Tolerance and Compromise in Social Networks

Garance Genicot

Georgetown University, National Bureau of Economic Research, and Center for Economic and Policy Research

Individuals typically differ in their identities—the behaviors they deem ideal for themselves and for the members of their network—and in their tolerance for behaviors that deviate from their ideals. This paper studies compromise—that is, departures from one's ideal point, to be accepted by others. I show that an individual's compromise in equilibrium is bounded by the difference between her tolerance level and the lowest tolerance level in society. Relatively intolerant individuals, who serve as "bridges," are critical for reciprocated compromise. When individuals with extreme identities are systematically less tolerant, societies polarize. In contrast, intolerance among moderates encourages cohesion.

I. Motivation

Bernard Crick defines tolerance as "the degree to which we accept things of which we disapprove" (Crick [1963] 1971). It is the ability or willingness to withstand opinions or behaviors that one may not necessarily agree with.

There is a large body of evidence documenting the presence of homophily in human relationships—the tendency to associate with those similar

Debraj Ray, Joan Esteban, Yann Bramoullé, Ian Gale, Marcel Fafchamps, Gaurav Bagwe, and Laurent Bouton provided very helpful conversations. I am also grateful for the feedback from participants to the Theoretical Research in Development Economics (ThReD) conference in Oslo, the Network Conference in Cambridge, and the Conference on Identity in Namur, as well as seminars at Iowa State University, New York University, Aix-Marseille School of Economics, the European Institute, George Washington University, Harvard University, and Cornell University. Genicot acknowledges support from the National Science Foundation under grant SES-1851758. This paper was edited by Emir Kamenica.

Electronically published November 17, 2021

Journal of Political Economy, volume 130, number 1, January 2022.

https://doi.org/10.1086/717041

to oneself (for recent examples, see Marmaros and Sacerdote 2006; Halberstam and Knight 2016; for a survey, see McPherson, Smith-Lovin, and Cook 2001).[1] This literature has long highlighted the role of preferences in explaining why people associate and bond more with those who are similar to themselves. However, there is an important distinction to be made between individuals caring about the innate identity of their friends—their type—and individuals caring about the conduct of their friends—their behavior. There are settings in which individuals care about others' type—for example, gender, ethnicity, religion, or sexual orientation (see Currarini, Jackson, and Pin 2009; Currarini, Matheson, and Vega-Redondo 2016). However, in many settings, individuals may care more about others' behaviors: their religious practices, the political opinions they voice, how conservatively they dress, or even how "white" they act (see Berman 2000; Lagunoff 2001; Austen-Smith and Fryer 2005; Carvalho 2013). This paper focuses on such preferences.

In this paper, individuals interact in social networks. There is a benefit from forming a link with others, what the paper calls "making friends," which includes actual friendships but also mutual help or working relationships. The common underlying factor among all these relationships is that they require social interactions and that people care about the behavior of individuals with whom they interact. Each individual has an ideal behavior—their *identity*. A person's utility depends on how close her own conduct, as well as the conduct of those in her network, is to her identity. An individual's *tolerance level* is the largest deviation from her identity that she can accept in a friend. Individuals can differ in their tolerance. Before forming social networks, individuals choose their behavior anticipating the need to fit in during the subsequent process of network formation. *Compromising* by adopting a conduct that differs from one's innate identity comes at a cost. The only motive for compromising in this model is to make friends. For instance, a religious person may abstain from displaying religious symbols in an environment with many nonbelievers, while an atheist may silence his criticism of the church to be accepted by devout friends. Conservative college students may express views that are more liberal than their own ideology on college campuses (Barker 1994; Braghieri 2020), as liberals might keep their leftist views to themselves in conservative small towns.

This paper explores the limits of compromise. In equilibrium, an individual's ability to compromise is bounded by the difference between her tolerance level and the lowest tolerance level in society. Consequently, if

[1] "Birds of a feather flock together" is attributed to Burton ([1651] 1927), but scholars have described the pattern starting in the antiquity: "we love those who are like themselves" (Aristotle's *Rhetoric*) or "similarity begets friendship" (Plato's *Laws*).

all individuals have the same tolerance levels, then compromise is impossible. In such a setting, everyone would choose their preferred behavior, and individuals would form links with each other if and only if they tolerate each other's ideal points.

The key insight is the incentive to only minimally compromise for others. During the network formation stage, individuals will be friends with anyone whose conduct is acceptable. Since compromise is costly, this gives individuals the incentive to compromise just enough to be accepted by their set of friends. Any individual who compromises in equilibrium must minimally compromise for at least one other individual. With symmetric tolerance levels, this other individual must also compromise. Moreover, this other individual's compromise must be larger for her to be a worthwhile friend. Each compromise implies the existence of another person who compromises by even more, which is not indefinitely possible.

Heterogeneity in tolerance is thus necessary for compromise. Introducing less tolerant individuals allows for the possibility of compromise in equilibrium. With differences in tolerance levels, tolerant individuals may value friendships with relatively intolerant individuals, even when the latter do not compromise, resulting in unilateral compromise. Moreover, relatively intolerant individuals can enable more tolerant individuals of the same tolerance level to compromise for each other. I show that these relatively intolerant "bridge" individuals are essential for reciprocal compromise.

The joint distribution of tolerance levels and identities matters for the patterns of compromise across different tolerance levels. If individuals with more extreme identities are systematically less tolerant, it can be shown that reciprocal compromise is impossible in equilibrium. In this case, behaviors tend to be more polarized than ideologies. In contrast, more tolerance at the extremes encourages a more connected society. These findings are related to the work of Esteban and Ray (1999; see sec. V), who study differences in the distribution of observable behaviors and the underlying distribution of extremist types in a model of conflict.

The next section discusses the related literature. Section III formalizes the model described above. Section IV describes simple examples that provide the intuition for the main results. Section V contains this paper's key results. Section VI discusses some of the assumptions of the model and implications of the results. Finally, section VII notes concluding thoughts.

II. Literature

This work relates to several strands of existing literature. First, this paper contributes to a growing theoretical literature on homophily in the formation of social networks (Gilles and Johnson 2000; Currarini, Jackson, and Pin 2009; Golub and Jackson 2012; Currarini, Matheson, and Vega-Redondo 2016; Gauer and Landwehr 2016; Iijima and Kamada 2017; Jackson 2019).

Like these papers, I assume that individuals prefer to associate with similar others. A crucial difference is that I assume that individuals care about others' conduct as opposed to their identity.

Second, this work pertains to the general framework of Akerlof and Kranton (2000), where individuals have identity-based payoffs that depend on their own actions and on others' actions and where they can modify their identity at a cost. However, in contrast to models of endogenous identity (e.g., Shayo 2009; Grossman and Helpman 2018) or intergenerational transmission of identity (e.g., Bisin and Verdier 2001; Carvalho 2016), individuals in this paper cannot change their innate identities. They can, at some cost, compromise and adopt a behavior that differs from their innate identity. Specifically, this paper provides a model where an endogenous social network may provide incentives for individuals to compromise.

Third, the analysis here is closely related to the literature on conformism. In models of conformism, individuals have a preferred conduct but also care about the extent to which their actions differ from those chosen by other members of their social group to gain esteem, conform to the norm (Jones 1984; Glaeser, Sacerdote, and Scheinkman 1996; Cervellati, Esteban, and Kranich 2010; Richter and Rubinstein 2021), or avoid revealing their true type (Bernheim 1994; Kuran 1995). In particular, a number of papers study conformism in exogenously given networks (Bisin, Horst, and Ozgur 2006; Patacchini and Zenou 2012; Ozgur, Bisin, and Bramoulle 2018), though only Boucher (2016) studies the formation of networks in the presence of conformism. This paper is significantly different in that individuals evaluate others' behaviors in comparison to their own ideals as opposed to a societal norm. Section VI.E discusses this crucial distinction and its implications.

A different strand of the literature in evolutionary economics studies the stability of norms of tolerance, where tolerance is defined as the willingness to interact with individuals from another group (Muldoon, Borgida, and Cuffaro 2012; Cerqueti, Correani, and Garofalo 2013). These papers use replicator dynamics to study why intolerance may persist even when being tolerant presents an advantage in trading environments. In contrast, this paper treats tolerance as an immutable attribute of an individual's preferences.

Finally, this paper contributes to the literature on diversity and social capital (see Dasgupta and Serageldin 1999; Putnam 2000; Portes and Vickstrom 2011). According to Putnam (2000), there is an important distinction between bridging (inclusive) and bonding (exclusive) social capital. Bonding social capital networks are inward looking and tend to reinforce exclusive identities and homogenous groups. On the other hand, bridging social capital networks are outward looking and include people across "diverse social cleavages." This paper highlights the role of intolerant individuals as bridging agents.

III. Premises of the Model

A. *Individuals and Preferences*

Consider a population of size N indexed by $i \in \{1, 2, \dots, N\}$. Each individual i has an ideal point $\iota_i \in [0, 1]$ that denotes her identity. Here identity represents an individual's ideal code of conduct and is immutable. In contrast, individuals select a code of conduct—a behavior, $a_i \in [0, 1]$. For example, an individual's religious identity would represent their intrinsic faith in god, whereas their religious behavior could range from carrying a rosary to burning Bibles. Another example is the expression of political opinions. In this case, one's identity corresponds to one's ideology, ranging from the extreme left to the extreme right, while one's behavior consists in the expressed political opinions.

I study a model of social interaction, where individuals first choose their conduct and then form their social network. Think of an individual's conduct as a set of public behaviors or expressed attitudes that individuals choose to convey their image or "adopted identity" within society.[2] Then individuals form their social network, by establishing links that allow them to interact with each other. I call these links *friendships* and the individuals in this network *friends*. These links may represent actual friendships but also work collaborations or other mutually beneficial relationships. Crucially, these relationships require interactions, and people care about the behavior of individuals with whom they interact on a regular basis.

While individuals value the benefits that flow from their friendships, they also care about the behavior of their friends. I assume that individuals judge all behaviors in comparison to their identity. In particular, an individual's utility strictly decreases in the Euclidean distance between a behavior and her ideal point. Individual i derives a benefit $v_i(|\iota_i - a_j|)$ from a link to an individual j with behavior a_j. The link has a strictly positive value when a_j is equal to ι_i, but this value strictly decreases as a_j differs from i's identity. Individuals may differ in the benefits they derive from a friendship and in their tolerance for behaviors that differ from their ideal point.

Finally, individuals also care about their own conduct. Choosing a behavior that departs from one's own identity comes at a cost $c(|\iota_i - a_i|)$ that strictly increases as a_i departs from ι_i.

Consider an individual i with ideal point ι_i, behavior a_i, and links to a friendship set S. Denoting the profile of behaviors of individual i's friends by $\mathbf{a}_S$, individual i's utility can be expressed as

$$u_i(a_i, \mathbf{a}_S) = U_i\left(\sum_{j \in S} v_i(|\iota_i - a_j|)\right) - c_i(|\iota_i - a_i|), \tag{1}$$

[2] This paper considers only one choice of conduct, but one could imagine individuals selecting different conducts for separate "societies" (say, their work and their neighborhood).

where the benefit from friendship v_i is continuous, strictly decreasing in the distance between ι_i and a_j with $v_i(0)$ strictly positive but finite; U_i is strictly increasing with $U_i(0) = 0$; and c_i is continuous, strictly increasing with $c_i(0) = 0$.

Note that this specification is quite flexible. It allows for the utility from friendships to be additively separable if U_i is linear but also allows for non-additive benefits, such as diminishing return to friendships if U_i is strictly concave (as in Currarini, Jackson, and Pin 2009).

A key attribute of the chosen specification in (1) is the notion of a *tolerance level*, which represents the largest deviation from individual i's ideal point that is still consistent with a valuable friendship. Formally, we define individual i's tolerance level t_i by $v_i(t_i) = 0$; it is the largest tolerable distance.[3]

Person i is happy to be friends with person j as long as j's behavior is within a distance t_i of i's ideal point: $|\iota_i - a_j| \leq t_i$. An individual's tolerance level reflects both how much the individual values a friendship and the extent to which she dislikes departures from her ideal behavior.

Let us define i's *tolerance window* as $\omega_i \equiv [\iota_i - t_i, \iota + t_i]$ and say that a belongs to i's tolerance window if $a \in \omega_i$. Figure 1 illustrates these concepts.

Individual i is said to compromise if her chosen behavior differs from her ideal point $|\iota_i - a_i| > 0$.

B. Timing

This model is a two-stage game with full information. In the first stage, individuals choose a code of conduct. In the second stage, they form their social network by choosing their friends. When people choose their behavior in the first stage, they anticipate the friendship networks they might form in the second stage. They may want to compromise in order to be accepted—to "belong."

This timing enables individuals to commit to a conduct—a public image—before interacting with others. Note that if individuals could revise their behavior after forming their network, they would never compromise and would always select their ideal points. But this alternative timing assumes away the possibility for individuals to stop interacting with anyone who does not bring them utility. A more realistic assumption that would be equivalent to our model would be to have individuals choose their behavior and possible friends simultaneously but simply permit individuals to revise their choice of friends afterward. This gives individuals the possibility to stop interacting with anyone who does not bring them utility. As a result, like in the original timing, no one in equilibrium has an incentive to adjust their action knowing that others can then respond to this by adjusting their friendships.

[3] This definition of tolerance may be closer to what Murphy (1997) calls toleration, though there is no consensus on the distinction between tolerance and toleration.

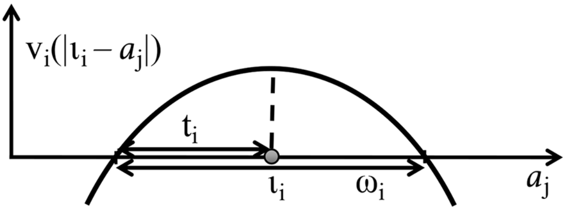

FIG. 1.—Tolerance.

In this simple model, friendships are not dynamic. But one could extend the model to allow for the formation of relationship capital among some pairs. This would make some friendships more valuable over time, thereby increasing one's tolerance toward these friends.

C. Network Formation

I assume that link formation is costless. As a result, the process of network formation is trivial. Given a vector of behaviors $\mathbf{a}$ in the population, individual i is happy to form a link with individual j if and only if $a_j \in \omega_i$.

Following most of the network literature, I consider networks that are pairwise stable. Jackson and Wolinsky (1996) defined a network to be pairwise stable if (i) no player would be better off if they severed one of their links and (ii) no pair of players would both benefit from adding a link that is not already in the network.

I assume that if both players are indifferent, they form a link. Then, for any given profile of behaviors $\mathbf{a}$, there is a unique pairwise stable graph G such that individuals i and j are friends if and only if their behaviors lie in each other's tolerance windows: $g_{ij} = 1$ if and only if $a_j \in \omega_i$ and $a_i \in \omega_j$.

IV. Examples

This section provides intuition for the main results of the paper through some simple examples. For the purpose of these examples, I assume the following linear functional form:

$$u_i(a_i, \mathbf{a}_{S(i)}) \sum_{j \in S(i)} \left[F_i - b_i \left|\iota_i - a_j\right|\right] - c_i|\iota_i - a_i|, \qquad (2)$$

where F_i represents the intrinsic value of a friendship for i, while c_i and b_i capture her aversion to behaviors that deviate from her ideal for herself and for others, respectively. In this setting, individual i's tolerance level is given by

$$t_i = \frac{F_i}{b_i}.$$

Note that this simple expression captures well the fact that an individual's tolerance level depends on the benefit that she derives from a friendship as well as her dislike of differences. The more she values friendships, the more she is willing to befriend individuals who differ from her ideal point.

A. No Compromise with Homogeneity

Assume that identities differ but $t_i = t$ for all i. When two individuals have the same level of tolerance, then either both individuals' identities belong to each other's tolerance window $\iota_i \in \omega_j$ and $\iota_j \in \omega_i$ or both individuals' identities lie outside of each other's tolerance window $\iota_i \notin \omega_j$ and $\iota_j \notin \omega_i$.

In figure 2*A*, individuals i and j are sufficiently tolerant such that their ideal points already belong to the other's tolerance window. They therefore have no incentive to compromise and can be friends in spite of their differences.

In contrast, figure 2*B* illustrates a situation where i and j do not belong to each other's windows though their tolerance windows do overlap. Thus, the only way for them to become friends is for both to compromise. Since there is no incentive to unilaterally compromise, it is easy to see why there exists an equilibrium without compromise. However, what may be surprising is that this no-compromise equilibrium is unique. This paper shows why this unique no-compromise equilibrium is not simply a coordination failure.

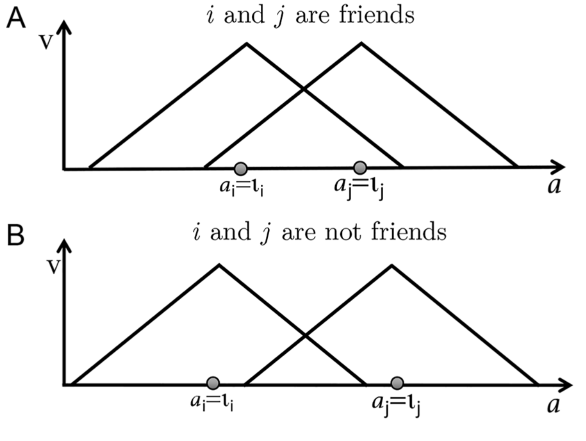

FIG. 2.—Homogeneity in tolerance.

Assume that i and j compromised for each other as illustrated in figure 3*A*. This cannot be an equilibrium, as individual i would benefit from reducing her compromise as shown in figure 3*B* (and so would individual j). Since compromise is costly, individuals have an incentive to minimally compromise for others—to do the least possible in order to be accepted. However, this implies that their friendship is not worth compromising for in the first place. Figure 3*C* shows that if i minimally compromises for j, her friendship is not valuable to individual j. The latter therefore had little incentive to compromise herself. It follows that individuals i and j in figure 2*B* cannot compromise for each other. In both of the above cases, individuals choose their preferred actions and are friends only if they belong to each other's tolerance window when tolerance levels are the same.

B. Heterogeneity Enables Compromise

To see how heterogeneity in tolerance levels enables compromise, take two individuals j and k who differ in tolerance levels. If j is more tolerant than k, then we can have k's ideal point belonging to j's tolerance window but not vice versa—$\iota_k \in \omega_j$ but $\iota_j \notin \omega_k$, as illustrated in figure 4*A*. Individual j values being friends with k even if k does not compromise. If j values this friendship enough so as to be willing to bear the disutility from compromising, j and k will be friends. If she compromises, j would clearly choose the smallest compromise needed to be friends with k: the behavior a_j in ω_k that is

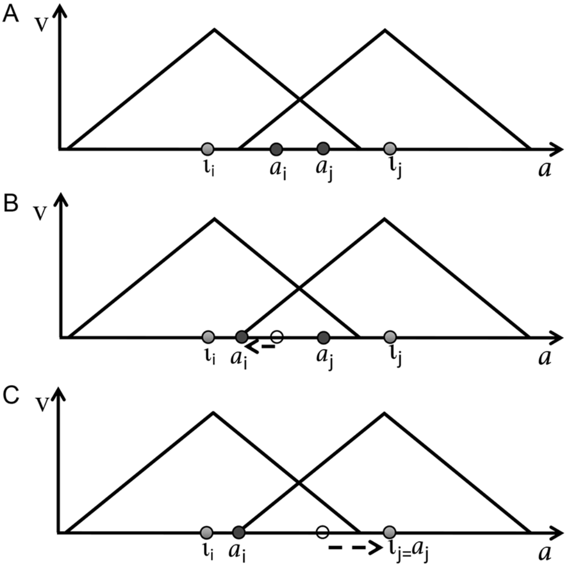

FIG. 3.—No compromise in equilibrium.

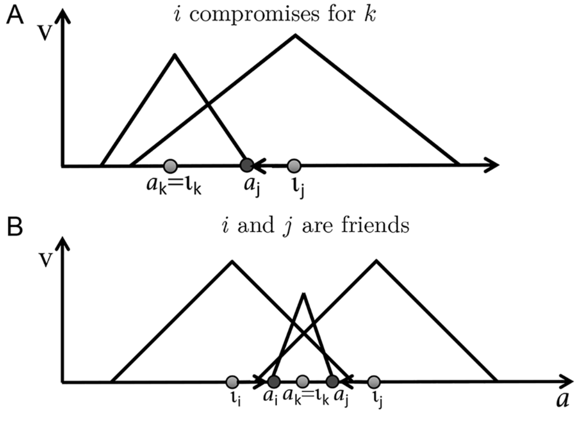

FIG. 4.—Compromising for an intolerant person.

the closest possible to ι_j as shown in figure 4*A*. Hence, j compromises and befriends k if

$$F_j - b_j|\iota_j - \iota_k| - c_j|\iota_j - a_j| \geq 0. \quad (3)$$

It naturally follows that the presence of less tolerant individuals can allow more tolerant individuals to compromise and become friends. Consider the example in figure 2 where i and j have the same tolerance levels. Now suppose that there were a less tolerant individual k placed between them so that $\omega_k \subseteq \omega_i$, $w_k \subseteq \omega_j$, $\iota_i \notin \omega_j$, and $\iota_j \notin \omega_i$, as in figure 4*B*. If i compromises to be acceptable to k, she becomes attractive to j as well and vice versa. Let l and r denote the left and right thresholds of k's tolerance window, respectively. There is an equilibrium where $a_i = l$, $a_k = \iota_k$, and $a_j = r$, and all three individuals are friends if the following two inequalities hold:

$$[F_i - b_i|\iota_k - \iota_i|] + [F_i - b_i|r - \iota_i|] \geq c_i|l - \iota_i|$$

and

$$F_i - b_i|\iota_k - \iota_i| \geq c_i|l - (\iota_j - t_j)|.$$

The first inequality requires the overall value of the compromise to be positive: the value of the friendships with j and k exceeds the cost of compromise. The second inequality guarantees that i prefers choosing l and being friends with both i and j, rather than choosing the left extremity of j's tolerance window, $\iota_j - t_j$, and being friends only with j. Both these constraints are satisfied for a sufficiently low cost of compromise. This equilibrium is illustrated in figure 4*B*.

Example.—Assume that i and j have ideal points $\iota_i = 0.2$ and $\iota_j = 0.8$. They are otherwise identical, with $b_i = b_j = 1$ and $F_i = F_j = 0.5$. Now consider individual k, who has an ideal position in between, $\iota_k = 0.5$, and is less tolerant than i and j, with $b_k = 5$ and $F_k = 0.5$. Individuals i and j have the same disutility from their own actions deviating from their ideal points, $c_i = c_j = 1.1$. Interestingly, i would not compromise for k alone, but $a_i = l = 0.4$, $a_k = \iota_k$, and $a_j = r = 0.6$ is an equilibrium.

C. *Compromise Builds on Compromise*

We have just seen examples of tolerant individuals who minimally compromise for a less tolerant person. But it is possible for a tolerant individual to minimally compromise for another tolerant individual in equilibrium. Indeed, there is a compromise multiplier encouraging reciprocal compromise.

This is illustrated in figure 5, where i and j are high tolerance, while k is a low-tolerance person. In this example, k does not compromise. Individual j minimally compromises for k, thereby becoming valuable to i. Individual i compromises not just for k but also for j. A complete network is achieved.

V. Main Results

After these illustrative examples, I return to the most general setup: agents $i \in \{1, 2, \dots, N\}$ with preferences represented by (1). Let t_i denote the tolerance level of individual i, and let $\underline{t}$ and $\overline{t}$ denote the lowest and highest tolerance levels in society, respectively.

A. *Limits to Compromise*

This section characterizes the limits to compromise. Before proceeding, I will introduce some definitions that are useful for the proof and the rest of the paper.

Definition 1. j is valuable to i if $a_j \in \,]\iota_i - t_i, \iota_i + t_i[$.

Definition 2. i compromises for j if $|\iota_i - \iota_j| > t_j \geq |\iota_j - a_i|$.

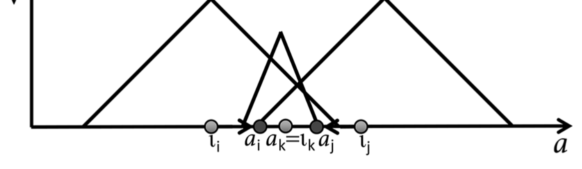

Fig. 5.—Compromise builds on compromise.

DEFINITION 3. i minimally compromises for j if i compromises for j and $a_i = \operatorname{argmin}_{a_i \in \omega_j} |a_i - \iota_i|$.

In words, an individual i is said to compromise for j if i's ideal point is outside of j's tolerance window but her chosen behavior is inside j's tolerance window. Individual i minimally compromises for j if she compromises for j while deviating as little as possible from her ideal point.

We now turn to our first proposition.

PROPOSITION 1. An individual i with tolerance t_i never compromises by more than $t_i - \underline{t}$ in equilibrium.

The detailed proof of this proposition is in the appendix, but the intuition for the proof is simple. Assume that there is an equilibrium vector of actions such that individual i violates the claim in the proposition. Lemma 1 tells us that if i compromises at all then there must exist a valuable individual j (in the sense of definition 1) for whom i minimally compromises (in the sense of definition 3). If not, i could compromise a bit less and still keep the same number of friends. Lemma 3 then uses the fact that j is valuable to i while i is just acceptable to j to show that the difference between j's compromise and his tolerance must be larger than the difference between i's compromise and her tolerance. This implies that j must also be compromising by more than $t_j - \underline{t}$. Repeating this argument would imply an infinite sequence of individuals $m = \{i, j, ...\}$ along which would be ever increasing, a clear contradiction to the finite number of individuals in the society.[4]

The incentive to minimally compromise for others drives this result. There is strategic complementarity between individuals' decisions to choose behaviors within each other's tolerance window but not to compromise beyond the minimum needed to be acceptable.

Proposition 1 has several implications. Since $\bar{t} \geq t_i$, proposition 1 implies an upper limit to the compromise that can be achieved in a society:

COROLLARY 1. Individual compromise cannot exceed $\bar{t} - \underline{t}$ in equilibrium.

Another straightforward but powerful corollary of proposition 1 is that there cannot be any compromise when all individuals have the same tolerance levels (though the underlying utility functions could differ). When all tolerance levels are the same, only individuals whose ideal points belong to each other's tolerance window can be friends in equilibrium. This result generalizes the example of section IV.A.

COROLLARY 2. If all individuals have the same tolerance, $t_i = t$, for all i,

1. compromise is not possible;
2. i and j are friends if and only if $i \in \omega_j$ and $j \in \omega_i$.

[4] Proposition 1 also holds with a continuum of agents, though the proof differs.

It also follows directly from proposition 1 that individuals with the lowest tolerance level, $t_i = \underline{t}$, do not compromise.

COROLLARY 3. The most intolerant individuals never compromise in equilibrium.

Finally, the bounds on compromise imply a maximal distance between the ideal points of any two friends (see the proof in the appendix).

COROLLARY 4. In equilibrium, $|\iota_i - \iota_j| \leq t_i + t_j - \underline{t}$ for all pairs $ij \in G$.

B. *Role of Bridges*

The previous section showed that equilibrium compromise is bounded and that no compromise is possible with homogenous tolerance levels. In contrast, the examples of sections IV.B and IV.C demonstrate that heterogeneity in tolerance levels makes compromise possible.

Now, clearly the presence of a variety of individuals leads to multiple equilibria. A relatively tolerant individual may compromise in one direction or the other depending on the behavior of others. This means that observed behaviors or the views expressed in a society can change rapidly, from moderate to extreme positions for instance, with little change in the underlying distribution of preferences. This would be consistent with the rise and rapid fall in political correctness observed over the last few years (Pew Research Center 2019). Also related is the evidence of "contingent extremists": individuals with long-held extreme ideology who will express their support (and vote) for extremist parties only when they think that a sufficient number of others share their views (see Jakli 2020).

Characterizing the full set of equilibria in a given setting is difficult, though proposition 1 and its corollaries help by limiting the range of possible behaviors. Proposition 2 helps by identifying the necessity of a bridge individual in between any two individuals who compromise for each other.

DEFINITION 4. Individuals i and j reciprocally compromise if i compromises for j and j compromises for i.

PROPOSITION 2. If i and j (with ideal points $\iota_i \leq \iota_j$) reciprocally compromise, then there exists an individual k such that

1. $\iota_k \in (\omega_i \cap \omega_j)$ and
2. $\iota_k - t_k$ or $\iota_k + t_k \in (\omega_i \cap \omega_j)$.

Proposition 2 tells us that there must be a bridge individual between any pair of individuals who reciprocally compromise: an individual whose ideal point and one edge lie in the intersection of the pair's tolerance windows. Note that this bridge individual is necessarily strictly less tolerant than the individuals she is bridging. Taken together with the relationship between identity and tolerance, this proposition has important consequences for whether compromise will contribute to the polarization or cohesion of a society.

C. *Extremism and Polarization*

There is some evidence pointing toward intolerance and extremism being correlated (see van Prooijen et al. 2015; van Prooijen and Krouwel 2017). A direct implication of proposition 2 is that if there is systematically more intolerance at the extremes, there cannot be any reciprocal compromise, leading to polarization.

Consider a form of "single-peaked" relationship between identity and tolerance:

Assumption [T]. There is a deterministic mapping $T:[0,1]\rightarrow\mathbb{R}_+$ from identity to tolerance such that, for any $\iota_i < \iota_j < \iota_k$, if $T(\iota_j) < T(\iota_k)$ then $T(\iota_i) \leq T(\iota_j)$ and if $T(\iota_i) > T(\iota_j)$ then $T(\iota_j) \geq T(\iota_k)$.

This mapping allows for two scenarios. $T(\iota)$ either is monotonic or increases then decreases over the interval. In the latter scenario, we say that the extremes are less tolerant.

Proposition 3. Under assumption [T], in equilibrium,

1. there is no reciprocal compromise;
2. if $ij \in G$ and $t_i \geq t_j$, then $\iota_j \in \omega_i$; and
3. when the extremes are less tolerant, behaviors are more polarized than identities.

The proof is simple. Assumption [T] implies that for any two individuals i and j with $\iota_i < \iota_j$, no individual k with $\iota_k \in [\iota_i, \iota_j]$ can be strictly less tolerant than both i and j. Without the possibility of a bridge between them, proposition 2 tells us that i and j cannot reciprocally compromise. So if two individuals are friends, it must be that one person's ideal point lies within the other one's tolerance window. This implies that more tolerant individuals necessarily compromise toward less tolerant individuals.

If tolerance is monotonic in identity, all compromise will be in one direction: toward the least tolerant individual. If tolerance first increases and then decreases along the interval, as illustrated in figure 6, then behaviors tend to be more polarized than identities in the following sense: there are identities ι_l and ι_r ($\iota_l < \iota_r$) such that all individuals with identities to the left of ι_l can compromise only to the left, while all individuals to the right of ι_r can compromise only to the right.[5] Individuals with identities

[5] Assume that T is continuous, and let $\bar{\iota}$ be (one of) the identity associated with the highest level of tolerance. The threshold ι_l is the maximum between zero and the largest $\iota \leq \bar{\iota}$ such that there exists $\iota' \geq \bar{\iota}$ with $T(\iota) = T(\iota') = |\iota' - \iota|$ or such that $T(\iota) = T(1)$. Similarly, let ι_r be the minimum between one and the smallest $\iota \geq \bar{\iota}$ such that there exists $\iota' \leq \bar{\iota}$ with $T(\iota) = T(\iota') = |\iota' - \iota|$ or such that $T(\iota) = T(0)$.

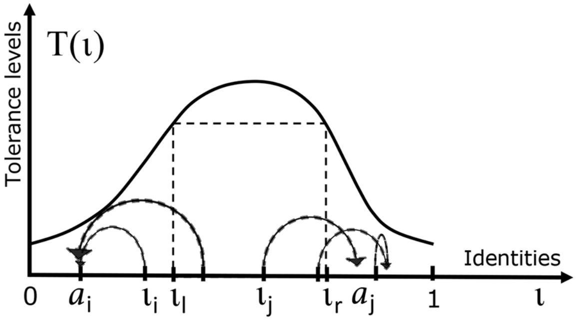

FIG. 6.—Intolerance at the extremes and polarization.

between ι_l and ι_r may compromise in one direction or the other depending on the circumstances or even the particular equilibrium considered. In figure 6, individual i's identity is to the left of ι_l. If i compromises, her behavior will be to the left of her ideal point. Figure 6 shows j compromising to the right. But since ι_j is to the left of ι_r, there could exist different equilibria or settings where j would compromise to the left.

Until recently, the literature had found more evidence of higher intolerance among conservatives, but recent research finds evidence for intolerance on both sides of the political spectrum (Brandt et al. 2014; Crawford and Pilanski 2014). Newspaper articles frequently document the rise of intolerance among liberals and have coined the term "cancel culture" to refer to this phenomenon.[6] Simultaneously, polarization in the United States has dramatically increased over the last few years (Boxell, Gentzkow, and Shapiro 2017). These patterns could easily be understood in the context of this model.

D. Centrism and Cohesion

Though we often associate intolerance with extremism, this need not be the case. For example, in postwar western Europe, there were calls for a strong commitment to moderation. In his very influential book *The Open Society and Its Enemies*, Popper ([1945] 2012) argued against the toleration of extremist ideas (see also Walzer 1997). Such a commitment could translate into lesser tolerance among moderates. Assumption [U] formally defines this premise, and proposition 4 derives its consequences.

ASSUMPTION [U]. There exists a deterministic mapping $T:[0, 1] \rightarrow \mathbb{R}_+$ from identity to tolerance, and T is decreasing and then increasing.

[6] See the *New York Times* podcast about cancel culture: https://www.nytimes.com/2020/08/10/podcasts/the-daily/cancel-culture.html.

Figure 7 provides an example of such a mapping. Let m be an individual with the lowest level of tolerance $\underline{t}$, and let $l = \max_j \iota_j - t_j$ and $r = \min_j \iota_j + t_j$.

PROPOSITION 4. Under assumption [U], behaviors are less polarized than identities. If $[\iota_m - \underline{t}, \iota_m]$ or $[\iota_m, \iota_m + \underline{t}] \subset \text{int}(\cap_{i \neq m} \omega_i)$, then for sufficiently low costs of compromise, equilibrium behaviors are

$$a_i = \begin{cases} l & \text{if } \iota_i < l, \\ \iota_i & \text{if } \iota_i \in [l, r], \\ r & \text{if } \iota_i > r, \end{cases} \tag{4}$$

and the equilibrium network is complete.

Proposition 4 tells us that assumption [U] encourages moderation and cohesion. Any individual who compromises will do so toward m. Intuitively, suppose that an individual to the left of m compromised to the left. Such a compromise would be worthwhile only if it were reciprocated by another individual to the left of m, which would be impossible without a bridge (in the sense of proposition 2) in between them (the formal proof is in the appendix).

When individual m acts as a bridge, moderating behaviors listed in (4) characterize the equilibrium when the cost of compromise is sufficiently low. Figure 7 illustrates this with an example. Here m and $r = m + t_m \subset \text{int}(\cap_{i \neq m} \omega_i)$. When the cost of compromise is sufficiently low, individuals to the right of r would unilaterally compromise to r. Thus, individuals to the left of l would compromise to l (in fig. 7, $l = \iota_j - t_j$) to be friends with everyone else. Individuals with identities that lie within $[l, r]$ have no reason to compromise. As a result, observed behaviors are more moderate in comparison with identities.

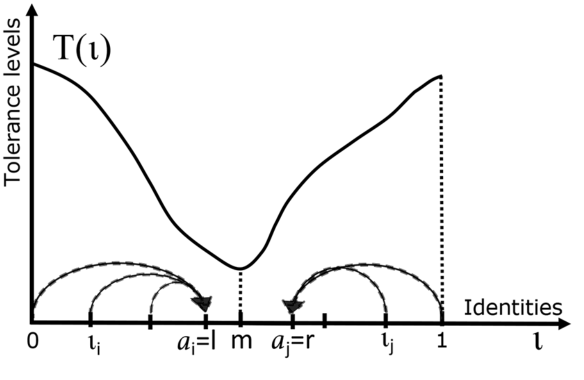

FIG. 7.—Intolerance at the center and cohesion.

Note that sufficiently low costs of compromise may implicitly assume individuals to be more sensitive to others' behaviors than their own. This suggests that some hypocrisy—judging others more harshly than oneself—may be required to have a cohesive society.

VI. Discussion

This section further discusses some of the implications and the assumptions of the model.

A. *Welfare*

It should not come as a surprise that there may be suboptimal equilibria with too little compromise. Take two individuals i and j with the same tolerance level, $t_i = t_j = t$. While corollary 5(1) tells us that no compromise is possible in equilibrium, it is easy to show that compromise could benefit them both.

A necessary condition for compromise between two individuals i and j to be optimal is that, for both i and j, the gain in their utility as the other moves toward them must be greater than their loss from moving away from their ideal position.

In the case of linear utility, as in the examples of section IV, this occurs when individuals are more sensitive to others' behaviors than their own: b_i higher than c_i in equation (2). While we probably all know of people who are stricter with others than they are with themselves—judging others for their lack of religiosity, their overindulgence with their children, or their promiscuity, while always having good reasons to allow themselves not to attend religious services, give in to their kids, or enter a new relationship—it seems unlikely to be the case for all.

With concave utility functions, an increase in the distance between an individual's ideal point and her friend's behavior causes more disutility the larger this distance is to begin with. In this case, compromise can be optimal even when the disutility from a person's own deviations from her ideal point is greater than or similar to the disutility she experiences from a friend's deviation from her ideal. This is illustrated in the following example.

Example.—Take the case of two individuals i and j with the same quadratic preferences:

$$u_k(a_k, \mathbf{a}_S) = \sum_{l \in S} [F - b|\iota_k - a_l|^2] - c|\iota_k - a_k|^2, \qquad c \geq b > 0. \tag{5}$$

Let $\lambda = c/b (\geq 1)$. In such a setting, the Pareto optimum is given by

$$a_i^* = \frac{\iota_j + \lambda\iota_i}{1 + \lambda} \qquad \text{and} \qquad a_j^* = \frac{\iota_i + \lambda\iota_j}{1 + \lambda}$$

when $|\iota_i - \iota_j| < \sqrt{((1+\lambda)/\lambda)}t$. In particular, if $b = c$, then meeting in the middle is optimal for i and j as long as they are sufficiently tolerant: $|\iota_i - \iota_j| < \sqrt{2}t$.

B. *Nonmonotonicity of Payoffs in $\underline{t}$*

We can build on the previous example to see that the payoffs to the tolerant individuals in a society are nonmonotonic in the tolerance levels of less tolerant individuals. Consider two individuals, i and j, whose identities lie just outside each other's tolerance window but for whom reciprocal compromise would be optimal. Now let us introduce a relatively less tolerant person, k, with an ideal point located between ι_i and ι_j.

If k is almost as tolerant as i and j, as illustrated in the left panel of figure 8, no one compromises. Individuals i and k are within each other's tolerance window and are friends. The same is true for individuals j and k. As we reduce k's tolerance level, we first reach a point where k's tolerance window lies just outside of i and k's ideal point. If compromise is not too costly, i and j will minimally compromise for k, and further reductions in k's tolerance level would decrease i and j's payoff. However, as k becomes even less tolerant, there is a point at which k's entire tolerance window lies within i and j. At that point, compromising for k allows i and j to be friends with each other (as shown in the right panel of fig. 8). In a given range, k brings the Nash equilibrium closer to the Pareto optimum and increases their utility by being less tolerant.

Figure 9 plots the tolerance level of individual k and the corresponding equilibrium payoff of i and j for the following parameter values: $\iota_i = 0.1$, $\iota_k = 0.5$, $\iota_j = 0.9$, and preferences given by (5) with $F = 0.5$ and $c = 1$. Assume that $b_i = b_j = 1$, which corresponds to a tolerance of $\bar{t} = 0.7$ for i and j. Let us steadily decrease the tolerance level for k by increasing b_k from one to high values (reading the graph from right to left). Initially, a decrease in k's tolerance level hurts i and j as it forces them to compromise to remain friends with k until the point where compromising for k allows them to all be friends. Further decreases in the tolerance of k benefits them.

Despite our focus on tolerance levels, it is worth noting that it is not the only characteristic of the preferences that determines compromise. Take two agents i and j such that $\iota_j \in \omega_i$ and $\iota_i \notin \omega_j$. Whether i wants to

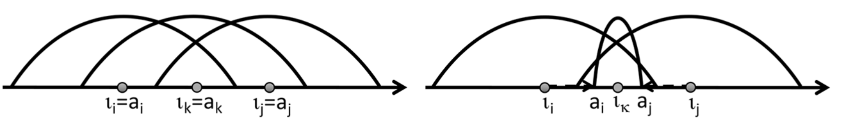

FIG. 8.—Nonmonotonicity of payoff in $\underline{t}$.

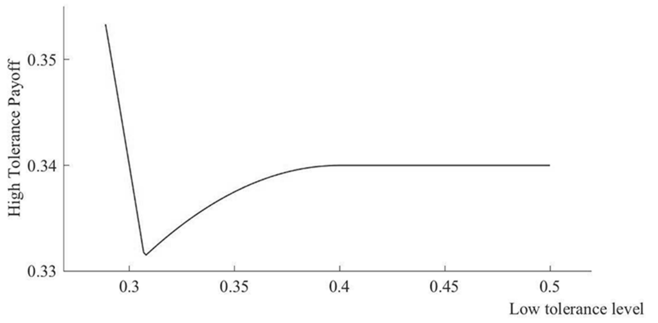

FIG. 9.—Nonmonotonicity of payoff in $\underline{t}$.

compromise for j will depend on the level of utility that he derives from the friendship and the cost of compromise.

C. *Uncertainty*

One might think that some uncertainty about the exact tolerance levels of the others may help reciprocal compromise. This section shows that uncertainty by itself is not enough.

Consider two individuals i and j with identities $\iota_i < \iota_j$ (such as the two individuals of sec. VI.A). Now assume that the utility of friendship and associated tolerance level t are private information. Suppose that there is a finite set $\mathcal{S}$ of types of agents, where a type is the utility of friendship. For each type s, the individual utility of friendship v_s is continuous, strictly decreasing, and bounded and has an associated tolerance level $t_s \in [\underline{t}, \overline{t}]$. The distribution from which these preferences are drawn is common knowledge, but the realized values are private information.

A perfect Bayesian equilibrium consists of strategy profiles $\boldsymbol{\mu}$ designating strategies for each type of individual i and j, $\mu_i : \mathcal{S} \to [0, 1]$ and $\mu_j : \mathcal{S} \to [0, 1]$, and inference functions ϕ_i and ϕ_j, characterizing the beliefs of individuals i and j at each information set, such that $\boldsymbol{\mu}$ is sequentially rational for each player given $\phi = (\phi_i, \phi_j)$ and $\boldsymbol{\phi}$ is derived from $\boldsymbol{\mu}$ using Bayes' rule whenever possible.

First, inference does not affect the network formation stage. Given behaviors $\mathbf{a} = (a_i, a_j)$, individuals i and j of any type are friends if and only if their behaviors lie within each other's (type-specific) tolerance windows. Let $1_{\mathbf{a},s_i,s_j}$ be the indicator that a link will be formed given $\mathbf{a}$ and the individual types.

Taking $\boldsymbol{\mu}_j$ as given, the utility that individual i of type s_i maximizes by choosing a_i is

$$u_{is_i}(a_i) = E_{s_j} 1_{\mathbf{a},s_i,s_j} v_{is_i}\left(\left|\iota_i - \mu_j(s_j)\right|\right) - c_i(|\iota_i - a_i|). \qquad (6)$$

OBSERVATION 1. If $\bar{t} < |\iota_j - \iota_i|$ or if $\underline{t} \geq |\iota_j - \iota_i|$, then the perfect Bayesian Nash equilibrium exhibits no compromise. Compromise can occur only if $\underline{t} < \iota_j - \iota_i < \bar{t}$.

Despite the uncertainty regarding the other person's tolerance level, observation 1 proves that there is no compromise if $\bar{t} < |\iota_j - \iota_i|$. As without uncertainty, the result is driven by the incentive to minimally compromise for others, which in turn reduces the value of the potential friendship. Because compromise is necessary from (some of the types of) both individuals, i and j are unable to become friends.

If $\underline{t} \geq |\iota_j - \iota_i|$, then i and j will be friends and do not need to compromise in order to do so. Individuals i and j may compromise for each other only if $\bar{t} > |\iota_j - \iota_i|$ and $\underline{t} < |\iota_j - \iota_i|$. The logic is similar to the role of heterogeneity in section IV.B. Compromise is sparked by the willingness of a relatively high-tolerance individual to compromise for a relatively low-tolerance type, even when the latter does not compromise.

D. *Externalities*

The assumption that individuals care only about the behavior of their friends may not seem plausible in all contexts. The behavior of those outside our friendship networks may affect us. The results of this paper hold even if we relax this assumption by allowing individuals to care about the behavior of all individuals. As long as individuals care more about those in their direct network than others, the results in the paper hold.

E. *Changing Identity*

In this paper, individuals' identities are set. Individuals may adopt behaviors that differ from their ideal points to fit in, but this changes neither who they are nor how they judge their friends.

If instead we allowed individuals to pick a new identity a_i and use this new identity to judge others' behaviors, we would get a model of conformism in social networks (see Boucher 2016). In such a "moving" model, an individual i with an adopted identity a_i would get a benefit $v_i(|a_i - a_j|)$ from his friendship with a person with adopted identity a_j. This modeling assumption might work well in some settings. For instance, identities could represent physical locations and individuals could choose their addresses to be close to others. The proximity of the new addresses would be the relevant distance in such a setting. Similarly, when kids choose extracurricular activities, they may have preferences over the activities themselves but also care about participating in the same activities as their friends (the context studied in Boucher 2016).

However, when we think of identities as normative values, individuals may not be able to change their identities at will. Although people's views on certain issues may evolve over the course of their lives, there is evidence that fundamental values and orientations are set early in life and may be hard to change. This is especially true in the context of religious values or political attitudes (Hamberg 1991; Lai et al. 2016; Ghitza, Gelman, and Auerbach 2019). A famous longitudinal study of women who studied at Bennington College (Alwin, Cohen, and Newcomb 1991, 64) finds that "through late childhood and early adolescence, attitudes are relatively malleable with the potential for dramatic change possible in late adolescence or early adulthood. But greater stability sets in at some early point, and attitudes tend to be increasingly persistent as people age." At the same time, there is some evidence supporting the contact hypothesis (Pettigrew and Tropp 2006) according to which a person's social network may affect their identity over time. Even assuming the premise, the model presented in this paper would characterize individual decision-making in the short- to medium-term better than a moving model.

Notice that the predictions from the two models are very different. In contrast to this paper, a moving model would predict that intolerant individuals would compromise the most, a potentially testable implication.

VII. Conclusion

This paper presents a model of compromise in social networks. Individuals' identities characterize their preferred conduct for themselves and for others. An individual's tolerance level is the largest deviation from her identity that she finds acceptable in a friend's behavior. The need to be acceptable to others can incentivize individuals to compromise and deviate from their preferred conduct.

The paper characterizes the limits to compromise and shows that the bounds to compromise decrease in the tolerance level of the least tolerant individual. When all individuals have the same tolerance level, there cannot be any compromise in equilibrium. In contrast, heterogeneity in tolerance levels enables compromise. The paper finds that relatively intolerant bridge individuals play a key role in reciprocated compromise. The paper explores how different joint distribution of tolerance and identity can lead to polarization or moderation in a society. Finally, it shows how welfare and compromise are nonmonotonic in the lowest level of tolerance in society.

Note that the distinction this paper makes between identity and behavior has implications for measures of diversity and tolerance in a society. Looking at the identity of the members of a person's social network may overestimate the tolerance exhibited by the person. The distance between a person's identity and her friends' behaviors would likely tell us more about her tolerance.

In future research, it would be interesting to study a dynamic version of the current model in which individual behaviors affect the intergenerational transmission of identity.[7] This model can also be used to study the effect of new opportunities for friendships owing to population growth or social media.

Appendix

Proofs

LEMMA 1. If i compromises in equilibrium, then the set of individuals who are valuable to i and for whom i minimally compromises, X_i, is nonempty.

Proof. Let S_i be i's set of friends in equilibrium. Denote as $V_i \subseteq S_i$ the subset of friends who are valuable to her (in the sense of definition 1) and as $X_i \subseteq V_i$ the subset of these for whom she minimally compromises (in the sense of definition 3). Suppose that X_i is empty: there is no j valuable to i and for whom i minimally compromises. Then, there is a_i' such that $|\iota_i - a_i'| < |\iota_i - a_i|$ and $|\iota_j - a_i'| \leq t_j$ for all $j \in X_i$. Clearly, a_i' represents a profitable deviation: i could compromise less and keep all her valuable links. QED

LEMMA 2. If i minimally compromises for j, then $|\iota_j - a_i| = t_j$ and $|\iota_i - \iota_j| = |\iota_i - a_i| + t_j$.

Proof. By definition, if i minimally compromises for j, then $|\iota_i - \iota_j| > t_j$ and $a_i = \text{argmin}_{a_i \in \omega_j} |\iota_i - a_i|$. The claim follows directly from these two facts. QED

LEMMA 3. If i minimally compromises for j and j is valuable to i, $|\iota_j - a_j| - t_j > |\iota_i - a_i| - t_i$.

Proof. Lemma 2 tells us that if i minimally compromises for j, then

$$|\iota_i - \iota_j| = |\iota_i - a_i| + t_j. \tag{A1}$$

If j is valuable to i, then $|\iota_i - a_j| < t_i$. Given that

$$|\iota_i - \iota_j| - |\iota_j - a_j| \leq |\iota_i - a_j|,$$

it follows that

$$|\iota_i - \iota_j| < |\iota_j - a_j| + t_i. \tag{A2}$$

Together, inequalities (A1) and (A2) imply that

$$|\iota_i - a_i| - t_i < |\iota_j - a_j| - t_j. \tag{A3}$$

QED

Proof of proposition 1. Assume that the proposition does not hold. That is, there is an equilibrium vector of behaviors **a** and a (nonempty) set of individuals

[7] I am grateful to an anonymous referee for raising this point.

K so that $|\iota_k - a_k| > t_k - \underline{t}$ for all $k \in K$. Let i be an individual in this set K with the largest $|\iota_i - a_i| - t_i$.

Since i belongs to K, $|\iota_i - a_i| > t_i - \underline{t}$, and therefore i compromises. Lemma 1 implies that there exists at least one individual j who is valuable to i and for whom i minimally compromises. Lemma 3 tells us that

$$|\iota_i - a_i| - t_i < |\iota_j - a_j| - t_j,$$

which contradicts the selection of i. QED

Proof of corollary 4.

$$|i_i - \iota_j| \leq |\iota_i - a_j| + |\iota_j - a_j|. \tag{A4}$$

It follows from proposition 1 that $|\iota_j - a_j| \leq t_j - \underline{t}$. Also, if i and j are friends, then $|a_j - \iota_i| \leq t_i$. Using these two inequalities in (A4) tells us that

$$|\iota_j - \iota_i| \leq t_i + t_j - \underline{t}.$$

LEMMA 4. If $ij \in G$, then a_i and $a_j \in (\omega_i \cap \omega_j)$.

Proof. For j to accept i's friendship, it must be that $a_i \in \omega_j$; $a_i \in \omega_i$ follows directly from proposition 1. QED

LEMMA 5. If i minimally compromises for j and j is valuable to i, $|\iota_j - a_i| = t_j > |\iota_j - a_j| > |\iota_j - \iota_i| - t_i$.

Proof. If i minimally compromises for j, then $|\iota_j - a_i| = t_j$. By capping compromise, proposition 1 implies that $|\iota_j - a_j| < t_j$. Finally, if i values j, then $|\iota_j - a_j| > |\iota_j - \iota_i| - t_i$ (see eq. [A2] in the proof of lemma 3). QED

Proof of proposition 2. Assume that the converse is true. Then, there must exist individuals i and j who reciprocally compromise but no intermediary individual k with identity ι_k in $(\omega_i \cap \omega_j)$ and an extremity, either $(\iota_k - t_k)$ or $(\iota_k + t_k)$, in $(\omega_i \cap \omega_j)$. Without loss of generality, assume that $\iota_i < \iota_j$ so that $\iota_i < a_i$.

The proof consists in showing that this would imply an infinite sequence of distinct individuals $\{x_m\}$ for $m \in \{1, 2, ...\}$ who compromise for each other—an impossibility. To simplify notation, we denote ideal points along the sequence ι_m, their tolerance t_m, and their actions a_m for $m \in \{1, 2, ...\}$. The sequence that we construct originates at i—that is, $x_1 = i$.

Since x_1 compromises, there exists a valuable individual x_2 with $\iota_2 > \iota_1$ (which may or may not be j) for whom x_1 minimally compromises (lemma 1). Since $a_1 > \iota_1$, this tells us that $a_1 = \iota_2 - t_2$. Lemma 4 implies that $\iota_2 - t_2 \in (\omega_1 \cap \omega_2)$. It must then be that $\iota_2 \notin (\omega_1 \cap \omega_2)$ (otherwise we would have a contradiction since $(\omega_1 \cap \omega_2) \subseteq (\omega_i \cap \omega_j)$). It follows that x_2 compromises and $\iota_2 > a_2$.

Take any $m \geq 2$. It follows from lemma 1 that if x_m compromises, with $a_m > (<) \iota_m$ for odd (even) m, then there exists a valuable x_{m+1} for whom m minimally compromises. Lemma 4 implies that $a_m = \iota_{m+1} - t_{m+1} (\iota_{m+1} + t_{m+1}) \in (\omega_m \cap \omega_{m+1})$ if m is odd (even). Lemma 5 tells us that $a_m = \iota_{m+1} - t_{m+1} < a_{m+1} < \iota_m + t_m = a_{m-1} (a_m = \iota_{m+1} + t_{m+1} > a_{m+1} > \iota_m - t_m = a_{m-1})$ if m is odd (even). This implies two things. First, x_{m+1} is a distinct individual from all x_n for $n \leq m$. Second, $(\omega_m \cap \omega_{m+1}) \subset (\omega_{m-1} \cap \omega_m)$ and therefore $(\omega_m \cap \omega_{m+1}) \subset (\omega_i \cap \omega_j)$. It must then be that $\iota_{m+1} \notin (\omega_m \cap \omega_{m+1})$ (otherwise we would have a contradiction) so that x_{m+1} compromises as well. Repeating this argument implies the existence of an infinite sequence of distinct individuals—a contradiction to the finite number of individuals. QED

Proof of proposition 4. Assume that an individual i to the left of m (m being an individual with the lowest level of tolerance) chooses $a_i < \iota_i$. Lemma 1 tells us that i must minimally compromise for some $j < i$, and assumption $[U]$ implies that $t_j \geq t_i$. It follows that i and j must reciprocally compromise. However, assumption $[U]$ rules out the possibility of the bridge that proposition 2 showed to be necessary for there to be reciprocal compromise. This leads to a contradiction. The same logic applies to show that no individual to the right of m could choose a behavior to the right of their identity. Hence, behaviors are less polarized than identities.

Let $l = \max_j \iota_j - t_j$ and $r = \min_j \iota_j + t_j$. Now assume that $[\iota_m - \underline{t}, \iota_m] \subset \operatorname{int}(\cap_{i \neq m} \omega_i)$ (a similar argument applies if $[\iota_m, \iota_m + \underline{t}] \subset \operatorname{int}(\cap_{i \neq m} \omega_i)$), which implies that $l = \iota_m - \underline{t}$.

Consider the following choice of behaviors:

$$a_i = \begin{cases} l & \text{if } \iota_i < l, \\ \iota_i & \text{if } \iota_i \in [l, r], \\ r & \text{if } \iota_i > r. \end{cases}$$

As long as the overall utility from friendships U_i is not bounded, a sufficient condition for any individual i to not want to deviate is that

$$c_i(|l - \iota_i|) < U_i((n-2)v_i(\iota_i) + v_i(m)) - U_i((n-2)v_i(\iota_i)) \quad \text{if} \quad \iota_i < l, \tag{A5}$$

$$c_i(|\iota_i - r|) < U_i((n-2)v_i(\iota_i) + v_i(l)) - U_i((n-2)v_i(\iota_i)) \quad \text{if} \quad \iota_i > r. \tag{A6}$$

Intuitively, the cost of compromise should be low enough that compromise would be worth it just for one additional friend with the least desirable equilibrium behavior.

It is also easy to check that if the inequalities in (A5) hold, this the only equilibrium. For any individual i with $\iota_i < l$, choosing l dominates any other choice, even if choosing l as opposed to ι_i serves in gaining the friendship of only a single individual (individual m). Given that, anyone to the right of r would want to unilaterally deviate to l even if it is just to gain the friendship of one individual whose behavior is at least l. Individuals between l and r have no reason to compromise. QED

Proof of observation 1. If $\underline{t} \geq |\iota_j - \iota_i|$, then the claim is trivial. Any type of agent who compromises in equilibrium would increase her utility by compromising less, as she does not need to compromise to be tolerable to the other.

Consider the situation where $\bar{t} < |\iota_j - \iota_i|$. The claim is that there is no compromise in equilibrium. Assume that the claim is incorrect. Then there must be some type of individual who compromises. Among all types of individuals who compromise, select i of type t_{s_i} to be the type of agent (or one of them if there are multiple) with the largest compromise minus tolerance $|\iota_i - \mu_i(s_i)| - t_{s_i}$. Since i compromises, the previously stated logic applies: type t_{s_i} must be minimally compromising for some type s_j of agent j, $|\iota_j - \mu_i(s_i)| = t_{s_j}$, and that agent must be valuable to i, $|\iota_i - \mu_j(s_j)| < t_{s_i}$. This implies that $t_{s_j} + |\iota_i - \mu_j(s_j)| < t_{s_i} + |\iota_j - \mu_i(s_i)|$. Since $\iota_i \leq \mu_i(s_i)$, $\mu_j(s_j) \leq \iota_j$, we have that $|\iota_i - \mu_j(s_j)| = |\iota_j - \iota_i| - |\iota_j - \mu_j(s_j)|$ and $|\iota_j - \mu_i(s_i)| = |\iota_j - \iota_i| - |\iota_i - \mu_i(s_i)|$. Using these equalities in the previous inequality yields $t_{s_j} + |\iota_j - \iota_i| - |\iota_j - \mu_j(s_j)| < t_{s_i} + |\iota_j - \iota_i| - |\iota_i - \mu_i(s_i)|$. Rewriting the latter gives $|\iota_i - \mu_i(s_i)| - t_{s_i} < |\iota_j - \mu_j(s_j)| - t_{s_j}$. Since j must compromise as well to be valuable, it contradicts the selection of i. QED

References

Akerlof, George A., and Rachel E. Kranton. 2000. "Economics and Identity." *Q.J.E.* 115 (3): 715–53.

Alwin, Duane F., Ronald L. Cohen, and Theodore M. Newcomb. 1991. *Political Attitudes over the Life Span: The Bennington Women after Fifty Years*. Madison: Univ. Wisconsin Press.

Austen-Smith, David, and Roland G. Fryer. 2005. "An Economic Analysis of 'Acting White.'" *Q.J.E.* 120 (2): 551–83.

Barker, K. 1994. "To Be PC or Not to Be? A Social Psychological Inquiry into Political Correctness." *J. Soc. Behavior and Personality* 9:271–81.

Berman, E. 2000. "Sect, Subsidy, and Sacrifice: An Economist's View of Ultra-Orthodox Jews." *Q.J.E.* 115 (3): 905–53.

Bernheim, B. Douglas. 1994. "A Theory of Conformity." *J.P.E.* 102 (5): 841–77.

Bisin, Alberto, Ulrich Horst, and Onur Ozgur. 2006. "Rational Expectations Equilibria of Economies with Local Interactions." *J. Econ. Theory* 127 (1): 74–116.

Bisin, Alberto, and Thierry Verdier. 2001. "The Economics of Cultural Transmission and the Dynamics of Preferences." *J. Econ. Theory* 97 (2): 298–319.

Boucher, Vincent. 2016. "Conformism and Self-Selection in Social Networks." *J. Public Econ.* 136 (C): 30–44.

Boxell, Levi, Matthew Gentzkow, and Jesse M. Shapiro. 2017. "Is the Internet Causing Political Polarization? Evidence from Demographics." Working Paper no. 23258, NBER, Cambridge, MA.

Braghieri, Luca. 2020. "Social Image, Information, and Political Correctness." Report, Stanford Univ., Stanford, CA.

Brandt, Mark J., Christine Reyna, John R. Chambers, Jarret T. Crawford, and Geoffrey Wetherell. 2014. "The Ideological-Conflict Hypothesis: Intolerance among Both Liberals and Conservatives." *Current Directions Psychological Sci.* 23 (1): 27–34.

Burton, Robert. (1651) 1927. *The Anatomy of Melancholy*. New York: Farrar & Rinehart.

Carvalho, Jean-Paul. 2013. "Veiling." *Q.J.E.* 128 (1): 337–70.

———. 2016. "Identity-Based Organizations." *A.E.R.* 106 (5): 410–14.

Cerqueti, Roy, Luca Correani, and Giuseppe Garofalo. 2013. "Economic Interactions and Social Tolerance: A Dynamic Perspective." *Econ. Letters* 120 (3): 458–63.

Cervellati, Matteo, Joan Esteban, and Laurence Kranich. 2010. "Work Values, Endogenous Sentiments Redistribution." *J. Public Econ.* 94 (9/10): 612–27.

Crawford, Jarret T., and Jane M. Pilanski. 2014. "Political Intolerance, Right and Left." *Polit. Psychology* 35 (6): 841–51.

Crick, Bernard. (1963) 1971. *Political Theory and Practice*. London: Allen Lane.

Currarini, Sergio, Matthew O. Jackson, and Paolo Pin. 2009. "An Economic Model of Friendship: Homophily, Minorities, and Segregation." *Econometrica* 77 (4): 1003–45.

Currarini, Sergio, Jesse Matheson, and Fernando Vega-Redondo. 2016. "A Simple Model of Homophily in Social Networks." *European Econ. Rev.* 90 (C): 18–39.

Dasgupta, Partha, and Ismail Serageldin. 1999. *Social Capital: A Multifaceted Perspective*. Washington, DC: World Bank.

Esteban, Joan, and Debraj Ray. 1999. "Conflict and Distribution." *J. Econ. Theory* 87 (2): 379–415.

Gauer, Florian, and Jakob Landwehr. 2016. "Continuous Homophily and Clustering in Random Networks." Working Paper no. 515, Center Math. Econ., Bielefeld Univ., Bielefeld, Germany.

Ghitza, Yair, Andrew Gelman, and Jonathan Auerbach. 2019. "The Great Society, Reagan's Revolution, and Generations of Presidential Voting." Working paper, Dept. Statis., Columbia Univ., New York.

Gilles, Robert P., and Cathleen Johnson. 2000. "Original Papers: Spatial Social Networks." *Rev. Econ. Design* 5 (3): 273–99.

Glaeser, Edward L., Bruce Sacerdote, and Jose A. Scheinkman. 1996. "Crime and Social Interactions." *Q.J.E.* 111 (2): 507–48.

Golub, Benjamin, and Matthew O. Jackson. 2012. "How Homophily Affects the Speed of Learning and Best-Response Dynamics." *Q.J.E.* 127 (3): 1287–338.

Grossman, Gene M., and Elhanan Helpman. 2018. "Identity Politics and Trade Policy." Working Paper no. 25348, NBER, Cambridge, MA.

Halberstam, Yosh, and Brian Knight. 2016. "Homophily, Group Size, and the Diffusion of Political Information in Social Networks: Evidence from Twitter." *J. Public Econ.* 143 (C): 73–88.

Hamberg, Eva M. 1991. "Stability and Change in Religious Beliefs, Practice, and Attitudes: A Swedish Panel Study." *J. Sci. Study Religion* 30 (1): 63–80.

Iijima, Ryota, and Yuichiro Kamada. 2017. "Social Distance and Network Structures." *Theoretical Econ.* 12 (2): 655–89.

Jackson, Matthew O. 2019. "The Friendship Paradox and Systematic Biases in Perceptions and Social Norms." *J.P.E.* 127 (2): 777–818.

Jackson, Matthew O., and Asher Wolinsky. 1996. "A Strategic Model of Social and Economic Networks." *J. Econ. Theory* 71 (1): 44–74.

Jakli, Laura Viktoria. 2020. "Contingent Extremism: How Perceptions of Party Popularity Activate Far Right Support." PhD diss., Univ. California, Berkeley.

Jones, Stephen R. G. 1984. *The Economics of Conformism.* New York: Basil Blackwell.

Kuran, Timur. 1995. *Private Truths, Public Lies: The Social Consequences of Preference Falsification.* Cambridge, MA: Harvard Univ. Press.

Lagunoff, R. 2001. "A Theory of Constitutional Standards and Civil Liberties." *Rev. Econ. Studies* 68:109–32.

Lai, Calvin K., Allison L. Skinner, Erin Cooley, et al. 2016. "Reducing Implicit Racial Preferences. II. Intervention Effectiveness across Time." *J. Experimental Psychology* 145 (8): 1001–16.

Marmaros, David, and Bruce Sacerdote. 2006. "How Do Friendships Form?" *Q.J.E.* 121 (1): 79–119.

McPherson, Miller, Lynn Smith-Lovin, and James M. Cook. 2001. "Birds of a Feather: Homophily in Social Networks." *Ann. Rev. Soc.* 27 (1): 415–44.

Muldoon, Ryan, Michael Borgida, and Michael Cuffaro. 2012. "The Conditions of Tolerance." *Polit. Philosophy and Econ.* 11 (3): 322–44.

Murphy, Andrew R. 1997. "Tolerance, Toleration, and the Liberal Tradition." *Polity* 29 (4): 593–623.

Ozgur, Onur, Alberto Bisin, and Yann Bramoulle. 2018. "Dynamic Linear Economies with Social Interactions." Working paper, Melbourne Bus. School.

Patacchini, Eleonora, and Yves Zenou. 2012. "Juvenile Delinquency and Conformism." *J. Law Econ. and Org.* 28 (1): 1–31.

Pettigrew, T. F., and L. R. Tropp. 2006. "A Meta-analytic Test of Intergroup Contact Theory." *J. Personality and Soc. Psychology* 90:751–83.

Pew Research Center. 2019. "Race in America." Report, Pew Res. Center, Washington, DC.

Popper, Karl. (1945) 2012. *The Open Society and Its Enemies.* London: Routledge.

Portes, Alejandro, and Erik Vickstrom. 2011. "Diversity, Social Capital, and Cohesion." *Ann. Rev. Soc.* 37:461–79.

Putnam, Robert D. 2000. *Bowling Alone: The Collapse and Revival of American Community.* New York: Simon & Schuster.

Richter, Michael, and Ariel Rubinstein. 2021. "Holding a Group Together: Non-Game Theory versus Game Theory." *Econ. J.* 131 (638): 2629–41.

Shayo, Moses. 2009. "A Model of Social Identity with an Application to Political Economy: Nation, Class, and Redistribution." *American Polit. Sci. Rev.* 103 (2): 147–74.

van Prooijen, Jan-Willem, and André P. M. Krouwel. 2017. "Extreme Political Beliefs Predict Dogmatic Intolerance." *Soc. Psychological and Personality Sci.* 8 (3): 292–300.

van Prooijen, Jan-Willem, André P. M. Krouwel, Max Boiten, and Lennart Eendebak. 2015. "Fear among the Extremes: How Political Ideology Predicts Negative Emotions and Outgroup Derogation." *Personality and Soc. Psychology Bull.* 41 (4): 485–97.

Walzer, Michael. 1997. *On Toleration.* New Haven, CT: Yale Univ. Press.