University of Groningen

The diversity puzzle

Mäs, Michael

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version
Publisher's PDF, also known as Version of record

Publication date:
2010

Citation for published version (APA):
Mäs, M. (2010). *The diversity puzzle: explaining clustering and polarization of opinions*. [s.n.].

IV. Argument Exchange and Demographic Faultlines[14]

Abstract

In the previous chapter, we presented a new theory of opinion polarization. In this chapter, we will apply this model to the research on demographic faultlines in work teams. We will test if the new model also predicts the polarizing effects of demographic faultlines that we have identified in the negative-influence model (see Chapter II).

We will show that the argument-exchange model supports the faultline hypothesis. However, the new model points out several conditions for this effect which previous contributions have overlooked. First, even with a very strong faultline, opinions will only polarize in groups where individuals tend to select similar interaction partners. Second, polarization is more likely the stronger opinions and demographic attributes in a team are correlated initially, that is, prior to interaction between the group members.

Furthermore, our new model implies that the short term effects of demographic faultlines differ crucially from their long term effects. Groups where demographic attributes are not perfectly correlated will eventually arrive at consensus even though they might suffer from polarization in the short run. Counter-intuitively, the model implies that the convergence process is faster the *stronger* the demographic faultline is.

IV.1. Introduction

Demographic diversity at the workplace is a major challenge for organizations and is becoming an increasingly important issue as the economy globalizes (for comprehensive reviews about theoretical and empirical research see: Bowers, Pharmer and Salas 2000; Milliken and Martins 1996; Pelled 1996; Stewart 2006; van Knippenberg and Schippers 2007; Webber and Donahue 2001; Williams and O'Reilly 1998). For work teams, demographic diversity can be beneficial, because it broadens the social and human capital

---

[14] This chapter was written together with Andreas Flache, Károly Takács and Karen A. Jehn and is currently under review. The title of the original article is: "In the short term we divide, in the long term we unite. Crisscrossing work team members and the effects of faultlines on subgroup polarization"

of the team. However the benefits do not accrue automatically. Demographic dissimilarity between team members may, at the same time, cause conflicts and tensions and thus threaten performance. This lead Milliken and Martins to conclude in their review of the field that "diversity thus appears to be a double-edged sword"(1996: 403).

In the search for conditions that explain why diversity sometimes increases team performance and reduces it at other times, Lau and Murnighan (1998; 2005) proposed that the effects of diversity may decisively depend on the way demographic attributes, like age and gender, are distributed among team members. Their main hypothesis is that diversity impairs team functioning when the distribution of demographic attributes generates a *strong faultline*. "Group faultlines increase in strength as more attributes are highly correlated, reducing the number and increasing the homogeneity of resulting subgroups. In contrast, faultlines are weakest when attributes are not aligned and multiple subgroups can form" (Lau and Murnighan 1998: 328). They argue that diversity (demographics not aligned) increases the potential of a team for creativity and good performance but when the diversity is in a group with a strong faultline (demographics aligned), this potential may not be realized. The team may split up into subgroups with polarized opinions that cause conflicts between team members (Bezrukova, Thatcher and Jehn 2007). An intriguing implication of this theory is that an ideal workgroup composition might exist such that large pools of social and human capital can be obtained, but the damaging effects of diversity on cohesion can be avoided.

We contribute to the faultline research by elaborating the explanation of the faultline hypothesis (Lau and Murnighan 1998) and thereby revealing crucial implications of faultline theory that have been overlooked so far. We start by reviewing two theories that have been used to explain faultline effects and that are based on fundamentally different arguments. The first, Lau and Murnighan's (1998) theory, highlights that the interplay of homophilious selection of interaction partners with social influence breeds subgroup polarization in work teams with strong faultlines. Subgroup polarization is our main dependent variable and is defined as the degree to which a work team separates into subgroups that hold opposing opinions on work-related issues. The second theory has been developed almost a century ago in the classical sociological and anthropological literature and focuses on the integrating function of "crisscrossing actors" (Colson 1954; Evans-Pritchard 1939; Flap 1988; Galtung 1966; Lijphart 1977; Nieuwbeerta and Flap 2000; Ross 1920; Simmel 1922 (1908)). These actors share at least one demographic

attribute with another demographic subgroup than their own and thus, function as a bridge over the subgroup split that was caused by the faultline.

Lau and Murnighan (1998) did not take into account how crisscrossing actors can reduce subgroup polarization when the faultline is strong. We show that while their theory implicitly considers this, by not examining crisscrossing actors, the authors failed to realize some crucial implications of their theory. Foremost, we show that their theory predicts subgroup polarization only in the *short* term. We propose that in the long run, crisscrossing actors help to overcome subgroup polarization and group splits even in teams with very strong faultlines. Moreover, we show that subgroup polarization in the short term depends upon two further conditions that Lau and Murnighan implicitly assumed, but which they did not examine theoretically. First, strong faultlines entail subgroup polarization in the short run only when employees exhibit sufficiently *strong homophily* when selecting partners in their communication with other team members. That is, team members have a sufficiently strong preference for interacting with colleagues that are similar to them on certain attributes (McPherson, Smith-Lovin and Cook 2001). Second, we propose that faultlines entail group splits in the short run only when there is sufficient initial congruency between work-related opinions and demographic attributes in a work team (Homan et al. 2007: 82; Phillips 2003: 7; Phillips et al. 2004: 503).

Faultline effects result from a complex interplay of the interactions between multiple team members. As Harrison et al. (2007) have suggested, agent-based computational modeling is a powerful research method that allows to cope with this theoretical complexity and to reveal counter-intuitive implications of a theory. Frank and Fahrbach (1999) developed an agent-based model of complex and interrelated network and opinion dynamics in organizations, based on mechanisms that are very similar to those we assume. We follow their lead and use a formal modeling approach to show how Lau and Murnighan's reasoning, on the one hand, and the theory of crisscrossing actors, on the other hand, can be reconciled. We present and analyze a computational model that is based on Lau and Murnighan's assumptions and test if our new propositions do consistently follow from their theory. Our results support the proposition that faultline effects occur only in the short run and that strong homophily and initial congruency are crucial conditions for the effect of faultlines on group polarization. Finally, our formal analyses reveal a counter-intuitive effect of faultline strength. It turned out that the same communication structures that trigger short-term subgroup polarization in teams with

strong faultlines accelerate the process of consensus formation in the long run. Teams with strong faultlines might arrive at a consensus faster than teams with weak faultlines.

IV.2. Two Explanations of faultline effects

IV.2.1.   Lau and Murnighan's explanation of faultline effects

Lau and Murnighan argued that all newly formed teams go through a "sensemaking process of understanding each other and their task" (1998: 332) to coordinate similar opinions about what their task is, how to fulfill it, and how to devise work. In this process, the interplay of two core mechanisms can cause problems in teams with a strong faultline. First, Lau and Murnighan assume *homophilious selection of interaction partners*. Team members tend to associate with colleagues who share relevant demographic attributes. This assumption is prominently supported by a large body of sociological research on homophily (Lazarsfeld and Merton 1954) or the tendency of "birds of a feather flock together" that has consistently been identified as a strong force in social interactions (McPherson, Smith-Lovin and Cook 2001). Studies in both educational (e.g. Kandel 1978; Moody 2001) and organizational settings (e.g. Bacharach, Bamberger and Vashdi 2005; Ibarra 1992; McPherson and Smith-Lovin 1987; Ruef, Aldrich and Carter 2003) have provided empirical confirmation of the homophily concept.

Homophilious selection of interaction partners implies that the communication structures in a team crucially depend on faultline strength. To show this, we constructed 6 hypothetical teams of 20 individuals (see Figure IV.1). This size is not too big to be unrealistic for a work team, but also large enough to allow for a sufficiently fine-grained variation in the strength of demographic faultlines. Each team member is described by three dichotomous demographic attributes (symbolized as *black* vs. *white*; *A* vs. *B* and *rectangle* vs. *circle*). Diversity in all teams and on all demographic dimensions is kept at its maximum. That is, all three dichotomous variables have a distribution in which both values of the attribute are equally frequent (50:50). Teams differ, however, in the strength of the demographic faultline. Faultline strength is denoted by the symbol *f* and is measured here by the pairwise Pearson-correlation between all pairs of demographic attributes. Applying a method proposed by Flache and Mäs (2008a; 2008b) we varied faultline strength between maximal (all pairwise Pearson correlations are 1) and minimal faultline strength (all Pearson correlations are 0).

Figure IV.1: 6 hypothetical teams with different faultline strength
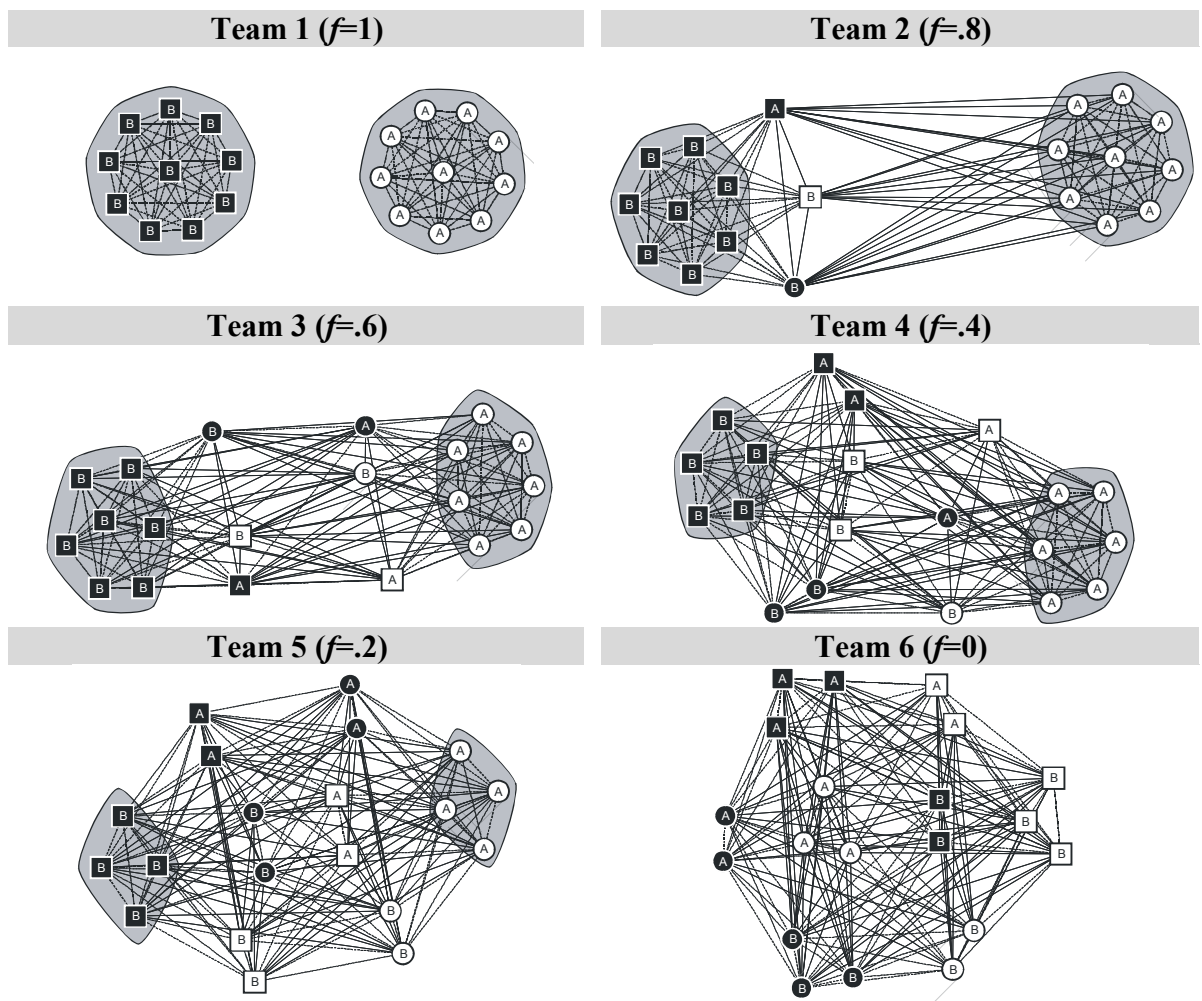
The network pictures shown in Figure IV.1 depict how faultline strength and the homophily mechanism shape the communication structure in the 6 work teams. Team members are represented by nodes. A pair of nodes is connected with a line if they have at least one demographic attribute in common. To depict the effects of homophilious selection of interaction partners, nodes have been arranged such that two individuals are nearer to each other the more attributes they share (Kamada and Kawai 1989; McFarland and Bender-deMoll 2007). The dashed circles identify the biggest subgroups of maximally similar individuals. In the team with the strongest faultline (team 1) the three demographic attributes correlate perfectly. Each pair of actors is either maximally similar or maximally *dis*similar and therefore either interacts frequently or never. On the team level, Figure IV.1 depicts two perfectly homogeneous but unconnected subgroups for team 1. As faultline strength decreases, however, this separation between subgroups becomes weaker and completely disappears as faultline strength is minimal (team 6). For instance, there are still two clearly distinct subgroups in team 3 (medium faultline strength). However, the

subgroups are smaller and there are also team members that can not be categorized into one of the subgroups. These actors share demographic attributes and therefore also interact with both subgroups. Note that these communication structures are a logical implication of faultline strength and the homophily mechanisms.

In addition to homophily, Lau and Murnighan assume that during interaction, team members exert influence on each others' opinions by *exchanging persuasive arguments* (Isenberg 1986; Myers 1982; Myers and Lamm 1976; Vinokur and Burnstein 1978). "Group members who support similar attitudinal positions will find that, as other members support that position using arguments different from their own, they each have more reason to become even more extreme than they were before" (1998: 332). Research on "polarization" (Myers 1982) has demonstrated how group members tend to become more extreme during group decision making. Faultline theory examines, however, not just polarization within one group, but it focuses on what we denote "subgroup polarization", the degree to which a work team separates into subgroups *holding opposing opinions* (Lau and Murnighan 1998). Subgroup polarization during a team's sense making process is problematic because it breeds emotional conflicts between the subgroups (Lau and Murnighan 1998) which, in turn, hamper good team performance (Jehn 1994).

The interplay of homophilious selection of interaction partners and influence with persuasive arguments can lead to subgroup polarization in groups with strong faultlines. As shown in Figure IV.1, homophily creates subgroups in teams with strong faultlines. Within subgroups, team members frequently exchange arguments but argument exchange *between* subgroups is rare. Lau and Murnighan argue that under these conditions, small initial opinion differences between the subgroups might be amplified during the sense making process. This is because subgroup members will mostly hear and share arguments that support their initial opinions (Stasser 1988), causing opinions in both subgroups to shift towards opposing ends of the opinion scale simultaneously. In other words, subgroup polarization increases. By contrast, in teams with weak faultlines, group members interact with colleagues who hold a variety of different opinions, such that no self-reinforcing dynamic towards emergent subgroup polarization can develop. This mechanism implies the central proposition of Lau and Murnighan's theory.

> **Proposition 1:** The stronger the faultline in a work team is, the stronger subgroup polarization will be.

IV.2.2.   The sociological explanation of faultline effects

Almost a century ago classical sociological and anthropological research on social order in stateless societies (Colson 1954; Evans-Pritchard 1939; Flap 1988; Galtung 1966; Lijphart 1977; Ross 1920; Simmel 1922 (1908)) revealed that strong faultlines may cause a problem for social integration. For instance, Ross argued in 1920 in a textbook:

"Suppose that at a given moment there is a certain strain along the line between Christians and Jews. If now, a strain appears along a quite different line, e.g. that between employers and workman, the religious opposition will become less intense. For Jewish bosses and Jewish workers will be estranged, likewise Christian bosses and Christian workman. On the other hand, Jewish and Christian capitalists will recognize that they are 'in the same boat', while Jewish and Christian workers will sympathize with one another as fellow victims of exploitation. Take the case of a tension between blacks and whites. If the lines of cleavage cross, each opposition will weaken the other. But if, as sometimes happens, all the employers are white and all the employed are black men, then one antagonism reinforces the other and the rift in society is deeper then ever. So, paradoxical as it may sound, a society riven by a dozen oppositions along lines running in various directions may actually be in less danger of early break-up than one split along just one line. For each new cleavage narrows the cross cleft, indeed, you might say that the society is sewed together by its inner conflicts"(Ross 1920: 164-165)

Although both faultline theory and the classical sociological and anthropological literature agree on the prediction that strong faultlines breed conflicts, they base their prediction on different sets of assumptions. Whereas faultline theory argues that the interplay of homophily and social influence may result in subgroup polarization, the sociological theory focuses on the integrating function of "crisscrossing" actors. Crisscrossing actors are individuals that share at least one demographic attribute with members of more than one demographic subgroup. Due to demographic similarity, they are attached to members of more subgroups and are thus able to conciliate in case of conflicts. Colson (1954), for instance, pointed to this in her studies of the African Tonga society in the 1940s. Each Tonga identified himself as a resident of his village and as a descendent of one of the Tonga-clans. However, societal rules prevented these two attributes from aligning. First, marriage between members of the same tribe was prohibited. Secondly, it was the man's privilege to choose the village where his family lived. Third, clan membership was organized matrilineally. That is, individuals belonged to their mother's descent group. Consequently each Tonga was attached by kinship to people living in different villages and at the same time by residence and his family members to descendents of different clans. Colson (1954) describes a conflict between the members of two clans that emerged after a member of one clan killed a member of the other clan. The conflict could be settled by persons that belonged to the victim's clan but lived in a village together with members of

the murderer's family. Because of their close relationships to both parties they could negotiate between the two groups and the conflict could be resolved.

From this sociological perspective, the faultline hypothesis follows because the more crisscrossing actors there are in a group, the stronger are the integrating forces that prevent conflicts. The number of crisscrossing actors in a group is, in turn, logically related to faultline strength. Figure IV.1 shows that the higher the number of those team members that are not part of one of the subgroups (i.e. crisscrossing actors), the weaker is the faultline. Teams with the maximal faultline strength (Team 1) consist of only two kinds of actors (*black*, *B*, *rectangles* and *white*, *A*, *circles*). There are no crisscrossing actors in this team. The number of crisscrossing actors, however, increases as faultlines become weaker. Team 2, for instance, still consists of two large subgroups. However, there are also three crisscrossing actors present.

IV.2.3.   *Integrating the two theories: Why time matters*

The processes that the two explanations of the faultline hypothesis propose appear to be fundamentally different. On the one hand, Lau and Murnighan (1998) argue that in teams with strong faultlines, subgroups form that develop increasingly different opinions that tear the team apart. The sociological theory, on the other hand, points to those actors that connect the subgroups and prevent conflicts. We argue that it is of great importance for our understanding of faultline effects to analyze how exactly these two processes are related to each other. We have shown that only groups with maximally strong faultlines have no crisscrossing actors. If crisscrossing actors can prevent group splits, does this then imply that their presence might neutralize the mechanism that Lau and Murnighan have described? Or would homophilious selection and persuasive arguments undermine the integrating effects of crisscrossing actors if the faultline is sufficiently strong?

It turns out that the same mechanisms from which Lau and Murnighan derive their faultline hypothesis can also be used to model the effects of crisscrossing actors. But, as we will show, when we explicitly integrate crisscrossing actors into Lau and Murnighan's reasoning, new consequences arise for the effect of faultline strength on the dynamics of subgroup polarization. With our integrating model, we can identify heretofore overlooked conditions under which the integrating effects of crisscrossing actors can be expected to prevail upon the dividing effect of a strong faultline.

Particularly, we argue that in teams with strong faultlines the processes that Lau and Murnighan's describe breed polarization only in the early stage of the sensemaking process.

Later, however, crisscrossing actors will help overcome group splits. Homophilious selection implies that crisscrossing actors interact with members of both subgroups, because they have some demographic similarity with members of each group. Based on persuasive argument theory (Myers 1982), we can expect that they will get arguments from all sides and will also communicate them to all subgroups they interact with. In this way, crisscrossing actors establish indirect communication between the subgroups who fail to interact directly. This may result in a gradual convergence of the subgroups' argument pools and also of their opinions, eventually reaching overall consensus in the work team.

This resonates with formal theories of the dynamics of social influence (Abelson 1964; Anderson 1991; French 1956; Harary 1959) This literature has suggested that social groups inevitably tend to reach consensus on initially controversial opinions, as long as the group has a connected interaction network in which no member is entirely cut off from influence by other group members or external sources. Based on Lau and Murnighan's assumption that demographic overlap implies interaction, crisscrossing actors can be seen as the link that integrates all group members into the network of mutual social influences. This suggests that, in principle, one single crisscrossing actor might suffice to create enough indirect communication between two subgroups to bring their opinions together. Thus, even in a group with a strong faultline, a small number of crisscrossing actors may ensure that no subgroup is entirely disconnected from outside influences. Accordingly, there should be *no* long run effect of faultline strength on subgroup polarization, except for the extreme case of a maximally strong faultline that divides the team into perfectly distinct subgroups. This absence of an effect of faultlines across almost the entire spectrum of possible teams is clearly contrary to what Lau and Murnighan (1998) suggest.

We propose that time is the critical factor and that crisscrossing actors help overcome group splits in the long run. But we also argue that in teams with a strong faultline, the polarizing dynamic of emergent subgroup splits will, in the short run, be stronger than the integrative dynamic of indirect interaction through crisscrossing actors. The reason is that in a team with a strong faultline, the members of the subgroups are by definition exposed to more other group members who are not connected to the outgroup than they are exposed to influences from crisscrossing actors. Accordingly, it is likely that consensus within the subgroups quickly develops and – based on the persuasive argument mechanism – subgroups initially polarize. At the same time, every member of the subgroup still has a positive probability of interacting with a crisscrossing actor at least from time to time. Whenever this happens, there is a chance that an argument from the outgroup is

adopted by ingroup members. This argument can subsequently rapidly spread in the ingroup. Due to the homophily principle, ingroup members are highly likely to interact with each other because they have both a high level of consensus and they are demographically similar. In other words, we propose that the same mechanisms that according to Lau and Murnighan imply subgroup polarization in the short term, also imply that subgroup splits are not stable in the long run if the group comprises crisscrossing members.

> **Proposition 2:** Subgroup polarization occurs only in the short run. In the long run, all teams where faultlines are not maximally strong will develop consensus and will overcome subgroup polarization.

IV.2.4. *Conditions of the short term effects of strong faultlines*

According to propositions 1 and 2, we expect that teams with strong faultlines will polarize in the short term, but will overcome the split in the long run. Moreover, we propose that the processes, that Lau and Murnighan describe, imply short-term polarization only if two necessary conditions are met. Following Flache and Mäs (2008a; 2008b), we argue that under Lau and Murnighan's assumptions the process of subgroup polarization crucially hinges on the assumption that initial congruency is sufficiently strong, meaning that opinions and demographic attributes in a team are already correlated initially, prior to interaction between the team members. If demographically similar group members do *not* share opinions more with each other than they do with demographically dissimilar others, then the exchange of arguments within demographic subgroups will not increase opinion differences between the groups because members of one subgroup do not learn more new arguments pro or con the original opinion than the actors in the other subgroup. As a consequence, subgroup polarization will not occur. Thus, an initial correlation between demographic attributes and opinions appears to be an essential condition for subgroup polarization in work teams.

> **Proposition 3:** Subgroup polarization increases in the beginning of the team dynamics only if the initial congruency is sufficiently high. Even teams with a strong faultline will not polarize if congruency is weak.

We furthermore propose that subgroup polarization can only take place if homophily is sufficiently strong. We define the strength of homophily as the *degree* to which interaction between similar actors is more likely than interaction between dissimilar actors. The strength of homophily in teams might be determined by the institutional context of work teams. For instance, in teams with high task interdependence workers are forced to

collaborate with both similar and dissimilar colleagues to fulfill their tasks. Thus, in these teams similarity will only weakly influence the choice of interaction partners. As a consequence, team members interact frequently with members who hold different opinions and will thus be influenced by them. Such a context would make it unlikely that the teams' opinions polarize, even if faultlines are strong (Molleman 2005).

> **Proposition 4:** Subgroup polarization increases in the beginning of the team dynamics only if homophily is sufficiently strong. Even teams with a strong faultline will not polarize if homophily is sufficiently weak.

IV.3. The Model

The exact logical implications of the combination of homophilious selection, persuasive influence, and faultline strength result from a complex interplay of these mechanisms simultaneously operating in multiple actors responding to each others' behavior. The method of computational agent-based modeling (Adner et al. 2009; Harrison et al. 2007; Macy and Willer 2002) provides multiple examples how in such a complex multi-agent system simple theoretical assumptions may generate counter-intuitive implications that would have been overlooked without model formalization (e.g. Frank and Fahrbach 1999; Siggelkow and Rivkin 2006; Stasser 1988). Accordingly, we conducted a strict test of the logical consistency of our reasoning (Adner et al. 2009) with a formal computational model. In fact, we test if the four propositions really follow logically from the assumptions of Lau and Murnighan's theory. Furthermore, our formal analyses revealed a new and unexpected effect of strong faultlines in work teams that has been overlooked in previous theorizing.

Our formal model is based on the two mechanisms of Lau and Murnighan's (1998) informal reasoning: homophilious selection of interaction partners and influence with persuasive arguments. In this model, each of the N team members is represented as an agent i characterized by D *demographic attributes* ($c_{i,d}$) and K *opinions* ($o_{i,k}$), where d refers to the d'th demographic dimension and k to the k'th opinion. The demographic attributes can either take the value 1 or -1 ($c_{i,d} \in \{-1;1\}$) and remain unchanged in the progress of interaction. The opinions of the actors vary between -1 and +1 ($-1 \le o_{i,k} \le +1$) and are open to influence.

Agents base their opinions on arguments $a_{k,l}$. For simplicity, we represent arguments as being either in favor of or against holding a pro-opinion ($o_{i,k} > 0$) on the corresponding issue ($a_{k,l} \in \{-1;1\}$). For each issue k there exist P pro arguments ($a_{k,l} = 1$) and C con

arguments ($a_{k,l}$ = -1). Which arguments exist in a given work team setting is summarized in the arguments matrix. This matrix has K columns and $P+C$ rows. Cells with a row number smaller than P+1 hold pro arguments, i.e. $a_{k,l}$ = +1. The remaining columns hold con arguments, i.e. a $a_{k,l}$ = -1. Matrix (a) in Figure IV.2 is an example of an argument matrix with one column (k=1) and 3 pro and 3 con arguments per issue (P=C=3).

Figure IV.2: Example of the updating process
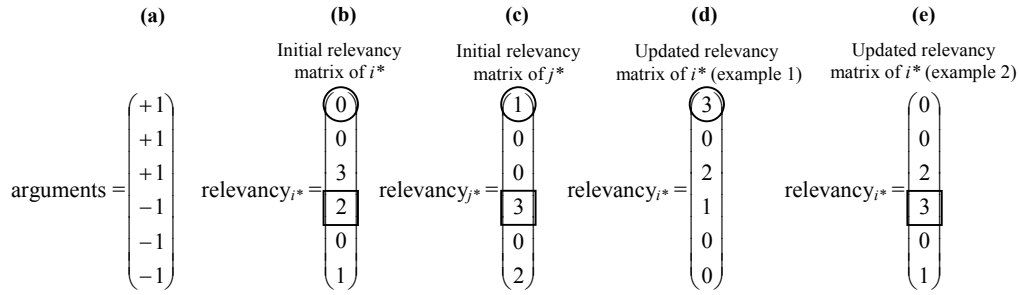
To take into account the limited cognitive capacities of humans, we assume that agents base their opinion not on all existing arguments but on a sample of S ($S \leq P+C$) arguments. Technically, an agent's opinion on issue k is the average value ($a_{k,l}$) of the arguments the agent considers as relevant (see equation 1). Thus, the more pro (con) arguments an agent's sample of arguments comprises the higher (lower) the value of the agent's opinion will be.

$$o_{i,k} = \frac{1}{S}\sum_{l=1}^{S} a_{k,l} \tag{1}$$

Furthermore, agents rank relevant arguments that address issue k according to their recency. As we show below, the more recent an argument is, the longer an agent will consider this argument for opinion formation. However, as equation (1) shows, the recency of an argument has no effect on the extent to which an argument shapes the current opinion. We denote recency ($s_{k,l,i}$) for agent i of the argument l that addresses issue k with integer values between 0 and S $s_{k,l,i} \in \{0,...S\}$. A recency value of $s_{k,l,i}$ = 0 indicates that the argument $a_{k,l}$ is *not* relevant for actor i. Values above zero indicate that this argument affects actor i's opinion on issue k. The most recent argument has the recency value of $s_{k,l,i}$ = S, the second most recent argument has the value S-1, and so on. Thus, if three arguments are relevant (S=3), then one has a recency of 1, one has a recency of 2, and one has a recency of 3. See matrix b in Figure IV.2 for an agent's relevance matrix for one issue (k=1), 6 existing arguments per issue (P=C=3) of which only three are relevant at a time (S=3). This

agent considers one pro argument and two con arguments. According to equation (1) the agent adopts an opinion of $o_{i,k}=-1/3$.

We model the sense making process of a team as a sequence of events, each event corresponding to one interaction between two agents. An interaction starts with the partner selection phase and is continued by the social influence phase. In the partner selection phase, two agents are matched for interaction, based on homophilious selection. Subsequently, an opinion of one of the interacting agents is updated, based on the persuasive argument mechanism.

IV.3.1.   Homophilious selection

We implement the partner selection phase as follows. In each event the computer program first randomly picks an agent i^*. All agents have at all events the same probability to be picked. Then an interaction partner j ($j \neq i^*$) is selected. To incorporate homophily, the probability that actor j is chosen as interaction partner depends on the similarity between i^* and j. As confirmed by empirical research we assume that both demographic similarity (Ibarra 1992; McPherson and Smith-Lovin 1987; McPherson, Smith-Lovin and Cook 2001; Ruef, Aldrich and Carter 2003) and opinion similarity (Byrne 1971) increase the probability to interact. Similarity sim_{i^*j} varies between 0 and 1. A similarity of zero means that the two actors are maximally dissimilar whereas a value of 1 indicates that both hold the same opinions and the same demographic attributes. We assume that all attributes are equally weighted in the calculation of similarity. Formally,

$$sim_{i^*,j} = \frac{1}{2 \cdot (D+K)} \left(\sum_{d=1}^{D} 2 - \left| c_{i,d} - c_{j,d} \right| + \sum_{k=1}^{K} 2 - \left| o_{i,k} - o_{j,k} \right| \right) \tag{2}$$

The probability that actor j is selected as interaction partner (p_j) is derived from the relative similarity of i^* and j compared to the similarities of i^* to all other actors, except i^* herself. To vary the *strength of homophily* we include furthermore a parameter h into the model ($h>0$). The higher the value of h, the more the relative similarity of the focal agent i^* and agent j increases the likelihood that j will be chosen as an interaction partner. Technically,

$$p_j = \frac{\left(sim_{i^*j} \right)^h}{\sum_{j=1, j \neq i^*}^{N} \left(sim_{i^*j} \right)^h} \tag{3}$$

Equation 3 shows that the more similar j is to $i*$, the higher is the probability that they will interact. If two actors differ maximally with regard to their opinions and their demographic attributes then the probability of interaction equals zero.

IV.3.2.   Persuasive Arguments

After the interaction partners $i*$ and $j*$ have been selected for the respective event, agent $i*$ is influenced by $j*$. We implement social influence through persuasive arguments in two steps. First, one of the arguments that $j*$ considers as relevant is adopted by $i*$. For this, one of the K opinions is selected randomly for update ($k*$), with the same probability ($1/K$) for all opinions. Then one of the S arguments that are relevant for j is picked ($a_{k*,j*}$) with equal probability ($1/S$) for all relevant arguments. Arguments that are not relevant for $j*$ are not chosen. The chosen argument is adopted by $i*$. Technically, the argument $a_{k*,j*}$ in $i*$'s relevance matrix adopts the value $S+1$ ($s_{k*,j*,i*} = S+1$).

When an agent's relevancy matrix has been updated repeatedly, it is likely that all existing arguments have been adopted at least once. However, it does not seem reasonable that after some time agents consider all arguments as relevant. To avoid this, we implemented a second step of the influence process. The second step assures that the number of arguments that are relevant for an agent remains constant at S during the whole sense making process. This implies that when an agent $i*$ has adopted an argument that has not been relevant before, one of the arguments that are currently relevant for $i*$ will be dropped. We assume that agents drop the argument that has been adopted least recently. This reflects the idea that every time an agent hears an argument from an interaction partner, the cognitive importance of that argument is reinforced. The longer ago an argument has been heard from another agent for the last time, the less important the argument is considered to be and sooner or later it will be seen as entirely unimportant. Technically, we implement this in the model such that the relevance matrix of $i*$ is updated by subtracting one from all non-zero recency values. The argument that was communicated between $i*$ and $j*$ in the present event adopts at the end of the iteration a recency value of S ($s_{k*,j*,i*} = S$). All other relevant arguments decline in recency. We also have tested alternative dropping rules, to assure the robustness of our results. Most importantly, we implemented that the argument for dropping is selected at random. Computational experiments revealed that all qualitative results reported below are robust to this modification of the model.

To illustrate the updating phase, Figure IV.2 contains two examples. Assume that matrix (b) is the initial relevancy matrix of agent i^*. Matrix (c) is the relevancy matrix of i^*'s interaction partner j^*. Before the update, the first argument is not relevant for i^* (see the circle in matrix (b)), but it is relevant for j^*. Hence, it is possible that i^* adopts the first argument, resulting in the updated relevance matrix for i^* shown in (d). Here the communicated argument is maximally recent (see the circle in matrix d). The recency of the remaining arguments has been reduced by 1. Note that this changed i^*'s opinion which shifted from -1/3 to +1/3 because i^* adopted a pro and dropped a con argument. As a second example, assume now that not the first argument is selected for update but argument number 4. This argument has already been relevant for i^* (see the square in matrix (b)). However, its recency has increased due to the interaction with j^* (see the square in matrix (e)). Note that this has no consequence on i^*'s opinion.

Interaction events are iterated until the system reaches equilibrium (Young 2001). Our model has exactly two equilibria, corresponding to perfect consensus or perfect subgroup polarization. Perfect consensus is reached when all agents hold the same opinions and, base these opinions on the same arguments. Then, no further change is possible. If all agents hold the same opinion but base that opinion on different arguments, then opinions can still change in upcoming events. Perfect subgroup polarization obtains if there are two subgroups, the members of each subgroup agree on all opinions and arguments with each other and the pairwise similarity ($sim_{i,j}$) between agents of different subgroups is zero. That is, the members of the subgroups maximally differ with respect to all demographic attributes and all opinions. If all members of the subgroups base their opinions on the same arguments, then this outcome is stable.

Obviously, the second equilibrium can only be reached in teams where faultline strength is maximal (f=1), because in these teams there are no crisscrossing agents. Crisscrossing agents share at least one demographic attribute with members of both demographic subgroups. Accordingly, if there is a crisscrossing agent and the two subgroups still disagree, a positive probability remains that arguments of the one subgroup are adopted by the other and the disagreement will vanish.

Some of our propositions and experiments focus on the duration of the sense making process, i.e. the time that it takes before consensus or perfect polarization has been reached. To be sure, we refrain from formulating statements about effects of the independent variables in our experiments on the *absolute* duration (e.g. in days or seconds) of the sense making process. We are not aware of any empirical evidence that would allow

assessing meaningfully the duration of a simulated interaction event in real time. However, it seems reasonable to assume that the more interaction events occur before equilibrium, the longer such a process also would take in real time. This allows us to compare the length of the process in terms of number of events under different conditions.

IV.4. Simulation Experiments

The central outcome variable of faultline theory is the level of subgroup polarization in work teams. To quantify subgroup polarization we use a measure called *polarization* (Flache and Mäs 2008a). It captures the degree to which the group can be separated into a small set of factions who are mutually antagonistic in the opinion space and have maximal internal agreement. To compute *polarization*, we use the variance of pairwise opinion agreement across all pairs of agents in the population, where agreement is ranging between -1 (total disagreement) and +1 (full agreement), measured as one minus the average distance of opinions (averaged across all K subdimensions). This measure obviously adopts its lowest level of zero for the case of perfect opinion consensus. The maximum level of opinion polarization (*polarization*=1) is obtained when the population is equally divided between the opposite ends of the opinion scale at -1 and +1 and all opinion dimensions are perfectly correlated.

To test whether our propositions follow consistently from the model, we conducted computational experiments varying three model parameters: the strength of faultlines (f), the initial correlation between demographic attributes and opinions (w) and the strength of homophily (h). For the remaining parameters we imposed values that are realistic and allow at the same time that the model generates sufficient variance in the outcome variables. Across all conditions, we assumed a team size of 20 (N=20) and used three fixed (demographic) attributes (D=3). Furthermore we assume that only one issue is relevant (K=1). Including further issues (K>1) makes it necessary to control for the correlations between the opinions and between each opinion and the demographic attributes. Since we focus here on the effects of demographic faultlines, we decided to keep the number of parameters varied in the experiment low and consider only one issue. Finally, we assumed that there exist always 10 pro (P=10) and 10 con (C=10) arguments. We assigned the value 4 to S in all conditions meaning that the actors base their opinions on 4 arguments.

To vary faultline strength (f) in the experiments, we used exactly the same distributions of demographic attributes that we used in Figure IV.1 (cf. Flache and Mäs

2008a; 2008b). We varied the Pearson correlation between each pair of demographic attributes from 0 to 1 in steps of .2. Of course, there are many alternative distributions of the three variables that result in the same *bivariate* correlations. The distributions we used, however, are the only ones that produce equal correlations between all pairs of demographic attributes and at the same time keep diversity maximal. We chose equal correlations to resolve a conceptual unclarity in Lau and Murnighan's definition of faultline strength. Do we, for example, speak of a strong faultline if two variables x and y are perfectly correlated but completely unrelated to a third variable z? Or, would we regard the faultline as stronger or weaker if x and y are correlated only with r=.8 but the correlation between x and z would rise to .6? These questions do not occur if all pairs or variables are equally correlated. Furthermore, considering unequal correlations between the demographic attributes does not affect the results. Also with unequal correlations, it holds that the weaker the correlation between the demographic dimensions, the smaller are the subgroups and the more crisscrossing actors there are in a team. Furthermore, as long as not all pairwise correlations are maximal (f=1), crisscrossing actors must be present.

To manipulate the level of initial congruency (w), we related the initial opinion to the first demographic attribute. The extent to which this affects the correlation of the opinion with the remaining demographic attributes depends on faultline strength (f). The stronger the faultline, the higher is the correlation between the first demographic attribute and the other demographic attributes. Accordingly, the stronger the faultline, the more similar are the correlations between the opinion and the first, second etc. demographic attribute. Technically, we assigned S arguments to each agent. For each of the S arguments, we assigned one of the existing pro arguments with the probability w when the agent holds the value 1 at the first demographic attribute and one of the con arguments otherwise. Agents with the value -1 at the first demographic attribute received a pro argument with probability 1-w. For instance, if w is 0.5, then pro and con arguments always have the same probability to be assigned. On average, this results in a Pearson-correlation between the first demographic attribute and the opinion of zero. However, as w increases, agents with the value 1 (-1) at the first demographic attribute more likely receive a pro (con) argument. This entails a higher Pearson-correlation between the first demographic attribute and the opinion as w increases. Under w=1, the opinion and the first demographic attribute perfectly align. More precisely, all agents that hold the value 1 at the first demographic attribute also hold opinion values of 1 and all agents who belong to the other demographic subgroup on the first dimension, hold opinion values of -1.

We varied *w* between .5 and 1 in steps of .1. We do not consider *w*-values below .5. Such values would lead to a negative correlation between the opinion and the demographic attributes. Since the actual values of the opinion and the demographic attributes have no substantial meaning, it makes no difference if opinions and demographic are positively or negatively correlated. To test the effects of the strength of homophily we manipulated the parameter *h* (see equation 3), varying it between 1 and 5 steps of 1. A value of *h*=1 expresses that agents have a weak preference to interact with similar team mates. The value of *h*=5 corresponds to a very strong homophily.

All in all we inspect 6*6*5 = 180 conditions in our computational experiments. The 5 conditions in which faultline strength is maximal (*f*=1) and the initial correlation between opinions and demographic attributes is maximal (*w*=1) have been excluded because the similarity (*sim_{i,j}*) between agents is under this condition either 1 or 0. In these cases, it is logically impossible that members of different subgroups will interact and opinions will therefore not change. For reliability, we conducted 500 independent replications per condition.

IV.5. Results

We present the results in three steps. In the first step, we present two ideal-typical simulation runs to illustrate model dynamics. We then turn to the consistency tests of the four propositions. Finally, we present additional analyses that revealed an unexpected effect of faultline strength.

IV.5.1. Ideal-typical simulation runs

Figure IV.3 demonstrates an ideal-typical simulation run with maximal faultline strength (*f*=1). To trigger subgroup polarization, we imposed conditions that, according to the propositions, make polarization very likely. We assumed strong homophily (*h*=5) and imposed a relatively strong correlation of initial opinions with demographic attributes (*w*=.8). The latter generated for this run an initial Pearson correlation between the opinion and the three demographic attributes of .77. Figure IV.3 shows the development of *polarization* and the distribution of the opinion at different stages of the simulation run. The histograms show the respective opinion distribution. The network pictures describe the resulting interaction structure. In the network pictures, each agent is represented by a circle. The color of a circle indicates to which of the two demographic subgroups the respective agent belongs. Each pair of agents that has a nonzero overall similarity (*sim_{i,j}*) is connected

by a line, symbolizing that there is a nonzero probability that these two agents interact. Because we focus here on the development of opinions in the team, the arrangement of the circles is only based on opinion similarity. Thus, circles are arranged in a way such that the nearer agents are placed to each other, the more similar their opinions are (Kamada and Kawai 1989; McFarland and Bender-deMoll 2007).

Figure IV.3: Ideal-typical run with maximal faultline strength (f=1, h=5, w=.8)
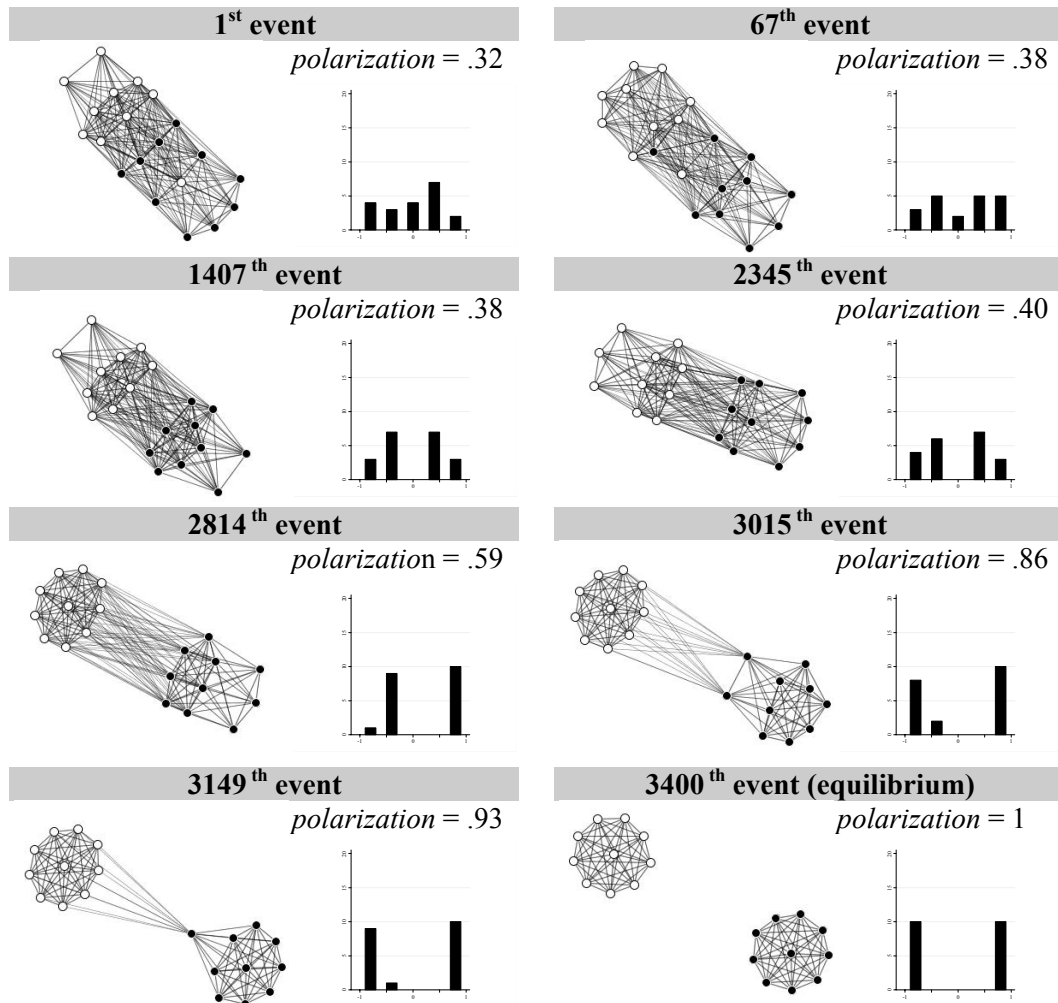
Initially (1st event in Figure IV.3), the opinion was almost uniformly distributed in this simulation run. Nevertheless, the corresponding network picture reveals that there are already initially systematic opinion differences between the demographic subgroups. The change of the histograms of the subsequent events shows that over time opinion differences between the subgroups increase. Consequently, the number of lines between the subgroups also decreases over time. Eventually (by event 3400) the subgroups hold maximally opposing opinions. The exchange of arguments between subgroups stopped at

this point, because there is neither an overlap in demographic attributes nor in opinions between agents from different subgroups. Opinion changes have now become impossible because agents only interact with team members hold the same opinion and arguments.

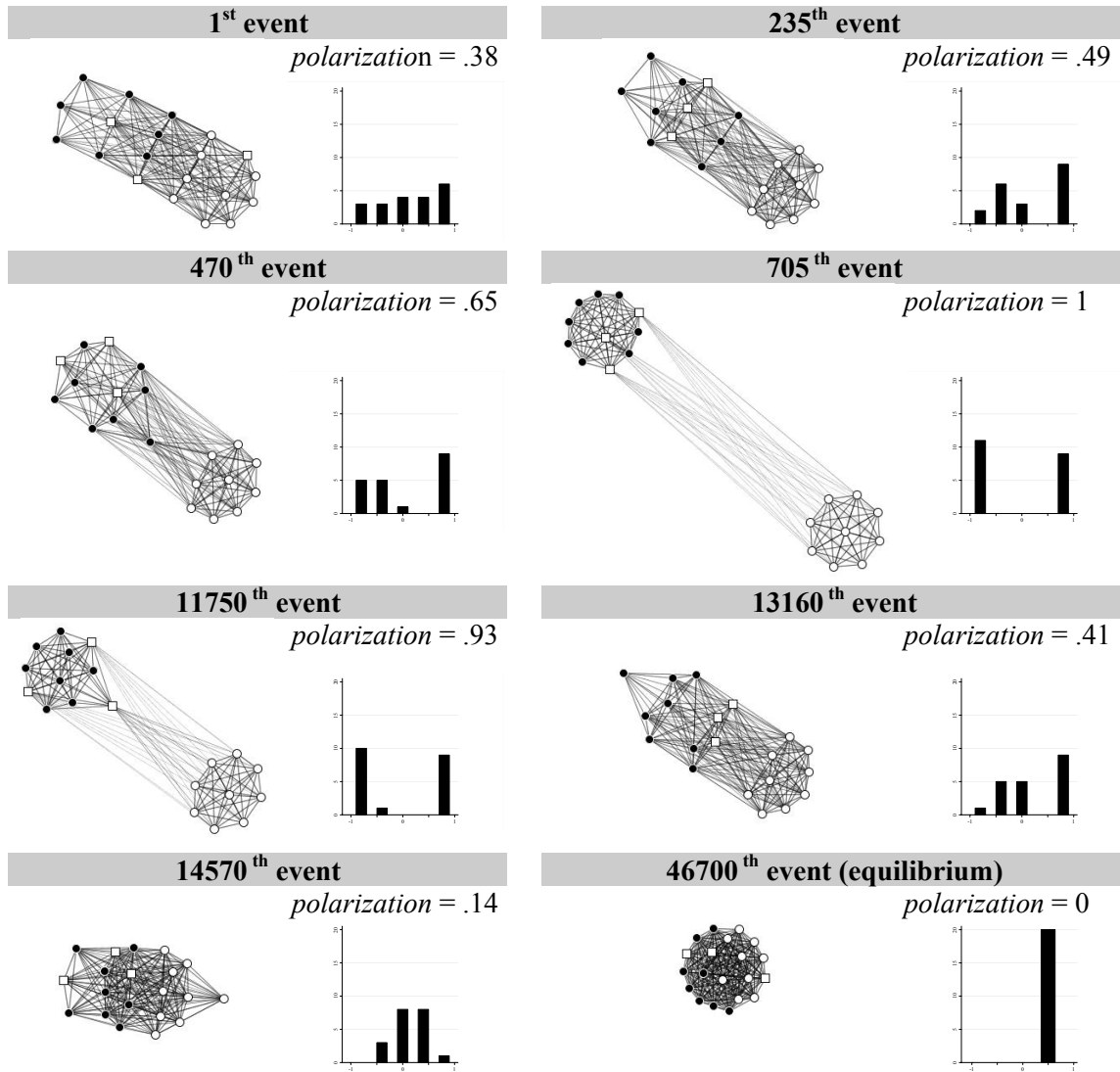**Figure IV.4:** Ideal-typical run with 3 crisscrossing agents (*f*=.8, *h*=5, *w*=.8)

Figure IV.3 depicts ideal-typical dynamics that ended in a stable group split. This shows that our model can generate the dynamics that Lau and Murnighan described in their informal reasoning if the faultline is maximally strong. Proposition 2, however, expects that dynamics differ crucially when crisscrossing actors are present. Figure IV.4 shows an ideal-typical run that supports the proposition. In this run, faultlines were slightly weaker than in the condition of Figure IV.3. For comparison, we retain all further parameters of the first illustrative run with maximal faultline strength (*h*=5, *w*=0.8), but we slightly reduce the

strength of the faultline to $f=0.8$. Now, the team contains three crisscrossing agents (see the three squares in network pictures of Figure IV.4). Initially (see 1st event), the opinion is again almost uniformly distributed and the demographic subgroups already hold somewhat different opinions. Again, we observe increasing subgroup polarization, just as Lau and Murnighan proposed. After the 705th event the work teams fell apart into two opposing subgroups with maximally different opinions. Within the two subgroups, the agents share the same opinions and also quickly coordinate on a common vector of arguments. However, the two subgroups are not completely unconnected: due to the three crisscrossing actors there is still some exchange of arguments between the subgroups. The network picture of the 11750th event demonstrates that one of the crisscrossing actors adopted an argument that changed his opinion. Subsequently, this argument spreads in the crisscrossing actors' subgroup and the opinion differences between the subgroups decrease (see event 13160). This convergence process continues until overall consensus is reached.

IV.5.2.   Consistency tests of the propositions

Long- term effects. According to Lau and Murnighan, higher faultline strength entails more subgroup polarization (proposition 1). Proposition 2, however, claims that in teams with non-maximal faultline strength this effect can only be observed in the short term. Our experiments clearly confirmed proposition 2. All simulated work teams with faultline strength below its theoretical maximum eventually ended in overall opinion consensus. That is, all team members held the same opinion and based it on exactly the same arguments.

In teams with maximally strong faultlines, however, we found perfect subgroup polarization, but not in all simulation runs. Figure IV.5 shows how the initial congruency (w) and the strength of homophily (h) affected the frequency of runs that ended in perfect subgroup polarization. More precisely, the size of the bubbles in Figure IV.5 corresponds to the percentage of runs under the respective condition that ended in stable group splits. For instance, 49.6% of 500 runs with very strong homophily ($h=5$) and very strong initial correlation between the opinion and the demographic attributes ($w=.9$) ended in a group split with two equally large subgroups and maximally opposing opinions. As suggested by the propositions 3 and 4, the higher the values of h and w are, the more likely subgroup polarization occurs. Figure IV.5 thus confirms that our implementation of Lau and Murnighan's mechanisms can explain stable subgroup polarization in teams with maximally strong faultlines. But, even with maximal faultline strength, group splits remain unlikely if

homophily is weak or the opinions are not already initially strongly aligned with the demographic attributes.

**Figure IV.5:** Percentage of runs that ended in stable splits (*f*=1)
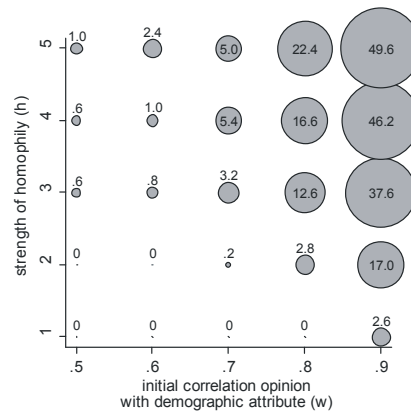
Short-term effects. Proposition 2 predicts increasing subgroup polarization in the short term even though teams reach consensus in the long term. To asses the short term polarization in the simulated work teams we measured the maximal value of *polarization* that teams exhibited during simulation runs and compared this value to the runs' initial value of polarization. The difference between these two values indicates to which degree groups split up in the short run independent on whether the split occurred right at the beginning of the run or later. To test if faultlines trigger short term polarization (proposition 2), Figure IV.6 shows bar graphs broken down by faultline strength (*f*). The gray part of each bar in Figure IV.6 depicts the average initial level of *polarization* in the teams. The black part of the bars shows the average increase in *polarization*. Both parts add up to the average of the maximal value of *polarization*. We excluded the conditions where *w*=1, because here *polarization* is initialized at it's logical maximum and cannot further increase. Figure IV.6 shows a stronger increase in the maximal value of *polarization* as faultlines become stronger (*f*) and thus supports proposition 2. At least in the short run, faultlines trigger subgroup polarization.

Figure IV.6: Average maximal opinion *polarization* over *f*, (15000 runs per bar)
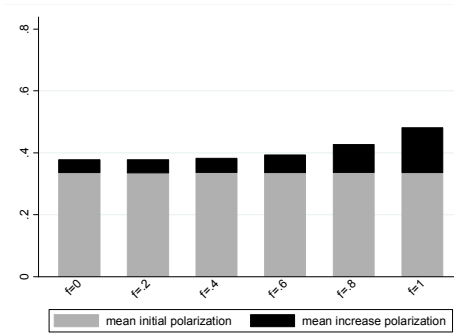
According to proposition 3, the higher the initial congruency, the stronger should be short term effects of strong faultlines on polarization. To test that, we display in Figure IV.7 the effects of congruency (*w*) on maximal *polarization* broken down by faultline strength *f*. The figure shows that the initial *polarization* depends on *w* (see the gray parts of the bars). This is a technical consequence of congruency that occurs because opinions align closer with the 50:50 split on the values of +1 and -1 in the first demographic attribute, as *w* increases. There is, however, no such relationship of initial opinion polarization to faultline strength because the distribution of the first demographic attribute is the same for all levels of faultline strength. It turns out that the maximal value of *polarization* (see the complete bars) in all subgraphs increases with *w*. However, as the size of the black areas shows, this is mainly the result our manipulation of *w*. If faultlines are not strong (*f*<.8), the mean increase of *polarization* declines with the initial correlation between opinion and the first demographic attribute. We believe that this results from a ceiling effect. If faultlines are weak, then most pairs of agents have a relatively high similarity (*sim*$_{ij}$) because of shared demographic attributes. The potential of opinion polarization in these teams is thus very low. If *w* is high, these teams start out close to their potential maximum of *polarization*. As a consequence, *polarization* can only rise moderately above the initial level and will decline soon thereafter. If faultlines are strong (*f*>.6), however, the model produces the effect of *w* that Proposition 3 expected. The black parts of the bars in the subgraphs for faultline strengths of .8 and 1 show that a higher initial correlation of opinion and first demographic attribute entails more opinion polarization.

Figure IV.7: Average maximal opinion *polarization* over *w*, by *f* (2500 runs per bar)
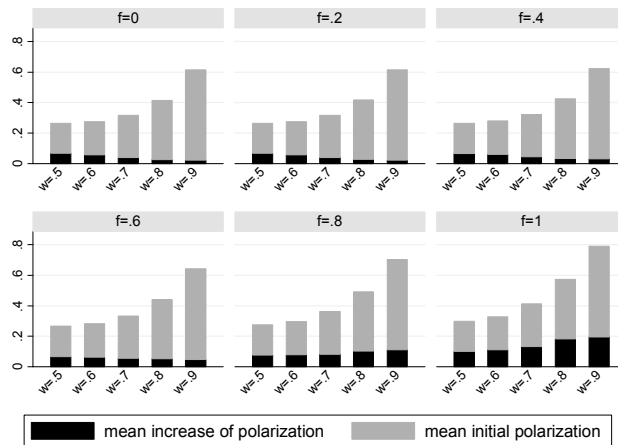
Proposition 4 suggests that subgroup polarization should increase with stronger homophily. Figure IV.8 confirms that the increase in *polarization* (see the black parts of the bars) is higher for stronger homophily (*h*). Comparison of different faultline levels also reveals that the effect of homophily strength increases in the strength of faultlines (*f*). If faultlines are weak, then even a very strong preference of the agents to interact with similar team members will cause only little increase in *polarization* in the short run. If faultlines are stronger, then strong homophily results in a larger increase in *polarization*.

Figure IV.8: Average maximal opinion *polarization* over *h*, by *f* (2500 runs per bar)
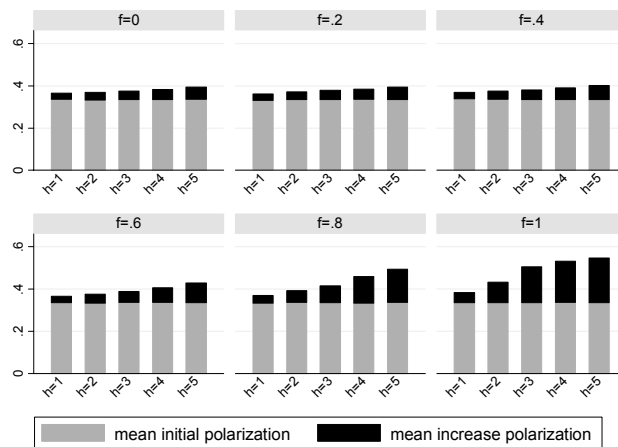

IV.5.3.   Relative time until convergence

The simulation experiments have confirmed that all teams that contain crisscrossing actors eventually arrived at consensus, even though many polarized in the short term. The analyses of the length of this convergence process, however, led to an unexpected and counter-intuitive result: the stronger the faultline in a team and the stronger homophily, the

faster the teams arrive at consensus. Figure IV.9 shows a bubble graph expressing the average number of events it took until the runs ended in overall consensus, broken down by faultline strength (*f*) and homophily strength (*h*). The graph shows that the less events were needed to reach consensus, the stronger the faultline was. It also shows that stronger homophily is associated with faster emergence of opinion consensus.

To confirm this counter-intuitive result, we conducted simulation experiments where teams started with perfect polarization (*w*=1) and varied faultline strength. In the runs with weak faultlines (*f*=0) the teams very quickly overcame the group split but it took them very long to arrive at consensus. By contrast, it took the teams with a strong faultline (*f*=.8) longer to overcome the initial group split. However, once the split was overcome, the teams quickly found a consensus.

Figure IV.9: Average number of events until the teams arrived at an overall consensus
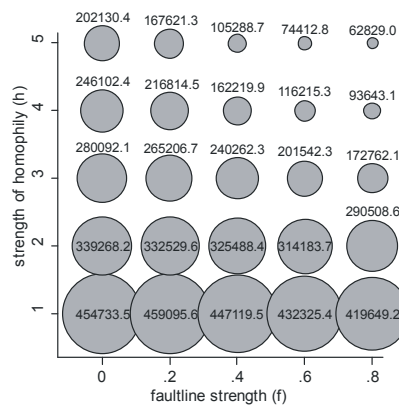
We explain this effect as a consequence of the interaction structure in teams with strong faultlines. The same interaction structure that causes subgroup polarization in the short term accelerates the convergence process as soon as the opinion split has been overcome. As we have shown in Figure IV.1, teams with strong faultlines consist of subgroups. Frequent exchange of arguments within the subgroups makes the subgroups reach internal consensus quickly. When there are crisscrossing agents, however, from time to time new arguments enter a subgroup and lead to changing opinions and a new subgroup consensus. If there are only a few crisscrossing agents present (strong faultlines), this process leads to a gradual convergence of opinions across the subgroups. Most importantly, this accelerates coordination on a single vector of arguments in the whole team much faster. The reason is that subgroups adopt a new argument and drop one of the arguments used before. Once an argument is dropped by a subgroup, it will not reoccur in later interactions. With weak

faultlines, however, there are more crisscrossing agents and new arguments enter the discussion within subgroups more frequently. This can be so frequent that subgroups do not manage to find consensus before a new argument enters. As a consequence, the number of arguments and opinion diversity within each subgroup remains high and the gradual convergence of subgroups that we found in groups with strong faultlines does not develop. Furthermore, frequent argument exchange with crisscrossing actors leads to a fast spread of arguments across the entire team. Thus, if a subgroup collectively drops an argument, this argument may still be used by other team members and might re-enter the discussion in the subgroup over and over again. Overall, the convergence of opinions occurs faster in the structured interaction network of a team with a strong faultline than in the unstructured communication pattern in a team with a weak faultline.

IV.6. Summary and Implications

Lau and Murnighan (1998) argued that teams with a strong demographic faultline likely experience subgroup polarization. We challenged this prediction, arguing that Lau and Murnighan overlooked the important role of crisscrossing actors in the sense making process of teams. Crisscrossing actors are team members who share some demographic attributes with multiple subgroups and can thus function as a bridge over the faultline. We showed that the faultline concept implies that even teams with very strong faultlines comprise at least a few crisscrossing actors. Accordingly, we argued that also in teams with strong faultlines there are processes that could prevent subgroup polarization or, if teams are polarized, help to overcome group splits. This led us to propose that strong faultlines breed subgroup polarization only in the short run. If there are crisscrossing actors in a team, even teams with strong faultlines will eventually overcome polarization. Moreover, we propose that Lau and Murnighan's theory implicitly factors crisscrossing effects in, although they did consider this explicitly. To underpin this claim, we developed a formal model based on the central behavioral assumptions of Lau and Murnighan's theory. We conducted computational experiments to test whether our new propositions follow consistently from the behavioral assumptions. Our analyses clearly confirmed this. We also proposed that faultline effects may crucially depend on core assumptions hidden in previous theoretical elaborations. We found that stronger faultlines only imply opinion polarization if demographic attributes are strongly correlated with the opinions of team members even before they influence each other. Moreover, to logically derive effects of strong faultlines, the assumption is needed that homophilious selection plays an important

role in interactions within the team. Finally, contrary to intuition, our simulations revealed that teams with strong faultlines might be faster in arriving at an opinion consensus.

Our analyses confirm that teams with strong faultlines experience more polarization than teams with weak faultlines. However, if faultlines are not maximally strong, effects of faultline strength occur only for the short term dynamics in a team. In the long run, group splits disappear sooner or later. This appears to be good news for managers. Nevertheless, we advise readers to interpret our results with caution. The main purpose of our analysis was to point to hidden implications of the mechanisms assumed by faultline theory. This does not preclude that other mechanisms not considered by the theory may lead to different consequences. Specifically, our formal model did not consider the possibility that social identities form around subgroups in the process of a group split. Members of the subgroups may then "act to legitimize the subgroups, and conflict between them may continue to be likely" (Lau and Murnighan 1998: 333). Strong subgroup identification may motivate team members to refuse communication with crisscrossing actors. Identification might also promote the development of stereotypes about the demographic subgroups. Since crisscrossing actors fit into none of the stereotypes, they may be rejected by members of both demographic subgroups. If such negativity arises then crisscrossing actors will not be able to conciliate.

Despite the possibility that identity formation may reduce the influence of crisscrossing actors, our results should also not be discarded too readily. We have shown that integrating effects of crisscrossing actors can in the long run only be precluded if these actors are perfectly excluded from the interaction networks within the subgroups. Even if subgroup identities form, it seems a rather extreme assumption that they can entirely prevent any subgroup influence via crisscrossing actors. It seems more plausible that the strength of subgroup identities affects how long it takes until the initial group splits can be overcome, but not the eventual outcome given the team has enough time to converge to a consensus. This also suggests that in the actual practice of work teams, crisscrossing actors may be important to overcome the negative effects of faultlines if management succeeds in creating conditions that support their integrating role. For example, an amicable and friendly environment in the work team may be important to reduce subgroup identifications and may therefore facilitate the exchange of arguments between the subgroups via crisscrossing actors.

This paper revealed that short term consequences of group dynamics might crucially differ from their effects in the long run. Other recent contributions also proposed effects

of time on consequences of demographic diversity in work groups. Most prominently, Harrison et al. (2002) argued that as team members get to know each other, the relevance of surface level (demographic) characteristics will diminish and members will base selection of interaction partners more on psychological similarity (personality, values, attitudes, beliefs). Like our reasoning, their argument suggests that the impact of demographic diversity and thus of demographic faultlines declines over time (see also: Pelled, Eisenhardt and Xin 1999). However, we have shown that this follows already from the elementary behavioral assumptions of faultline theory, without the need to necessarily include additional mechanisms such as the distinction between surface similarity and psychological similarity. This demonstrates that already relatively simple models of social processes can be too complex to grasp their logical consequences by informal reasoning. Formal methods, therefore, are useful to study such complex systems and to reveal unexpected consequences of theories that may remain undiscovered otherwise.