Social Integration in a Diverse Society

Flache, Andreas

Published in:
New Perspectives and Challenges in Econophysics and Sociophysics

DOI:
10.1007/978-3-030-11364-3_15

Chapter 15
Social Integration in a Diverse Society: Social Complexity Models of the Link Between Segregation and Opinion Polarization

Andreas Flache

Abstract There is increasing societal and scholarly interest in understanding how social integration can be maintained in a diverse society. This paper takes a model of the relation between opinion polarization and ethnic segregation as an example for social complexity. Many argue that segregation between different groups in society fosters opinion polarization. Earlier modeling work has supported this theoretically. Here, a simple model is presented that generates the opposite prediction based on the assumption that influence can be assimilative or repulsive, depending on the discrepancy between interacting individuals. It is discussed that these opposite results from similar models point to the need for more empirical research into micro-level assumptions and the micro-to-macro transformation in models of opinion dynamics in a diverse society.

Introduction

Migration both within and between countries has strongly increased in recent decades [11]. For many Western societies this comes with more ethnic and cultural diversity of their population. Other societies, like India, know high levels of diversity already for many centuries. Ethnic and cultural diversity have many benefits, for example in terms of a broader pool of talent or more creativity in diverse teams in organizations [16]. But diversity also constitutes a challenge for societies. Often diversity is associated with high levels of segregation between different groups [9], or with differences between groups in attitudes on fundamental issues such as civil rights of homosexuals, legalization of abortion, or gender equality [38].

There are no easy answers to the question under which conditions diversity can endanger societal integration and foster instead persistent disagreement or even polarization. Polarization can be described as the tendency of a population to fall apart

A. Flache (✉)
Department of Sociology/ICS, University of Groningen, Groningen, The Netherlands
e-mail: a.flache@rug.nl
URL: https://flache.gmw.rug.nl/

F. Abergel et al. (eds.), *New Perspectives and Challenges in Econophysics and Sociophysics*, New Economic Windows,
https://doi.org/10.1007/978-3-030-11364-3_15

into a small number of subgroups with large agreement within and disagreement between them [6]. Whether and under which conditions polarization arises is notoriously hard to predict. The evolution of the distribution of opinions in a society results from numerous simultaneous interactions between individuals both within their own cultural subgroups as well as across intergroup boundaries. Some of these interactions may drive groups apart, others may foster consensus. For example, research in the tradition of contact theory emphasizes that intergroup contacts reduce prejudice and promote agreement [1, 14, 40], but negative interactions at the individual level can also result in deeper divisions between groups [43, 44]. Polarization can thus be an outcome that results from the interactions of multiple individuals who neither expect nor intend to bring it about. One reason is the possibility that small changes in opinion distributions can have large unexpected consequences, for instance when disagreement emerging between some individuals in a local region of a network spreads and then quickly becomes amplified by social contagion [7, 20]. Identifying the conditions and mechanisms under which social influence dynamics in a diverse population result in polarization is therefore a major scientific issue with a long tradition of vivid debate [34].

In search for tools to tackle the inherent complexity of collective opinion dynamics, researchers used in recent decades increasingly agent-based computational modeling [5, 22, 28]. While this has greatly helped to understand the complex interplay of individual-level social interactions with macro-level outcomes, it also highlighted that the outcomes of opinion dynamics can sensitively depend on the exact assumptions researchers make about the process of social influence at the micro-level. In this paper, I will illustrate this with a model of the relationship between segregation and polarization in a diverse society. According to many, segregation between different subgroups is one of the important reasons for persistent disagreement between groups. Segregation is the separation of different groups, for example between residential areas of a city [9], or between different clusters in a social network [12, 36]. Segregation can reduce the extent of intergroup contact [8] and thus exacerbate prejudice. A further problem is that segregation can create 'bubbles' within which only like-minded people meet and interact. As former U.S. president Obama pointed out in his farewell address, in such a bubble we are "surrounded by people who look like us and share the same political outlook and never challenge our assumptions", such that "we become so secure ... that we start accepting only information, whether it's true or not, that fits our opinions" [39].

The argument that segregation fosters polarization seems compelling, but social complexity models showed how different equally plausible micro-level theories of social influence can generate radically different implications. A number of formal models is consistent with Obama's intuition. Building on persuasive argument theory [37, 46], models proposed by Mäs and coauthors [31, 33] assume that agents with more similar opinions are more likely to persuade each other to strengthen their already prevailing opinion tendency. Simulations demonstrated how then opinions in different subgroups can be pushed towards opposing poles of an opinion spectrum if agents prefer interacting with similar others, based on the principle of homophily [35]. Similar dynamics have also been derived from models of "biased assimilation"

[4, 13] in which agents are assumed to put more weight on those influences in the process of assimilation that are in line with their current opinion.

Models of persuasive arguments and biased assimilation suggest that segregation fosters opinion polarization. Building on earlier work [17], I will show in this paper that radically different conclusions can be drawn from another class of models. I follow a number of studies [2, 3, 15, 18, 20, 26, 29, 30, 41] which incorporated into models of social influence the assumption that influence can not only be assimilative, reducing opinion differences, but also repulsive. When influence is repulsive, individuals strive to be dissimilar to people they dislike, accentuating disagreement with others. But this only happens when those others are perceived as being too discrepant, otherwise influence is assimilative. This combination of assimilative and repulsive influence is suggested by theories of fundamental psychological processes in the formation of attitudes, like Heider's balance theory [23] or Festinger's theory of cognitive dissonance [19]. In a number of formal models elaborating this idea, it has in particular been assumed that perceived discrepancy not only arises from disagreement in opinions between individuals, but also from 'demographic' differences representing fixed characteristics like gender, race or ethnicity [17, 21, 30].

I will demonstrate in what follows that a model combining assimilative and repulsive influence implies that more segregation reduces opinion polarization between groups. The model will be presented in section "Modelling the Link Between Segregation and Opinion Polarization", results are described in section "Results". Section "Discussion and Conclusion" concludes with a more general reflection on the role of social complexity models for our understanding of social integration in a diverse society.

Modelling the Link Between Segregation and Opinion Polarization

First, the micro-level assumptions about social influence are introduced in section "Microlevel Assumptions About Social Influence". Second, the model of spatial network segregation is described in section "Modeling the Spatial Structure: Local Interaction and Segregation".

Microlevel Assumptions About Social Influence

The model contains a population of N individuals i who are throughout members of either group 0 or group 1, indicated by group membership $g_i \in \{0, 1\}$. For simplicity, I assume that both subgroups are always equally large. Every individual i adopts at every time point t an opinion o_{it}, with $0 \leq o_{it} \leq 1$. Following [17, 20], individuals are connected in a static interaction network (see section "Modeling the Spatial Structure:

Local Interaction and Segregation" for details) and can only interact with network neighbors.

Dynamics of the model unfold in consecutive discrete time steps t. In every time step, a pair of two different network neighbors i and j is selected at random with equal probability. All individuals k who are not involved in an interaction at time point t do not change their opinions, thus $o_{k,t+1} = o_{kt}$. If i and j interact, then both can modify their current opinions to move closer towards or away from the opinion of the interaction partner as given by Eqs. (15.1) and (15.2).

$$o_{i,t+1} = o_{it} + \Delta o_{it} = o_{it} + \mu w_{ijt} \left(o_{jt} - o_{it}\right) \tag{15.1}$$

$$o_{j,t+1} = o_{jt} + \Delta o_{jt} = o_{jt} + \mu w_{jit} \left(o_{it} - o_{jt}\right) \tag{15.2}$$

The parameter μ ($0 < \mu \leq 0.5$) in Eqs. (15.1) and (15.2) defines the rate of opinion change and will be kept at $\mu = 0.5$ in the present paper. The influence weights w_{ijt} and w_{jit} in Eqs. (15.1) and (15.2) represent the direction and magnitude of the influence of i on j and j on i, respectively. Weights are constrained by $-1 \leq w_{ij} \leq 1$. A positive weight w_{km} entails assimilative influence (k moving her opinion closer towards m's opinion), whereas a negative weight imposes repulsive influence (k moving her opinion away from m's opinion). A zero weight implies no change, reflecting indifference towards the source of influence. In this basic form, Eqs. (15.1) and (15.2) allow interactions to push the opinion outside of the opinion interval $[0, 1]$ if weights are negative. In this case, the resulting opinion is truncated to the interval boundary that was crossed by the opinion shift. In some models that combine assimilation and repulsive influence, opinions are constrained with smoother functions [20, 21, 25], but this seems to have little effect on the main model dynamics.

The link between diversity, disagreement and social influence is implemented as follows. The influence weight w_{ijt} expresses the similarity that i experiences at time point t between herself and j. More precisely, the influence weight declines in the current level of disagreement $|o_{jt} - o_{it}|$, and is reduced if i and j belong to different groups. Equation (15.3) formalizes the computation of influence weights.

$$w_{ijt} = 1 - 2\left(\beta_O |o_{jt} - o_{it}| + \beta_D |g_j - g_i|\right). \tag{15.3}$$

Equation (15.3) shows that influence becomes repulsive when the discrepancy $\beta_O |o_{jt} - o_{it}| + \beta_D |g_j - g_i|$ exceeds 0.5, half of the theoretical maximum of 1. The parameters β_O and β_D in Eq. (15.3) scale the relative impact that respectively, *opinion* disagreement and *demographic* differences have on discrepancy. For convenience, I impose the constraint $\beta_O + \beta_D = 1$.

The model assumed here uses a simple linear transformation of discrepancy into influence weights w_{ijt}. Some studies have adopted a non-linear weight function in an otherwise similar framework [26, 32], but did only consider disagreement in opinions. Future work should combine a non-linear weight function with both disagreement and intergroup differences to explore possible new implications.

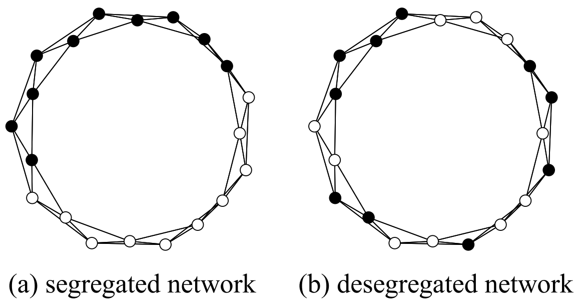


Fig. 15.1 Example for result of desegregation algorithm with desegregation rate $d = 0.5$ (right), starting from initially maximally segregated ring lattice (left). $N = 20$, range of interaction $= 2$

Modeling the Spatial Structure: Local Interaction and Segregation

The key condition of interest in the simulation experiments is segregation between groups. Modeling local interaction in a simple way, I employ a ring-lattice network in which all agents have the same number of local network neighbors to the left and to the right, called range of interaction r. Figure 15.1a shows the baseline condition of maximal segregation between the two groups for $r = 2$, $N = 20$, and two equally large subgroups. In both subgroups, only 4 out of 10 agents have any outgroup-neighbor among their 4 network neighbors. Of those 4 agents, half have 2 outgroup neighbors and the other half has only 1.

The degree of segregation is manipulated as follows.[1] Starting from a maximally segregated network (see Fig. 15.1a), a subset of N_s distinct agents from group 0 is randomly chosen for relocation. N_s is given by the desegregation rate d, $(0 \leq d \leq 0.5)$, rounded to the integer nearest to $d\ {}^{N}\!/_{2}$. For every chosen agent of group 0, a unique randomly selected agent from group 1 is picked. In all the selected pairs thus formed, network positions of the group 0 agents are swapped with those of the group 1 agents. Figure 15.1b shows an example for a network generated with a desegregation rate of $d = 0.5$.

To quantify segregation, a segregation measure S is computed. S indicates the fraction of same-group neighbors among all network neighbors of an agent, averaged over all agents and divided by the theoretically possible maximal fraction of ingroup neighbors, given N and r. The exact value of S given d varies randomly, depending on which pairs of agents were selected for a position swap.

The relation between desegregation rate d and average segregation S is non-linear. The closer the desegregation rate comes to 0.5, the less impact further increase has on the segregation level S. To account for this, the average value of S per level of d will be used to show how segregation affects polarization in the experiments that follow, whereas segregation will be manipulated with stepwise variation in the desegregation rate d.

[1] All computations, simulations and graphics in this paper were produced with Wolfram Mathematica©Version 11.2.

Results

In Section "Design and Measures", design and outcome measures of the computational experiments are described. Two experiments are conducted in both of which segregation S is manipulated. In the first experiment it is assumed that there is no initial group-specific disagreement in opinions. In the second experiment, a moderate group-specific disagreement is introduced. Results of experiment 1 and experiment 2 are described in sections "Experiment 1" and "Experiment 2", respectively. Section "Robustness Tests" is devoted to a brief description of some robustness tests.

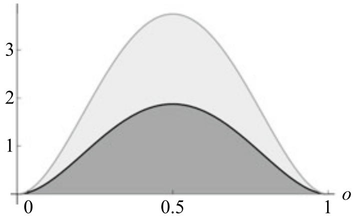
Design and Measures

The following simple baseline scenario was used in both experiments. Population size was set to $N = 100$ with 50 members in groups 0 and 1, respectively. The network was a circular ring lattice with interaction range $r = 5$. The relative impact of demographic dissimilarity on the influence weight w_{ijt} was set to $\beta_D = 1/3$. With this value, polarization between groups is possible but not trivial.

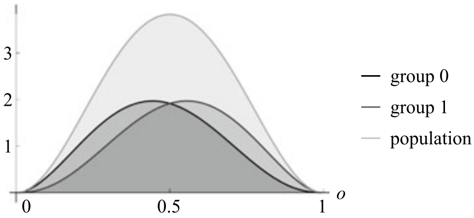
In both experiments, the desegregation rate d was varied from 0 to 0.5 in steps of 0.025, over 21 different levels. This resulted in variation of the average segregation measure S between $S = 1$ at $d = 0$ and $S \approx 0.522$ at $d = 0.5$. Except for $d \geq 0.45$ the 95% confidence intervals of mean S were non-overlapping for consecutive levels of d in a sample of 500 independent realizations per condition. For every level of d, 500 independent realizations of the simulation model were conducted, each running for 1000 $N = 100,000$ time steps. This was more than enough for all conditions in experiment 1 and 2 to reach stable outcomes.

It is an important question whether opinion polarization between groups can arise even if these groups have no systematic disagreement prior to interaction. For this reason, I drew in experiment 1 initial opinions randomly from the same Beta distribution $Beta(3, 3)$ for both groups, shown in Fig. 15.2a. This distribution has expected mean value of 0.5 and a standard-deviation of about 0.189. For culturally salient issues it is, however, more plausible that different groups also have different initial opinion tendencies. To model this, initial opinions were in experiment 2 randomly drawn from two symmetric Beta distributions $Beta(3, 3.5)$ and $Beta(3.5, 3)$ for groups 0 and 1, respectively, as shown in Fig. 15.2b. Mean opinions were about 0.462 for group 0 and 0.538 for group 1. Initial opinions in both groups had the same expected standard deviation of approximately 0.182.

The key outcome of interest in the simulation experiments was the degree of polarization both within the population as a whole and between the two groups. To assess between-group polarization, I measured the absolute value of the difference between the mean opinions in both groups, $P_t^g = |\overline{o_{t,g=1}} - \overline{o_{t,g=0}}|$. If this difference is close to one, this is a clear sign of strong between-group polarization. A low difference between the mean opinions of the groups, however, does not necessarily



(a) Experiment 1: both groups same initial distribution



(b) Experiment 2: moderate initial disagreement between groups

Fig. 15.2 Initial opinion distributions

show that there is no polarization at all. The population can also fall apart into two opposed factions that both contain members of both groups. To distinguish this form of 'population polarization' from between-group polarization, population polarization P_t^p at time point t is computed as the variance of all pairwise opinion distances in the population (adapted from [20]), as given by Eq. (15.4).

$$P_t^p = \frac{4}{N^2} \sum_{ij}^{i=N,j=N} \left(|o_{jt} - o_{it}| - \overline{|o_{kt} - o_{mt}|} \right)^2 . \tag{15.4}$$

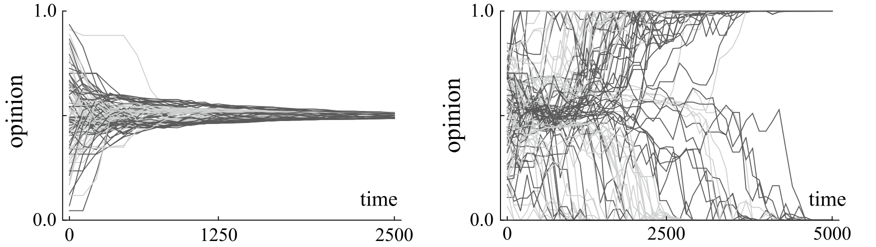
In Eq. (15.4), $\overline{|o_{kt} - o_{mt}|}$ denotes the average opinion distance across all pairs (km) in the population. The minimum level of polarization ($P = 0$) obtains when all pairwise distances are zero, corresponding to full consensus in the population. P^p obtains its maximal value of 1 if the population is split into two equally large factions with maximal mutual disagreement and full agreement within each of the factions.

Experiment 1

I begin with showing the dynamics for two prototypical runs. Figure 15.3a shows a run with maximal segregation $S = 1$, Fig. 15.3b displays a run with minimal segregation $S \approx 0.522$.

Figure 15.3 reveals remarkable differences between the two runs. In the maximally segregated population (Fig. 15.3a), members of both groups were quickly drawn to almost perfect population-wide consensus on an opinion at approximately 0.5. After 100,000 time steps, the standard deviation of opinions declined to practically zero.[2] This outcome occurred in about 90% of all runs in this condition. In the maximally desegregated population (Fig. 15.3b), the result was strikingly different.

[2]Perfect consensus is only obtained in the time limit. The simulation program computed a standard deviation of about 2.31 10^{-9} after 100,000 time steps for this run.



(a) Maximally segregated network ($S = 1$) (b) Maximally desegregated network ($S \approx 0.522$)

Fig. 15.3 Change opinions in single runs without initial difference between group means ($N = 100$, $\beta_D = 1/3$, $r = 5$). Dark: group 0, Light: group 1

The population was split almost perfectly into two opposing camps after 5000 time steps. Population polarization reached in this condition a level of $P_t^p \geq 0.99$ in 97% of all runs at $t = 100{,}000$. Yet, the emergent camps were not perfectly divided between groups. On average, between-group polarization was about $P_t^g = 0.497$ in the final state.

The strong difference between the two scenarios can be explained as follows. With high segregation, only few agents had neighbors who belong to another group. If interacting agents i and j belong to the same group, it is highly unlikely that their initial opinion disagreement is large enough to trigger repulsive influence ($w_{ij} < 0$). This happens only when their disagreement exceeds $|o_{jt} - o_{it}| = 0.75$. However, with the initial distribution of $Beta(3, 3)$ this was practically impossible.[3] Thus, within both groups influence was overwhelmingly assimilative, pulling all agents towards the mean value of the initial distribution (0.5). Only those few agents who were located on the interface between groups had outgroup-neighbors. With outgroup-neighbors, disagreement only needed to exceed $|o_{jt} - o_{it}| = 0.5$ to trigger repulsive influence. In a randomly chosen pair of neighbors from different groups, this happens at the outset with a probability of about 0.056. The few events of repulsive influence that occurred pushed agents to move away from each other towards the extremes of the opinion space. However, in most cases they were pulled back towards less extreme opinions in subsequent interactions with moderate ingroup members. This explains why in this condition about 90% of all runs ended in consensus. Yet, in the remaining approximately 10% of runs, interactions on the interface of groups became repulsive, driving agents on opposite sides of the boundary towards opposite extremes in the opinion space. Consensus within groups remained high at the same time. As a consequence, opinions of the two groups were driven apart. These runs ended in almost perfect between-group polarization.

In a highly desegregated population outcomes were different. With $S \approx 0.522$ agents had on average about 50% outgroup-neighbors. Likely, on at least some places in the network neighboring agents disagreed enough to develop repulsive influence. As a consequence, they increasingly shifted opinions away from each other, towards

[3]The probability was about 0.00155.

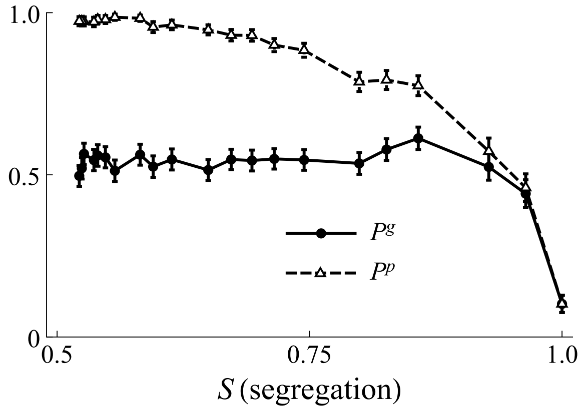


Fig. 15.4 Experiment 1: Effect of segregation on polarization measures. Bars indicate 0.95 CI around mean values

opposing extremes on the opinion scale. The dynamic of increasing differentiation between neighboring agents occurred simultaneously in different local regions of the network, because groups were well-mixed in this condition. This explains why population polarization reached its maximum of $P^p = 1.0$ here. At the same time, members of the two groups differentiated in different ways from each other in different local regions of the network. Thus, within both groups, members moved to both extreme ends of the opinion spectrum. Within the same group different poles were adopted at different places in the network. This was the reason why between-group polarization fell far below its theoretical maximum, with about $P^g = 0.5$ on average in the final state.

Figure 15.4 reports the results of experiment 1 for the entire range of segregation levels that were inspected. More precisely, the figure shows how the level of segregation S affected between-group polarization P^g and population polarization P^p in the final state, averaged across 500 realizations per condition. In line with the explanation given above, less segregation was on the whole associated with more population polarization. Also between-group polarization is on the whole higher in desegregated networks than in highly segregated ones. However, Fig. 15.4 also shows a non-linear association. When segregation increased from its minimum of $S \approx 0.522$, average between-group polarization remained fairly constant up to about $S \approx 0.8$, then increased to its peak-level at $S \approx 0.85$, to finally drop to a minimum of $P^g = 0.102$ in maximally segregated networks.

Figure 15.5 helps explaining the non-linearity identified by Fig. 15.4. Figure 15.5 shows the effect of segregation on the proportion of three types of outcomes in the final state, consensus ($P^p \leq 0.01$), group-split ($P^g \geq 0.99$) and population-split ($P^p \geq 0.99$). The share of runs with population-split and with consensus were largely mirror images of each other in this experiment. The more runs generated population-split, the less runs ended in consensus. In other words, lower segregation increasingly drove populations into a polarized state. But this state did not need to be group-split. In the region between about $S = 0.8$ and $S = 0.9$ population-split decreased with

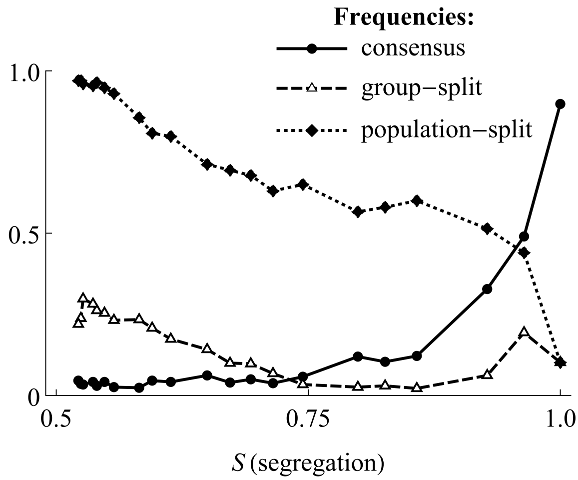


Fig. 15.5 Experiment 1: Effect of segregation on proportions of three outcome types

more segregation, while group-split increased. Above $S = 0.9$ both lines move again in the same direction.

The reason for the difference between group-split and population-split is the spatial coherence of groups under high segregation. This can be best understood by traversing the change of outcome proportions from right to left in Fig. 15.5, starting from a maximally segregated population. As the figure shows, moderate amounts of 'mixing' induce more polarization. This is due to more between-group interactions. But moderate mixing does not yet disrupt the spatial connectedness of groups, which therefore can still develop internal consensus. Thus, population polarization largely was found to be between-group polarization between $S = 0.9$ and $S = 1$. Once the segregation level was reduced further below $S = 0.9$, more interactions across group boundaries fueled more polarization, while consensus within groups was disrupted at the same time, due to more disconnectedness within groups. This explains the simultaneous decline of group-split and increase of population polarization when segregation moves downward from $S = 0.9$ to $S = 0.8$. Only when segregation levels further declined, even more individuals were spread across the network so that again most had at least some members of their own group in their local network, allowing for more within-group coordination in the process of population polarization. As a consequence, group-split and population-split moved again in the same direction when segregation levels were lower than about $S = 0.8$. However, the low levels of group-split between about $S = 0.75$ and $S = 0.9$ do not show that there was no systematic disagreement between groups at all. With an interaction range of $r = 5$, most individuals are connected with ingroup peers at all levels of segregation. Thus, some degree of coordination remains, explaining that on average between-group disagreement never fell below about 0.5, as shown by Fig. 15.4. Moreover, declining levels of group-split between $S = 0.5$ and $S = 0.8$ did not show up in declining average between-group polarization P^g, because also the proportion of runs in consensus declined in favor of more runs with medium-levels of between-group polarization.

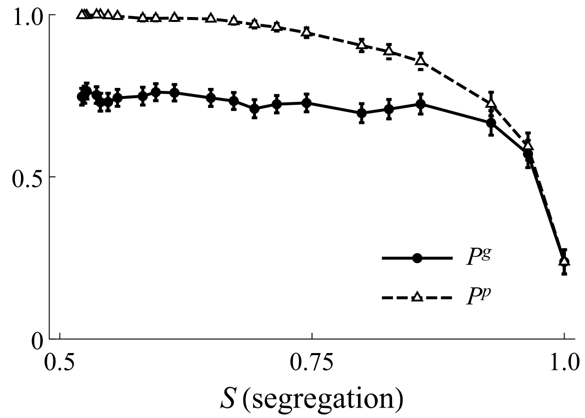


Fig. 15.6 Experiment 2: Effect of segregation on polarization measures. Bars indicate 0.95 CI around mean values

Experiment 2

Experiment 1 showed that considerable levels of between-group polarization came about in an unsegregated population, even when there was no systematic initial disagreement between groups. Experiment 2 tested whether the effect of segregation remained the same when mean opinions of the two groups differed from the outset.

Figures 15.6 and 15.7 report results of the ceteris-paribus replication of experiment 1, the only difference being that initial opinions were randomly drawn from the Beta-distributions shown in Fig. 15.2b. Comparison of Fig. 15.6 with the corresponding result for experiment 1 shows that on average between-group polarization in the final state was considerably higher across all levels of segregation. While in experiment 2 between-group polarization declined from about $P^g = 0.75$ to $P^g = 0.25$ between the lowest and the highest level of segregation, this decline happened at a lower level ($P^g \approx 0.5$ to $P^g \approx 0.1$) in experiment 1. Also population polarization was consistently higher in experiment 2, but this difference was less pronounced. A further noteworthy difference was that there was no longer a discernible non-monotonous effect of segregation on between-group polarization.

Figure 15.7 further confirms these differences and helps to explain them. The share of runs ending in group-split was slightly but consistently above the levels found in experiment 1, while the share of runs ending in consensus was slightly but consistently below this level. Population-split clearly was at a considerably higher level. The most striking qualitative difference was that group-split did no longer increase when small amounts of mixing were added to a maximally segregated network, but instead started to drop immediately. This illustrates the most important explanation for the differences between the experiments. In experiment 2, initial between-group differences were high enough to trigger mutual distancing on the interface between groups. Thus about 25% of runs were ending in group-split in the maximally segregated networks. While reducing segregation from this point fueled more population polarization - like in experiment 1 - it also blurred the boundaries between the groups

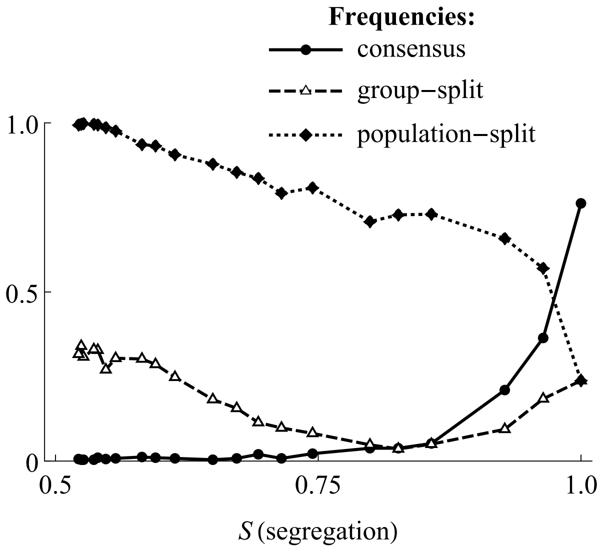


Fig. 15.7 Experiment 2: Effect of segregation on proportions of three outcome types

at the global level. Locally, agents from different groups were even more prone to end up on opposing sides of the spectrum than in experiment 1, but this was globally less coordinated than in the maximally segregated networks. This explains why starting from a maximally segregated network, mixing groups immediately decreased group-split in experiment 2, unlike it did in experiment 1.

Robustness Tests

The results of experiment 1 and experiment 2 rest on a number of assumptions about the model and the specific scenario. A full exploration of the robustness of results against meaningful variations is impossible in the space of this paper. As a start, I conducted two main robustness tests. First, a ceteris-paribus replication of both experiments was conducted with the range of interaction reduced from $r = 5$ to $r = 1$. A smaller range of interaction greatly reduces the interface between groups in highly segregated populations and inhibits the spreading of locally emergent extreme opinions in the network. The robustness test showed that this did not change the main qualitative effects of segregation found in experiments 1 and 2. More precisely, replicating experiment 1 with $r = 1$, it was found that increasing segregation from its minimal level first slightly increased, then reduced between-group polarization, while population polarization was reduced monotonously. Similarly, replicating experiment 2 it was found that there was no more non-monotonicity under higher initial between-group disagreement, but more segregation still reduced polarization both in the population and between groups. However, it should be noted that N1000 time steps were not enough to obtain stable outcomes in all conditions with $r = 1$. This is especially not the case in highly segregated networks, where polarization that starts on the interface between groups can take a long time to spread and pull all group members into opposing camps in the sparse network with $r = 1$.

The second robustness test was to reduce the relative impact of demographic group differences on discrepancy in the influence process. A relatively lower value of β_D can be expected to reduce the overall potential for polarization, because individuals from different groups need more disagreement to develop a mutually negative relationship. To assess this, a ceteris-paribus replication of experiments 1 and 2 was conducted, setting $\beta_D = 1/4$ (vs. $\beta_D = 1/3$ in the baseline condition). As expected, both forms of polarization declined. Most importantly, segregation still reduced polarization, where the difference between segregated and desegregated networks was actually larger than for $\beta_D = 1/3$ across both experiments.

Discussion and Conclusion

Intuitive reasoning as well as a number of formal models of opinion dynamics suggest that cultural diversity can under certain conditions be a threat to societal consensus, despite all its undoubted benefits. It has been argued that polarization between groups in a diverse society may be particularly likely when the society is highly segregated, echoing concerns raised by former U.S. president Barack Obama and results obtained with formal models of socially complex opinion dynamics. In these models, interactions between like-minded people can make them more and more convinced of their prevailing opinion tendencies, resulting in opinions that are increasingly extreme and different from those outside of their segregated world [13, 31, 33]. I presented in this paper a formal model drawing on social-psychological theories of cognitive balance that points to the opposite conclusion. Building on earlier work [17, 20, 21, 26, 29], this model combines assimilative with repulsive social influence, assuming that mutual disagreement between interacting agents is particularly likely to become accentuated and extreme when they interact with members of other groups that are separated from them by socially salient boundaries.

The point of my paper is not to show that the one or the other line of modeling is right or wrong about the link between segregation and polarization. What I would like to demonstrate is that formal modeling of socially complex dynamics can help us to better understand counter-intuitive and often unanticipated consequences of simple and familiar principles of social interaction. Principles such as influence, repulsion, persuasion, homophily or xenophobia are well known from our daily lives and from research conducted by social scientists. However, their possible implications at the societal level are often less well understood. An important contribution of social complexity models is that they can focus attention of empirical researchers on testing those assumptions in models that can be particularly critical for key social outcomes, such as polarization between groups.

The two lines of modeling work discussed in this chapter serve as an example in case. The obvious contradiction between their implications has motivated researchers in recent years to conduct systematic behavioral experiments. While this endeavor is still in progress, some of this work speaks to the models presented here. For example, while several empirical studies point to some evidence for repulsive influence in experimental and field settings [24, 27], recent experimental research tested repulsive influence more systematically in a controlled lab setting and found no support [10, 45]. This suggests that repulsive influence may be less easily triggered in social interactions than most formals models assume. At the same time, experimental tests have been conducted that lend some support to models of argument persuasion [31]. However, it would be premature to therefore entirely discard the possibility that segregation may sometimes preclude polarization. Experimental tests hitherto could not capture situations of strong between-group antagonism nor could they observe groups with mutually strongly exclusive social identities, conditions that appear to be plausible candidates for triggering repulsive influence that may drive groups apart.

Social complexity models have revealed important challenges for our scientific understanding of polarization. In line with calls from recent reviews of the field [22, 42], I believe that for tackling these challenges, we need to move forward towards a deeper connection of formal models with empirical insights from behavioral experiments and field research in the social sciences. The potential threat from polarization in diverse societies is an issue important enough to merit this effort.

Acknowledgements An earlier version of this paper has been presented at the Econophys-2017 & APEC-2017 Conference, held in November 2017 at the Jawaharlal Nehru University and Delhi University, New Delhi, India. The author wishes to thank the participants and especially the organizers of the conference, as well as the editors of this volume, for creating a pleasant and intellectually stimulating environment for this work.

References

1. Allport, G.W.: The Nature of Prejudice. Addison-Wesley, Cambridge (1954)
2. Altafini, C.: Consensus problems on networks with antagonistic interactions. IEEE Trans. Automat. Contr. **58**(4), 935–946 (2013). https://doi.org/10.1109/TAC.2012.2224251
3. Baldassarri, D., Bearman, P.: Dynamics of political polarization. Am. Sociol. Rev. **72**(5), 784–811 (2007)
4. Banisch, S., Olbrich, E.: Opinion polarization by learning from social feedback. J. Math. Sociol., (2018). https://doi.org/10.1080/0022250X.2018.1517761
5. Bianchi, F., Squazzoni, F.: Agent-based models in sociology. Wiley Interdiscip. Rev. Comput. Stat. **7**(4), 284–306 (2015). https://doi.org/10.1002/wics.1356
6. Bramson, A., Grim, P., Singer, D.J., Fisher, S., Berger, W., Sack, G., Flocken, C.: Disambiguation of social polarization concepts and measures. J. Math. Sociol. **40**(2), 80–111 (2016). https://doi.org/10.1080/0022250X.2016.1147443
7. Centola, D.: The social origins of networks and diffusion. Am. J. Sociol. **120**(5), 1295–1338 (2015). https://doi.org/10.1086/681275
8. Christ, O., Schmid, K., Lolliot, S., Swart, H., Stolle, D., Tausch, N., Al Ramiah, A., Wagner, U., Vertovec, S., Hewstone, M.: Contextual effect of positive intergroup contact on outgroup

prejudice. Proc. Natl. Acad. Sci. **111**(11), 3996–4000 (2014). https://doi.org/10.1073/pnas.1320901111

9. Clark, W.A.V., Fossett, M.: Understanding the social context of the Schelling segregation model. Proc. Natl. Acad. Sci. U. S. A. **105**(11), 4109–4114 (2008)
10. Clemm von Hohenberg, B., Maes, M., Pradelski, B.: Micro influence and macro dynamics of opinion formation (2017). https://ssrn.com/abstract=2974413
11. Collier, P.: Exodus. Immigration and Multiculturalism in the 21st Century. Penguin, London (2013)
12. Currarini, S., Jackson, M.O., Pin, P.: Identifying the roles of race-based choice and chance in high school friendship network formation. Proc. Natl. Acad. Sci. U. S. A. **107**(11), 4857–4861 (2010)
13. Dandekar, P., Goel, A., Lee, D.T.: Biased assimilation, homophily, and the dynamics of polarization. Proc. Natl. Acad. Sci. U. S. A. **110**(15), 5791–5796 (2013)
14. Dovidio, J.F., Love, A., Schellhaas, F.M.H., Hewstone, M.: Reducing intergroup bias through intergroup contact: twenty years of progress and future directions. Group Process. Intergroup Relat. **20**(5), 606–620 (2017). https://doi.org/10.1177/1368430217712052
15. Eger, S.: Opinion dynamics and wisdom under out-group discrimination. Math. Soc. Sci. **80**, 97–107 (2016). https://doi.org/10.1016/j.mathsocsci.2016.02.005
16. Ellemers, N., Rink, F.: Diversity in work groups. Curr. Opin. Psychol. **11**, 49–53 (2016). https://doi.org/10.1016/j.copsyc.2016.06.001
17. Feliciani, T., Flache, A., Tolsma, J.: How, when and where can spatial segregation induce opinion polarization? Two competing models. JASSS **20**(2), (2017). https://doi.org/10.18564/jasss.3419. http://jasss.soc.surrey.ac.uk/20/2/6.html
18. Fent, T., Groeber, P., Schweitzer, F.: Coexistence of social norms based on in- and out-group interactions. Adv. Complex Syst. **10**, 271–286 (2007). https://doi.org/10.1142/S0219525907000970
19. Festinger, L.: A Theory of Cognitive Dissonance. Stanford University Press, Stanford (1957)
20. Flache, A., Macy, M.W.: Small worlds and cultural polarization. J. Math. Sociol. **35**(1–3), 146–176 (2011). https://doi.org/10.1080/0022250X.2010.532261
21. Flache, A., Mäs, M.: How to get the timing right. A computational model of the effects of the timing of contacts on team cohesion in demographically diverse teams. Comput. Math. Organ. Theory **14**(1), 23–51 (2008). https://doi.org/10.1080/0022250X.2010.532261
22. Flache, A., Mäs, M., Feliciani, T., Chattoe-Brown, E., Deffuant, G., Huet, S., Lorenz, J.: Models of social influence: towards the next frontiers. JASSS **20**(4), (2017). https://doi.org/10.18564/jasss.3521. http://jasss.soc.surrey.ac.uk/20/4/2.html
23. Heider, F.: Attitudes and cognitive organization. J. Psychol. **21**(1), 107–112 (1946)
24. Hovland, C.I., Harvey, O.J., Sherif, M.: Assimilation and contrast effects in reactions to communication and attitude change. J. Abnorm. Soc. Psychol. **55**(2), 244–252 (1957). https://doi.org/10.1037/h0048480
25. Huet, S., Deffuant, G.: Openness leads to opinion stability and narrowness to volatility. Adv. Complex Syst. **13**(3), 405–423 (2010). https://doi.org/10.1142/S0219525910002633
26. Jager, W., Amblard, F.: Uniformity, bipolarization and pluriformity captured as generic stylized behavior with an agent-based simulation model of attitude change. Comput. Math. Organ. Theory **10**(4), 295–303 (2005)
27. Liu, C.C., Srivastava, S.B.: Pulling closer and moving apart: interaction, identity, and influence in the U.S. Senate, 1973–2009. Am. Sociol. Rev. **80**(1), 192–217 (2015). https://doi.org/10.1177/0003122414564182
28. Macy, M.W., Flache, A.: Social dynamics from the bottom up: agent-based models of social interaction. In: Bearman, P., Hedström, P. (eds.) Oxford Handbook of Analytical Sociology, pp. 245–268. Oxford University Press, Oxford (2009)
29. Macy, M.W., Kitts, J., Flache, A., Benard, S.: Polarization and dynamic networks. a hopfield model of emergent structure. In: Breiger, R., Carley, K., Pattison, P. (eds.) Dynamic Social Network Modeling and Analysis: Workshop Summary and Papers, pp. 162–173. The National Academies Press, Washington (2003)

30. Mark, N.P.: Culture and competition: homophily and distancing explanations for cultural niches. Am. Sociol. Rev. **68**(3), 319–345 (2003)
31. Mäs, M., Flache, A.: Differentiation without distancing. Explaining bi-polarization of opinions without negative influence. PLoS One **8**(11), (2013). https://doi.org/10.1371/journal.pone.0074516
32. Mäs, M., Flache, A., Kitts, J.A.: Cultural integration and differentiation in groups and organizations (2014). https://doi.org/10.1007/978-3-319-01952-9_5
33. Mäs, M., Flache, A., Takács, K., Jehn, K.A.: In the short term we divide, in the long term we unite: demographic crisscrossing and the effects of faultlines on subgroup polarization. Organ. Sci. **24**(3), 716–736 (2013). https://doi.org/10.1287/orsc.1120.0767
34. Mason, W.A., Conrey, F.R., Smith, E.R.: Situating social influence processes: dynamic, multidirectional flows of influence within social networks. Personal. Soc. Psychol. Rev. **11**(3), 279–300 (2007). https://doi.org/10.1177/1088868307301032
35. McPherson, M., Smith-Lovin, L., Cook, J.M.: Birds of a feather: homophily in social networks. Annu. Rev. Sociol. **27**(1), 415–444 (2001)
36. Moody, J.: Race, school integration, and friendship segregation in America. Am. J. Sociol. **107**(3), 679–716 (2001). https://doi.org/10.1086/338954
37. Myers, D.G.: Polarizing effects of social interaction. In: Brandstätter, H., Davis, J.H., Stocker-Kreichgauer, G. (eds.) Group Decision Making, pp. 125–161. Academic Press, London (1982)
38. Norris, P., Inglehart, R.F.: Muslim integration into western cultures: between origins and destinations. Polit. Stud. **60**(2), 228–251 (2012). https://doi.org/10.1111/j.1467-9248.2012.00951.x
39. Obama, B.: Farewell address (2017). Retrieved 06 07 2018. https://obamawhitehouse.archives.gov/Farewell
40. Pettigrew, T.F., Tropp, L.R.: A meta-analytic test of intergroup contact theory. J. Pers. Soc. Psychol. **90**, 751–783 (2006)
41. Proskurnikov, A.V., Matveev, A.S., Cao, M.: Opinion dynamics in social networks with hostile camps: consensus versus polarization. IEEE Trans. Automat. Contr. **61**(6), 1524–1536 (2016). https://doi.org/10.1109/TAC.2015.2471655
42. Sobkowicz, P.: Modelling opinion formation with physics tools: call for closer link with reality. J. Artif. Soc. Soc. Simul. **12**(1), 11 (2009). http://jasss.soc.surrey.ac.uk/12/1/11.html
43. Stark, T.H., Flache, A., Veenstra, R.: Generalization of positive and negative attitudes toward individuals to outgroup attitudes. Personal. Soc. Psychol. Bull. **39**(5), 608–622 (2013). https://doi.org/10.1177/0146167213480890
44. Stark, T.H., Mäs, M., Flache, A.: Liking and disliking minority-group classmates: explaining the mixed findings for the influence of ethnic classroom composition on interethnic attitudes. Soc. Sci. Res. **50**, 164–176 (2015). https://doi.org/10.1016/j.ssresearch.2014.11.008
45. Takács, K., Flache, A., Mäs, M.: Discrepancy and disliking do not induce negative opinion shifts. PLoS One **11**(6), e0157,948 (2016). https://doi.org/10.1371/journal.pone.0157948
46. Vinokur, A., Burnstein, E.: Depolarization of attitudes in groups. J. Pers. Soc. Psychol. **36**(8), 872–885 (1978)