij \in g | s_i = \sigma_g; s_j = \sigma_h\}|$ and $m_{++1} = \Sigma_g \Sigma_h m_{gh1}$ are, respectively, the number of edges between the social groups, σ_g and σ_h, and the number of all edges. A segregation index closer to 1 indicates more segregation, a value closer to 0 indicates less.

Given the random graph benchmark used in Freeman's segregation index, it can be interpreted as a measure of the segregation introduced by agent preferences. Letting the network form under incomplete information, however, introduces an additional uncertainty effect, which would be confounded in Freeman's segregation

index. To disentangle the effect of biased beliefs from pure uncertainty effects or agent preferences we propose an additional *incremental segregation index* defined as follows:

$$S_{IS} = 1 - \frac{p_{biased}}{p_{base}} \qquad (16)$$

where p_{biased} denotes the proportion of inter-group edges (Equation 15) under our network with biased beliefs and where p_{base} represents the proportion of such edges in a baseline network. We use one baseline network under complete information to see if segregation is introduced in excess of what agent preferences imply. A second baseline network under incomplete information but with rational expectations allows us to gauge if segregation is introduced in excess of what agent preferences and pure uncertainty imply.

Degree Centrality. In order to measure how well-connected the graph is, we make use of degree centrality, which is the number of edges for a given agent. To focus on the macro-characteristics of the network as a whole, we report average degree centrality across all agents.

Discovery. A distinguishing feature of our model lies in the memory matrix M (Equation 4) capturing varying degrees of information present in the system. As agents become less likely to form edges when the expected costs of doing so increase, the amount of information discovered will suffer as a result. We therefore report on the proportion of discovery, defined as:

$$discovery(N, M) = \frac{\Sigma_i \Sigma_j M_{ij}}{|N|^2} \qquad (17)$$

which corresponds to the sum of the number of agents each person knows, normalised by the maximum possible such sum given a population of $|N|$ agents.

6 SIMULATION RESULTS

This section presents the results of our experiments. Averaged results over 30 converged networks are reported throughout. Section 6.1 investigates the effect of biased beliefs, pointing to their key role in generating segregation under incomplete information. Section 6.2 provides additional intuition around the effect of the type distribution on segregation dynamics while Section 6.3, in contrast, shows that social group distributions do not significantly determine segregation dynamics.[4]

6.1 Segregation under Uncertainty

We first focus on networks with two equally sized social groups which are each equally divided between two private types (Table 2). In Figure 3 we compare the network dynamics under biased beliefs with those under rational expectations and complete information.

Segregation. The segregation indices (SIs) in the top panels (Figure 3) confirm that biased beliefs play an important role in segregation under uncertainty: Freeman's SI remains distinctly positive as intra-type and inter-type costs change. Moreover, as both the incremental SIs with respect to rational expectations and complete information are positive and closely track the changes in Freeman's

---

[4]Full network dynamics for Figures 4 to 6 are provided in the supplementary materials.

Table 2: Population composition. Array entries in the first row indicate size of a particular social group. Array entries in the second row indicate size of a particular private type given a social group.

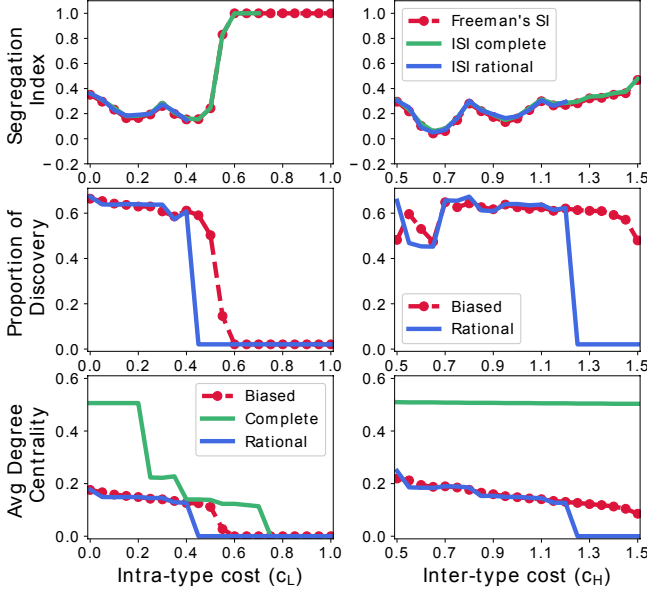| Parameter	Value
Social group size	[24, 24]
Private type distribution	[12, 12] per group

Figure 3: Biased beliefs sustain segregation. Incremental SI (ISI) cannot be computed when baseline consists of singletons (as $p_{base} = 0$). Freeman's SI and ISI approach 1 when the biased network collapses to singletons (as $p_{biased} = 0$).

SI, this suggests that the segregation is not a result of pure uncertainty or agent preferences respectively.

Discovery. The middle panels (Figure 3) show how willingly agents form edges and discover each other's types under biased beliefs and rational expectations (under complete information all types are discovered from the outset). This does not imply that edges that are initially formed survive once types are revealed. Biased beliefs sustain edge-forming between agents for both higher intra-type and inter-type costs compared to the case under rational expectations.

Intuitively, biased beliefs enable agents to hold different beliefs, such as those who are optimistically biased. These agents have lower expected costs and are thus more willing to risk forming edges. In comparison, agents with rational expectations all follow the same beliefs. As such, biased beliefs induce edges in higher cost ranges relative to rational expectations.

Connectedness. The lower panels (Figure 3) show how well connected the network is, i.e., how many edges survive. Firstly, networks under biased beliefs and rational expectations have a similar

degree of connectedness for the lower part of the cost ranges. Given the evolution of the SIs, this suggests that biased networks support as many edges as networks with rational expectations, but with a greater number of edges between agents of the same social group.

Secondly, the greater willingness to connect at higher costs under biased beliefs compared to under rational expectations also results in a correspondingly later collapse of the biased network.

Lastly, networks under complete information are better connected in equilibrium. This is intuitive, as unconnected agents of the same type will always form and maintain an edge as long as the intra-type cost lies below the decay factor – a decision that is independent of inter-type costs as indicated by the flat degree centrality curve on the right-hand side of the panel.

Degree of Bias. In Figure 4 we vary the degree of bias by changing the β shape parameter, where a lower β-value corresponds to a greater degree of bias. We observe that while increasing the bias is clearly associated with an upward shift in Freeman's SI (where the network has not collapsed), a reduction in the bias still yields distinctly segregative results, suggesting that even mildly biased beliefs are strong drivers of segregation.
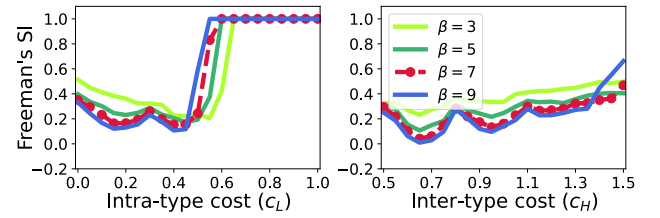

Figure 4: An increase in bias (a reduction in the β shape parameter) leads to greater segregation. Freeman's SI approaches 1 when the network collapses to singletons.

6.2 Compositional Impact: Private Types

We compare segregation dynamics under different distributions of private types with (i) a higher number of types ("4 Types"), (ii) a type that dominates across groups ("Dominant"), and (iii) types that each dominate a different group ("Correlated"). A population breakdown is shown in Table 3.

Table 3: Population composition. Array entries in the first row indicate size of a particular social group. Array entries in the second row indicate size of a particular private type given a social group.

Parameter	Base	4 Types	Dominant	Correlated
Social group size	[24, 24]	[24, 24]	[24, 24]	[24, 24]
Private type distribution	[12, 12] per group	[6, 6, 6, 6] per group	[18, 6] per group	Group 1: [18, 6] Group 2: [6, 18]

Figure 5 shows that segregation dynamics are sensitive to the distribution of private types. This is expected given their direct relationship to agent preferences.
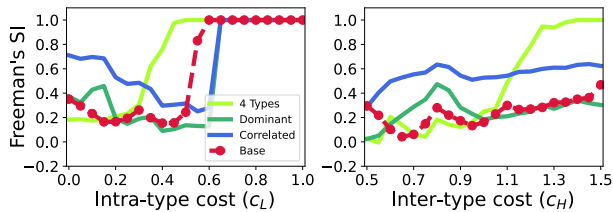
Figure 5: Correlating types with social groups raises segregation. Increasing the number of types reduces edge-forming incentives. Freeman's SI approaches 1 when the network collapses to singletons.

Figure 6: Social group compositions do not strongly influence segregative dynamics. Freeman's SI approaches 1 when the network collapses to singletons.

The results indicate that increasing the number of types (the "4 Types" case) appears to have a disincentivising effect on network formation. Intuitively, a larger number of rival types dilutes incentives to form edges, as all edges with agents of a different type are penalised. This is reflected in a rapid rise in Freeman's SI as costs increase and an earlier collapse of the network to singletons.

Initialising the network with a type dominating each social group (the "Dominant" case) does not impact dynamics significantly: networks tend to be slightly better connected and collapse to singletons at slightly higher intra-type costs compared to the base case.

In contrast, in the "Correlated" case in which each social group is dominated by a different private type, there is a marked increase in Freeman's SI. This is not surprising, as this case naturally encourages segregation given that social group memberships are informative about types. Agents of the majority type in a given group would sensibly stay in their social group.

6.3 Compositional Impact: Social Groups

We compare segregation dynamics under different social group compositions with (i) multiple smaller social groups ("4 Groups") and (ii) a dominant group ("Dominant"). An overview of the populations is shown in Table 4.

Table 4: Population composition. Entries of arrays indicate size of a particular social group.

Parameter	Base	4 Groups	Dominant
Social group size	[24, 24]	[12, 12, 12, 12]	[12, 36]
Private type distribution		Even split within each group	

Figure 6 shows that changes to the social groups do not significantly alter base case dynamics. Note that Freeman's SI is calculated with respect to the number of inter-group edges expected in a randomly formed network. An increase in social groups or creating a dominant group implicitly change the expected number of random inter-group edges, so that the similarity in SIs suggests that agents must have acted to adjust the number of inter-group edges they form to keep proportions constant.
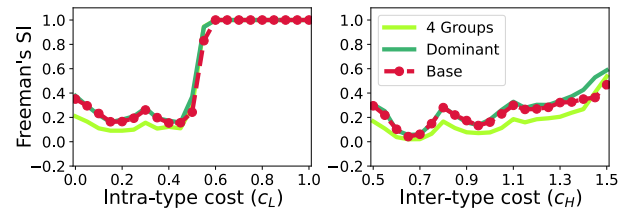
7 CONCLUSION

We have contributed an agent-based network formation model for the study of group segregation relaxing the common assumption of complete information. We have done so by introducing notions of expectations and beliefs that are allowed to be heterogeneous across agents. The resulting model can be parameterised to nest a complete information model directly and is also capable of generating all complete information network graphs under incomplete information, thus providing greater modelling flexibility. Beyond this, we have demonstrated that socially sub-optimal segregation equilibria can be supported when agents are sufficiently biased. We have proceeded to a simulation-based approach to investigate the dynamics of segregation and found clear evidence pointing to the role of biased beliefs in generating segregation under incomplete information.

This has not been an exhaustive study of beliefs. The belief mechanisms we have assumed throughout have been relatively static to enable our focus on the presence and absence of complete information. A promising avenue of further research would be to imbue our mechanisms with greater realism, such as by assuming that beliefs update gradually, such as in Zhang and van der Schaar [27], or by assuming that agents have imperfect memory and forget each other over time, or by including a budget constraint on the number of edges an agent can form. A second line of extensions could emanate from incorporating more complex relationships between social groups and beliefs, where a spatial cost topology, such as in Johnson and Gilles [16], could be applied to model expected costs between ordinal costs, or additional heterogeneity in benefits could be introduced, such as in Galeotti et al. [8]. This could be useful in moving the focus from overall segregative patterns as in this paper to studying segregation between specific groups. A third topic of future research could revolve around the stability conditions underlying our model. While pairwise stability offers some attractive properties, it would be worthwhile investigating the effect of coalitional stability concepts, such as in Jackson and van den Nouweland [12].

REFERENCES

[1] Dominic Aits, Alexander Carver, and Paolo Turrini. 2019. Group segregation in social networks. *Proceedings of the International Joint Conference on Autonomous Agents and Multiagent Systems, AAMAS* 3, Aamas (2019), 1524–1532.

[2] Joshua Evan Blumenstock, Guanghua Chi, and Xu Tan. 2019. Migration and the value of social networks. *CEPR Discussion Paper* (2019), No. DP13611. https://ssrn.com/abstract=3360078

[3] Michał Bojanowski and Rense Corten. 2014. Measuring segregation in social networks. *Social Networks* 39, 1 (2014), 14–32. https://doi.org/10.1016/j.socnet.2014.04.001

[4] Kenneth Chay and Kaivan Munshi. 2013. Black Networks After Emancipation: Evidence from Reconstruction and the Great Migration. *Working Paper* (2013).

[5] Gongmeng Chen, Kenneth A. Kim, John R. Nofsinger, and Oliver M. Rui. 2007. Degrees of Uncertainty: An Overview and Framework for Future Research on Experience-Based Choice. *Journal of Behavioral Decision Making* 20 (2007), 425–451. https://doi.org/10.1002/bdm

[6] Joan de Martí and Yves Zenou. 2015. Network games with incomplete information. *Journal of Mathematical Economics* 61 (2015), 221–240. https://doi.org/10.1016/j.jmateco.2015.10.002

[7] Joan de Martí and Yves Zenou. 2017. Segregation in Friendship Networks. *Scandinavian Journal of Economics* 119, 3 (2017), 656–708. https://doi.org/10.1111/sjoe.12178

[8] Andrea Galeotti, Sanjeev Goyal, and Jurjen Kamphorst. 2006. Network formation with heterogeneous players. *Games and Economic Behavior* 54, 2 (2006), 353–372. https://doi.org/10.1016/j.geb.2005.02.003

[9] Mark Granovetter. 1983. The Strength of Weak Ties: A Network Theory Revisited. *Sociological Theory* 1 (1983), 201–233. https://doi.org/10.1007/978-3-658-21742-6{_}55

[10] David M. Grether. 1980. Bayes Rule as a Descriptive Model: The Representativeness Heuristic. *The Quarterly Journal of Economics* 95, 3 (1980), 537. https://doi.org/10.2307/1885092

[11] Matthew O. Jackson and Brian W. Rogers. 2005. The Economics of Small Worlds. *The Economic Journal* 3, 2-3 (2005), 617–627. https://doi.org/10.2307/2223611

[12] Matthew O. Jackson and Anne van den Nouweland. 2005. Strongly stable networks. *Games and Economic Behavior* 51, 2 SPEC. ISS. (2005), 420–444. https://doi.org/10.1016/j.geb.2004.08.004

[13] Matthew O. Jackson and Alison Watts. 2002. The evolution of social and economic networks. *Journal of Economic Theory* 106, 2 (2002), 265–295. https://doi.org/10.1006/jeth.2001.2903

[14] Matthew O. Jackson and Asher Wolinsky. 1996. A strategic model of social and economic networks. *Journal of Economic Theory* 71, 1 (1996), 44–74. https://doi.org/10.1006/jeth.1996.0108

[15] Matthew O. Jackson and Leeat Yariv. 2007. Diffusion of behavior and equilibrium properties in network games. *American Economic Review* 97, 2 (2007), 92–98. https://doi.org/10.1257/aer.97.2.92

[16] Cathleen Johnson and Robert P. Gilles. 2000. Spatial social networks. *Review of Economic Design* 5, 3 (2000), 273–299. https://doi.org/10.1007/PL00013690

[17] Daniel Kahneman and Amos Tversky. 1973. On the psychology of prediction. *Psychological Review* 80, 4 (1973), 237–251. https://doi.org/10.1037/h0034747

[18] Kaivan Munshi. 2011. Strength in numbers: Networks as a solution to occupational traps. *Review of Economic Studies* 78, 3 (2011), 1069–1101. https://doi.org/10.1093/restud/rdq029

[19] Kaivan Munshi and Mark Rosenzweig. 2016. Networks and Misallocation: Insurance, Migration, and the Rural-Urban Wage Gap. *American Economic Review* 106, 1 (2016), 46–98. https://doi.org/10.1257/aer.20131365

[20] Thomas C. Schelling. 1969. Models of Segregation. *The American Economic Review* 59, 2 (1969), 488–493.

[21] Thomas C. Schelling. 1971. Dynamic Models of Segregation. *Journal of Mathematical Sociology* 1 (1971), 143–186.

[22] Jan Skopek, Florian Schulz, and Hans Peter Blossfeld. 2011. Who contacts whom? Educational homophily in online mate selection. *European Sociological Review* 27, 2 (2011), 180–195. https://doi.org/10.1093/esr/jcp068

[23] Yangbo Song and Mihaela van der Schaar. 2015. Dynamic network formation with incomplete information. *Economic Theory* 59, 2 (2015), 301–331. https://doi.org/10.1007/s00199-015-0858-y

[24] Amos Tversky and Daniel Kahneman. 1974. Judgment under Uncertainty : Heuristics and Biases. *Science* 185 (1974), 1124–1131.

[25] Thierry Verdier and Yves Zenou. 2017. The role of social networks in cultural assimilation. *Journal of Urban Economics* 97 (2017), 15–39. https://doi.org/10.1016/j.jue.2016.11.004

[26] Yves Zenou. 2015. A dynamic model of weak and strong ties in the labor market. *Journal of Labor Economics* 33, 4 (2015), 891–932. https://doi.org/10.1086/681098

[27] Simpson Zhang and Mihaela van der Schaar. 2017. From Acquaintances to Friends: Homophily and Learning in Networks. *IEEE Journal on Selected Areas in Communications* 35, 3 (2017), 680–690. https://doi.org/10.1109/JSAC.2017.2672238