

2011

Network Structure, Interracial Contacts, and the Evolution of Social Norms

David J. Hendry

University of Illinois at Urbana-Champaign, hendry2@illinois.edu

Follow this and additional works at: http://opensiuc.lib.siu.edu/pnconfs_2011

Paper presented at the 4th Annual Political Networks Conference, University of Michigan, Ann Arbor, June 17-18.

Recommended Citation

Hendry, David J., "Network Structure, Interracial Contacts, and the Evolution of Social Norms" (2011). 2011. Paper 33.
http://opensiuc.lib.siu.edu/pnconfs_2011/33

This Article is brought to you for free and open access by the Conference Proceedings at OpenSIUC. It has been accepted for inclusion in 2011 by an authorized administrator of OpenSIUC. For more information, please contact opensiuc@lib.siu.edu.

Network Structure, Interracial Contacts, and the Evolution of Social Norms*

David J. Hendry[†]

[Draft copy of a work in progress. Please do not cite without permission of the author.]

Abstract

In this paper I explore the underlying mechanisms of the changes in public discourse with respect to the issue of racial equality that have been observed in the United States over the course of its history, with a particular focus on the changes that occurred in the latter half of the twentieth century. Specifically, I provide a formal model of social interactions in which agents are assigned to non-homophilic networks, are heterogeneous with respect to preferences for equality between the races, and have preferences both to express their true preferences and to not appear deviant from the group. In a series of numerical experiments, results indicate that the probability of a transition in norms from an equilibrium around inequality to an equilibrium around equality is increasing in the size of the minority population and decreasing in the size of groups to which individuals are assigned.

*Prepared for delivery at the 4th Annual Political Networks Conference, University of Michigan, Ann Arbor, June 17–18, 2011.

[†]Ph.D. student, Department of Political Science, University of Illinois at Urbana-Champaign. hendry2@illinois.edu; <https://netfiles.uiuc.edu/hendry2/www/>.

Introduction

In the twentieth century, the United States underwent a massive cultural shift. Norms of public discourse with respect to racial equality were completely reversed. The idea that the races could and should be considered social and political equals went from being a view rarely communicated outside of a small number of intimate social groups to being the only acceptable mode of public expression regarding race. This transformation has been deep, wide, and pervasive, effecting large-scale change in discourse across virtually all contexts of public communication. In most white social circles—in casual friendships, in school, in the workplace, in politics—whereas open expression of racial egalitarianism had once been a cause for social censure, eventually any statements of an inegalitarian nature would be met with strong disapproval, and could potentially cause irreparable damage to a reputation.

An area where the transition in norms is particularly evident is in speech related to political competition. In the latter years of the nineteenth century, the vast majority of candidates for political office in the United States could anticipate strong challenges from their political rivals and possible retribution from the voting public for espousing a personal view of racial equality. From the first years of the twentieth century to well past the mid-century mark (in most local areas), white politicians and members of the mass public alike would be subject to varying degrees of admonishment for advocating the advancement of blacks' social and political rights, the integration of public and private institutions, equal employment and housing opportunities for blacks, or any other measures that would breach the white-dominated racial hierarchy. By contrast, in the latter years of the twentieth century, the idea of racial equality was virtually universal in mainstream political discourse. Changing at varying rates across localities, and undergoing a process of progression and digression depending on the social and political climate of the day, it is readily apparent that the transition in norms with respect to socially acceptable speech regarding race is now complete. In modern American politics, any discourse even bordering on racial insensitivity is sure to be met with vociferous denunciation from a wide swath of powerful organized interests.

The scope of this cultural change is difficult to understate. The body politic of the United States in the twentieth century experienced a complete displacement of a prevailing social norm

that had helped to define and constrain both mass and elite rhetoric since the founding of the nation. From establishment, through the Civil War, through Reconstruction, and into the twentieth century, agreement on the inherent superiority of whites over African Americans was a necessary condition for entry into national, state, and most local politics. In particular, for the vast majority of white elected officials of the Southern United States, at least a nominal defense of racial separation was necessary until the 1960s, and, in some cases, beyond. The necessity of these political stances was a direct consequence of the then-prevailing social norm of racial inequality. In rapid historical fashion, however, the idea of racial equality, once held by only an extremely small minority of the population, grew in the latter half of the twentieth century to be the prevailing social norm, thereby displacing the norm of racial inequality.

That this shift in American culture has occurred is virtually incontrovertible. But given the status quo of American thought on matters of racial differences (and particularly of white racial attitudes) in the mid-twentieth century, the complete turnaround by the century's end certainly does not appear to be a priori self-evident. In fact, given the level of expressed white resistance to the idea of blacks as social and political equals, and the widespread willingness of elites to exploit and take on leadership roles in that resistance for electoral gains, the eventual entrenchment of the norm of racial equality is a puzzling phenomenon. The rapidity with which the change occurred is even more profound. And the implications for American politics have been far-reaching, permeating social interactions between and among elites and masses, changing the nature of political campaigns nationwide, altering the cleavage structure of major-party political competition, and helping to redefine the coalitions of electoral support for the major political parties.

How did this process of norm displacement occur and, just as importantly, how did it occur so quickly? In this paper, I address an important piece of the answer to these questions by situating the process of the displacement of norms of public speech within a broader story of cultural evolution and focusing on the theoretical impact of network structure on the motivations of individuals. In particular, I explore individual agents' sometimes competing desires to publicly express views that are consistent with their own internal preferences on the one hand, and to maintain social standing within their networks of social relationships on the other. Under a system of homophilic

selection into networks of relationships within a population that is relatively homogeneous with respect to internal preferences, desires for internal consistency and social acceptance are unlikely to be in tension. However, when institutional mechanisms construct non-homophilic networks of interactions, these two desires may frequently be opposed. When individuals are forced to interact with non-familiar others, and when exogenous events lead individuals to be relatively uncertain about the distribution of preferences in the population, it is reasonable to expect that publicly expressed opinions may not necessarily comport with privately held values. Individual idiosyncracies in estimating the distribution of preferences along with the desire to maintain social status may establish conditions conducive to the rapid transition from one norm of public expression to another. And importantly, the evolution of norms of public expression can occur independently of any changes in internal values.

The paper proceeds as follows. In the following section I briefly discuss the historical context of the norm change that is the basis of my theoretical perspective in order to provide a direct link between the substantive phenomenon and the more abstract model presented herein. Following this, I situate the individual-level mechanisms at play in the theory within the broader social scientific literature on cultural evolution and social pressures to develop insights into the individual-level mechanisms underlying key aspects of the dynamics of norms of public speech. Then, I present a simplified model that formalizes these individual-level mechanisms and some of the features of social structure within small-scale institutions that are thought to be important components to norms of public discourse on sensitive social issues. After this, I use the model to perform numerical experiments that demonstrate how the proposed individual-level mechanisms lead to population-level outcomes as a function of two key parameters. Following the numerical experiments, I offer concluding remarks.

The Norm of Racial Equality in Public Speech

As stated previously, the transition from a norm of racial inequality to a norm of racial equality in public speech has permeated all levels of society, having failed to penetrate only a small minority of closed social groups. In the most general terms, the process of change unfolded over a period of several hundreds of years, operating at a relatively slow overall rate for much of U.S. history, and

finally achieving a rapid pace in the latter half of the twentieth century (see, e.g., Mendelberg 2001, chs. 2 and 3 for a more detailed historical development). Accepting this description as an accurate characterization of the macro-level process of norm transition does not preclude the existence of significant variation in the pace of micro-level change. In fact, the society-wide displacement of the norm of racial inequality by the norm of racial equality was marked by different rates of piecemeal progression and digression across social groupings and geographic locations. Over time, the complex relationship between elite leadership, public opinion, and mass political action gradually advanced the social and political rights of African Americans and established an ever-changing set of legal codes and institutions that structure interracial interactions.

With respect to changes in institutions, of particular significance was the patchwork of public policies aimed at desegregating previously racially segregated institutions. Policies to combat racial discrimination in various aspects of society along with the eventual vigorous enforcement of judicial orders to desegregate schools led to a system of interracial contacts that had previously been unimaginable. A great deal of research has been conducted on how policies to desegregate have impacted the system of interracial interactions (e.g., Clotfelter 2004; Orfield 1969, 1983; Orfield and Monfort 1992) and how changes in interracial contacts have in turn impacted a variety of individual and social outcomes (e.g., Granovetter 1986; Patchen 1982). But no work has focused on the specific impact that changes in the degree and character of interracial contacts has had on the social acceptability of speech espousing racial inequality. In survey research on racial attitudes, the potential social desirability biases of direct measures are widely known and often mentioned, but a coherent theory or body of empirical evidence on the origin of such biases is absent.

Micro-level Cultural Transmission: Small Groups and Social Conformity

To develop a model of the micro-level process of norm development and evolution that has general appeal, and that can be readily applied to the case of the United States, I begin with the assumption that conformist transmission (also referred to as frequency-dependent bias) is adaptive under

a wide range of conditions. That is, rather than develop a specific model of the genetic development of conformist transmission in humans, I appeal to the literature with respect to its origins and dynamics (Boyd and Richerson 1985, 2005; Henrich 2001; Henrich and Boyd 1998, 2001) and begin at the starting point of conformist transmission as a pervasive human phenomenon. In brief, conformist transmission is the tendency of individuals in a population to acquire a particular cultural variant due to the relative frequency of that variant among the individuals' cultural "parents" (Boyd and Richerson 1985, p. 10). Boyd and Richerson (1985, 2005) and Henrich and Boyd (2001) note that conformist transmission is not identical to the concept of social conformity as it has typically been examined in social psychology (e.g., Asch 1951; Sherif 1935), and the difference is rooted in the motivation of individual agents. Specifically, conformist transmission results in the adoption of cultural traits that are popular because individuals use popularity as an indirect measure of a trait's value. Social conformity, on the other hand, results in the adoption of traits that are popular because individuals consciously try to avoid appearing deviant (Henrich and Boyd 2001). The former represents an internalization of a cultural variant, while the latter can be a simple change in observable behavior unaccompanied by internal changes.

If we consider norms of public speech to be cultural variants, the distinction between the concepts of conformist transmission and social conformity becomes murky. It is plausible to think that changes in public speech with respect to a social referent can and do occur without always being accompanied by an associated change in internal attitudes with respect to that same referent. Using the characterization of Henrich and Boyd (2001), such a process would suggest that patterns of public speech at the micro level could be the result of social conformity. At the macro level, however, large-scale changes in norms of public speech with respect to a social referent are the result of a broad internalization of what does and what does not constitute deviant behavior. Notice that I am *not* claiming that changes in norms of public speech must be accompanied by large-scale shifts in attitudes. Rather, I am explicitly maintaining a distinction between public speech and internal attitudes. Further, though it is clear that public speech and internal attitudes with respect to the same social referents must be related in some ways, it is also reasonable to expect that certain environmental conditions may cause the two to evolve on different—and sometimes even opposing—trajectories. Specifying those conditions theoretically

and deriving empirical implications are some of the goals of the present study.

The substantive situation of interest here is one in which a transition occurred to from one norm of public speech to another. I contend that a major factor in this transition was widespread social conformity at the level of the small group. Following Sherif (1935) and Asch (1951), a long line of research in social psychology and related fields has demonstrated that social conformity is in fact empirically verifiable in the laboratory. Sherif (1935, 1936) employed an optical illusion known as the autokinetic effect (Adams 1912), in which a stationary pinpoint of light is projected onto the wall of a darkened room, and to almost all human observers, the light appears to move. Employing confederates, Sherif demonstrated that experimental subjects were quite susceptible to the power of suggestion with respect to judgments about how far the light had traveled. That is, when subjects were placed in a group setting and asked to publicly express their estimation of how far the light had traveled after confederates had already done so, the judgments of subjects tended to fall in line with those of the confederates, even when the expressed judgments of confederates were wildly different than baselines established in a control condition. The Sherif studies demonstrated that when humans are asked to express judgments about relatively ambiguous phenomena, there is a strong tendency to look to the majority opinion as a rule of thumb for the appropriate response.

But what about relatively unambiguous phenomena? Working in the tradition of Sherif (1935, 1936), Asch (1951, 1952) examined social conformity in situations in which the judgment task had clear right and wrong answers. Specifically, subjects in the Asch studies were asked to express which among a set of drawn lines was identical in length to another drawn line. Line lengths were purposely drawn so that subjects would easily be able to differentiate their lengths, and subjects making the judgment in an isolated control condition rarely expressed the wrong answer. Again, subjects were placed in a group setting and asked to express their judgment out loud, this time following a group of confederates who expressed the wrong answer. Though a majority of subjects gave the correct response in this group condition, a surprisingly large proportion of subjects gave the same wrong answer as the group of confederates. Given little evidence that the subjects examined by Asch (1951, 1952) were unsure about the correct answer, the results point to a strong tendency for humans in public settings to conform to expectations about the appropriate

social response.

Scholars in a variety of fields have argued over how to interpret the results from these research programs, with some asserting that subjects are publicly complying with the group while privately maintaining a different viewpoint (e.g., Kiesler and Kiesler 1969), and others contending that subjects are unsure about the correct answer and are using the views of fellow group members to learn the the appropriate response (e.g., Thelen, Dollinger, and Kirkland 1979). Boyd and Richerson (1985), on the other hand, argue convincingly that one should leverage the difference in the level of ambiguity between the two decision-making tasks and that doing so may lead to different conclusions about the phenomenon exhibited by subjects. Specifically, since the judgment situation under study in the Asch experiments has an answer that is clearly correct, while the judgment situation from the Sherif experiments does not, it may be the case subjects who conform to the view of the confederates in the Asch experiments are exhibiting mere public compliance, while subjects who conform in the Sherif experiments are exhibiting a form of social learning.

While the results from the research programs spawned by Asch and Sherif are extremely useful as the largest and most prominent collection of empirical evidence on small-group conformity, it is not clear whether the type of conformity demonstrated has a direct connection to norms of public speech with respect to salient social referents (including race), which are the focus of the present study. In particular, the typical decision-making task under examination in these studies is highly arbitrary, and cannot reasonably be expected to carry any social significance for the subject outside of the laboratory. In one way, this can be seen as a strength of the experimental designs in that the researcher is able to eliminate subjects' preconceived notions about particular referents. In another way, however, if we consider those preconceptions to be worthy of study in their own right, and if we consider the clear qualitative difference between expressed judgments on arbitrary tasks and expressions of opinion on salient—and sometimes sensitive—attitudinal measures, existing work on social conformity leaves open the question of how small-group pressures might operate on norms of public speech regarding race. Since a variety of studies have provided evidence that expression of opinions is influenced by individual perceptions of the aggregate distribution of opinions in the society at large (e.g., Mutz 1998;

Noelle-Neumann 1974, 1984), and that the use of mass media information and opinion formation within social groups are often subject to the influence of opinion leaders (e.g., Katz 1957; Katz and Lazarsfeld 1955; Lazarsfeld, Berelson, and Gaudet 1948), it seems plausible to think that the type of social conformity demonstrated by Asch and Sherif with respect to arbitrary judgment tasks may also operate on salient attitudinal measures. But there currently is no body of empirical evidence to support this idea.

Of particular significance for norms of public speech with respect to race is the question of whether changes in norms of public speech are necessarily tied to changes in internal attitudes. I argue that norms of public speech with respect to race and internal attitudes with respect to race need not evolve at the same rate or even in the same direction. And in fact, early in a process of norm transition it is plausible to think that they do not. In a large body of work, Kuran (1987, 1989, 1990, 1991, 1995), has used formal logic, anecdotal accounts, and qualitative analysis of political events to argue convincingly that individual public compliance with a belief that is incongruent with internal attitudes is a theoretically plausible phenomenon. At the individual level, one can think of agents earning a type of expressive benefit by stating publicly their true preferences with respect to social referents and incurring a type of expressive cost by stating publicly a view that goes against the majority opinion (Kuran 1995). The primary mechanisms leading to these costs and benefits are negative and positive reinforcement by other group members. Under very general circumstances, it can be shown that norms of public compliance that run counter to some or all of a group's members' internal attitudes can persist if the costs of going against the group are sufficiently high, and that factors exogenous to the group can lead norms to transition.

The dynamics of norms of public speech at the level of the small group, as they have been presented here, can be summarized as follows. Norms of public discourse are developed and maintained through a process of social learning and cultural transmission. Over successive instances of publicly expressing one's opinion with respect to a particular social referent (e.g., race), individuals use the frequency of a publicly expressed value as an indirect measure of its appropriateness in social settings. Positive social reinforcement of the expressed view acts to strengthen the enculturation process. Additionally, individuals identify unacceptable public pronouncements by either witnessing or directly experiencing the social censure of unpopular expression. When dis-

approval of a particular expressed value is sufficiently frequent or intense, individuals internalize the unacceptability of that value as a social fact (Durkheim [1895] 1965). In this way, positive feedback and social censure act as reinforcement mechanisms that condition the attainment of particular cultural variants (Bandura 1977). However, internalizing the norms of appropriate social behavior does not necessarily indicate the internalization of congruent internal attitudes. Critically, norms of public expression can and often do run counter to privately held individual preferences. Public compliance with the expressed opinions of others with whom one interacts is considered to be a type of social conformity that results from a cultural evolutionary process of conformist transmission.

The Model

I now present a model of the public expression of attitudes with respect to race that accounts for the features of non-homophilic social interactions (including interracial interactions) structured by institutions and the preferences of individuals to conform to the behaviors of others with whom they are socially connected. The purpose is to capture the types of intimate interactions that would have been forced upon a particular population following a policy shift toward the desegregation of institutions. Following Akerlof (1980) and Jones (1984), I represent the social status of agents by the utility that they derive from behaving in accordance with others. But as in Bernheim (1994) and Kuran (1995), preferences for conformity must be balanced by preferences to exhibit behavior that is consistent with internal values.¹

Formally, let $N = \{1, \dots, n\}$ be a set of agents represented as the nodes of a graph, $g \in G(N)$, where $G(N)$ is the set of all undirected graphs on N . Agent $i \in N$ is characterized by a fixed, two-element vector (v_i, r_i) , where $v_i \in [0, 1]$ and $r_i \in \{0, 1\}$. Specifically, the social context to which the model is intended to apply is characterized by a racial hierarchy and racial separation. Parameter v_i represents agent i 's internal values with respect to the question of racial equality, such that values closer to 1 represent a preference for the maintenance of the racial hierarchy and

¹In other words, the balance is between conformity with others and consistency between internal values and expressed behavior. For a similar take on a different phenomenon, Bednar et al. (2010) and Page, Sander, and Schneider-Mizell (2007) use the same terminology of conformity and consistency to develop models of the balance between conformity with others and internal consistency across a variety of issue dimensions.

values closer to 0 represent a preference for empowerment of the less powerful group. Parameter r_i represents the racial group to which agent i belongs, with $r_i = 1$ indicating that agent i belongs to the more powerful group and $r_i = 0$ indicating that agent i belongs to the less powerful group.

Each agent in the population is forced to select some publicly observable behavior $b \in B \equiv \{0, 1\}$. Agents have intrinsic preferences over B represented by the payoff function $f(b_i, v_i)$, where $f(\cdot)$ achieves a maximum when $v_i = b_i$ and is assumed to be symmetric. In addition, agents have extrinsic preferences over B represented by the payoff function $h(b_i, \phi_i, g)$, where $h(\cdot)$ achieves a maximum at $b_i = \phi_i$ and is assumed to be symmetric, and where ϕ_i is a summary representation of agent i 's estimate of the behaviors of others in her neighborhood (to be explained below). Given the assumptions imposed on $f(\cdot)$ and $h(\cdot)$, a convenient representation is to let $f(b_i, v_i) = -(b_i - v_i)^2$ and $h(b_i, \phi_i, g) = -(b_i - \phi_i)^2$. Agent i chooses b_i to solve $\max_{b_i \in B} U(b_i, v_i, \phi_i, g)$, where

$$\begin{aligned} U(b_i, v_i, \phi_i, g) &= f(b_i, v_i) + h(b_i, \phi_i, g) \\ &= -(b_i - v_i)^2 - (b_i - \phi_i)^2. \end{aligned}$$

The form of $U(\cdot)$ shows that it is simply a convex combination of intrinsic and extrinsic payoffs.

For a specified number of time steps, agents are ordered according to their social status as determined by their extrinsic payoff in the previous period, and in this order each chooses an observable behavior. Let $S \equiv \sigma(N) = (s_1, \dots, s_n)$ be a permutation of N that lists agents in descending order according to their status. For example, suppose in a population of three agents that agent 3 has the highest status, agent 1 the second highest, and agent 2 the lowest. We would then have $S = (s_1, s_2, s_3) = (3, 1, 2)$. In a given time step, after each agent s_j chooses b_{s_j} her action is observable by all other agents in her neighborhood of direct connections. Therefore, any agent s_k such that $j < k$ and $s_j \in N_{s_k}(g)$ will be able to incorporate the observation of b_{s_j} into her estimate of the distribution of behaviors in the neighborhood, ϕ_{s_k} . A further assumption is imposed to capture a particular feature of interracial vs. intraracial interactions. That is, agents in the dominant group having internal preferences for the maintenance of racial inequality take into account only intraracial connections when considering how the distribution of behaviors will affect their social status. Agents in the dominant group having internal preferences for racial equality, on the other hand, take into account all of their connections, regardless of race. Formally, for

any graph $g \in G(N)$, let $N_i(g)$ denote the neighborhood of direct connections and $N'_i(g)$ the neighborhood of direct intraracial connections. Also, let m_i be the total number of agents in the direct neighborhood and m'_i be the total number of agents in the direct intraracial neighborhood of agent i . Then, at time step t , agent $i = s_k$ develops an estimate of the distribution of behaviors in her neighborhood according to the following rule:

$$\phi_{s_k,t} = \begin{cases} \frac{1}{m_{s_k}-1} \left[\left(\sum_{j < k: s_j \in N_{s_k}(g)} b_{s_j,t} \right) + \#\{j > k : s_j \in N_{s_k}(g)\} p_{s_k,t-1} \right], & \text{if } v_i < 0.5; \\ \frac{1}{m'_{s_k}-1} \left[\left(\sum_{j < k: s_j \in N'_{s_k}(g)} b_{s_j,t} \right) + \#\{j > k : s_j \in N'_{s_k}(g)\} p_{s_k,t-1} \right], & \text{if } v_i \geq 0.5; \end{cases}$$

where $p_{s_k,t-1}$ is the mean of observable behaviors in agent i 's neighborhood in the previous time period. Specifically,

$$p_i = \frac{1}{m_i - 1} \sum_{j \neq i: j \in N_i(g)} b_j.$$

Substantively, the form of ϕ_{s_k} shows that agents establish an estimate of the distribution of behaviors in the population by observing others who have gone before them in the current round, and combining this with the information from the entire neighborhood in the previous round.

After all agents in N choose a behavior to solve $\max_{b_i \in B} U(b_i, v_i, \phi_i, g)$, each receives an actual payoff of $U(b_i, v_i, p_i, g)$. That is, the extrinsic payoff actually received is based on the average behavior of other agents in agent i 's neighborhood of direct connections rather than the estimate employed to choose a behavior. Agents' extrinsic payoffs in time t then determine the order in which agents choose behavior in time $t + 1$, and so on. The operation of the model dynamics can be summarized as follows.

1. Define $N = \{1, \dots, n\}$ and choose $g \in G(N)$.
2. For each i in 1 through n , assign $v_i \in [0, 1]$, $b_i \in \{0, 1\}$, and an initial status, $q_0 \in [0, 1]$.
3. Define $S = \sigma(N) = (s_1, \dots, s_n)$ so that agents are placed in descending order according to their social status.
4. Agents s_1 through s_n calculate ϕ_{s_j} and choose b_{s_j} to solve $\max_{b \in B} U(b_i, v_i, \phi_i, g)$.
5. Agents receive payoff $U(b_i, v_i, p_i, g)$ and $h(b_i, p_i, g)$ becomes agent i 's current status.
6. Repeat steps 3 through 5.

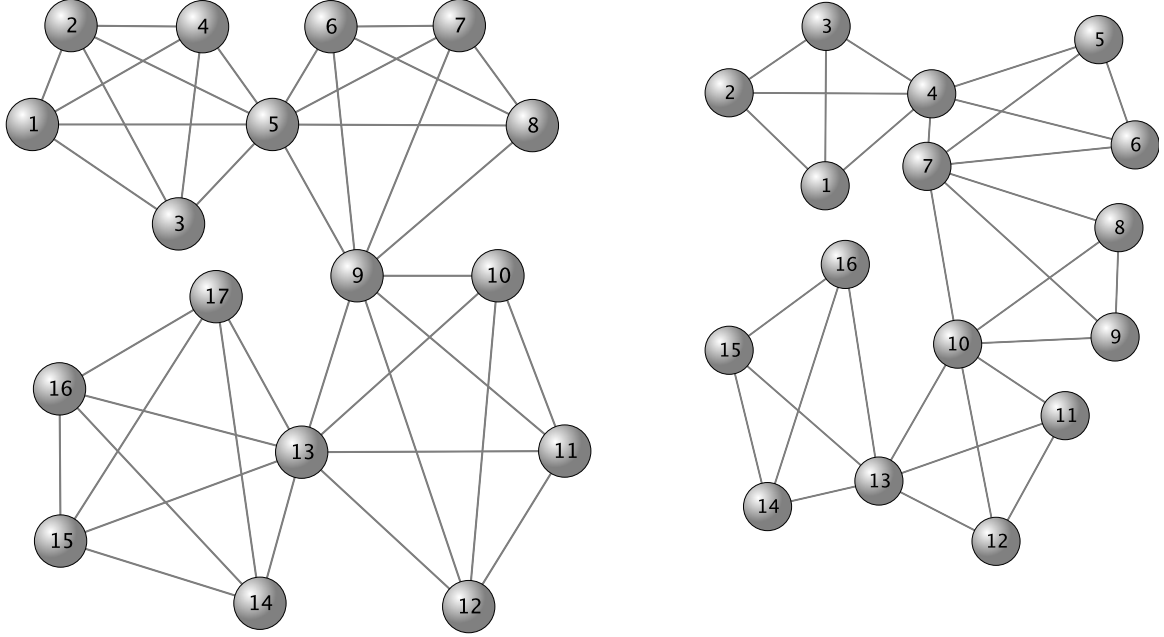


Figure 1: Example networks with $N = 17$ and $k = 5$ (left) and $N = 16$ and $k = 4$ (right).

Numerical Experiments

A set of numerical experiments were constructed to explore variation in some of the key features of the social context that are expected to impact the distribution of expressed behaviors in the population over time. Graphs were constructed to ensure the presence of cliques of varying size. Specifically, a base clique size, k , was chosen so that every agent in the population was randomly assigned to at least one clique of size k , and one individual in each clique was assigned to two cliques of size k . The resulting networks of social interactions take the form of those presented in Figure 1. These contrived examples depicts two societies: one of 17 agents assigned to base cliques of size 5, with agents 5, 9, and 13 having been assigned to two cliques; the other of 16 agents assigned to base cliques of size 4, with agents 4, 7, 10, and 13 having been assigned to two cliques. The purpose of random assignment is to capture non-homophilic interactions and the purpose of random assignment to cliques is to capture an important feature of schools, workplaces, and other social settings that structure social interactions. That is, institutions composed of many actors will structure interactions around tasks that must be performed, and those tasks will call for certain interactions to be intimate and others to range from casual to non-existent. Additionally, certain agents will have more social connections than others.

All of the experiments described herein used a population of 200 agents. Since the intention is to capture the effects of movement from segregated to desegregated institutions, the population was initially composed only of members from the dominant (majority) racial group. Agents' internal values, represented by the parameter v , were drawn from a truncated (0,1) normal distribution with a mean of 0.55 and a standard deviation of 0.1. The purpose of this choice was to slightly skew the internal values of the population toward the maintenance of racial hierarchy, but at the same time ensure that the bulk of agents fall into the moderate range of the distribution. The model was then allowed to iterate for several periods so that a behavioral norm among the dominant group and the associated social status for individuals could develop endogenously. Then after several periods, a small number of agents from the less powerful (minority) group were introduced to replace agents of the dominant group. In other words, the total number of agents was unchanged with the introduction of members of the less powerful group. The model was then allowed to iterate a specified number of time steps, and the network of connections was reformed every several periods.

Figure 2 presents a summary of results from a set of experiments in which a small percentage of agents from the less powerful group and the base clique size were varied. Agents from the less powerful group were introduced after 5 periods and the network of connections was reformed every 5 periods thereafter. Plots present the first 100 iterations. For each percentage of agents from the less powerful group and base cliques size combination, 100 simulations were performed. Lines on the plots indicate the average behavior of the society for each simulation.

Scanning Figure 2 across rows illustrates the effect of increasing the base clique size while holding the size of the minority population constant. The starkest pattern arises for the cases in which the minority population was defined as 2.5% of the overall population. In these experiments, the population quickly settles into an equilibrium in which all agents choose behavior 0, or in which all of the agents of the minority race choose behavior 0 and all of the members of the majority race choose behavior 1. The distribution of average behaviors across simulations provides clear evidence that the system is biased toward the latter equilibrium, and this is increasingly apparent as the size of the base cliques to which agents are assigned increases. Indeed, in the case of 2.5% minority agents, 19% of simulations settled on the equilibrium in which all agents choose

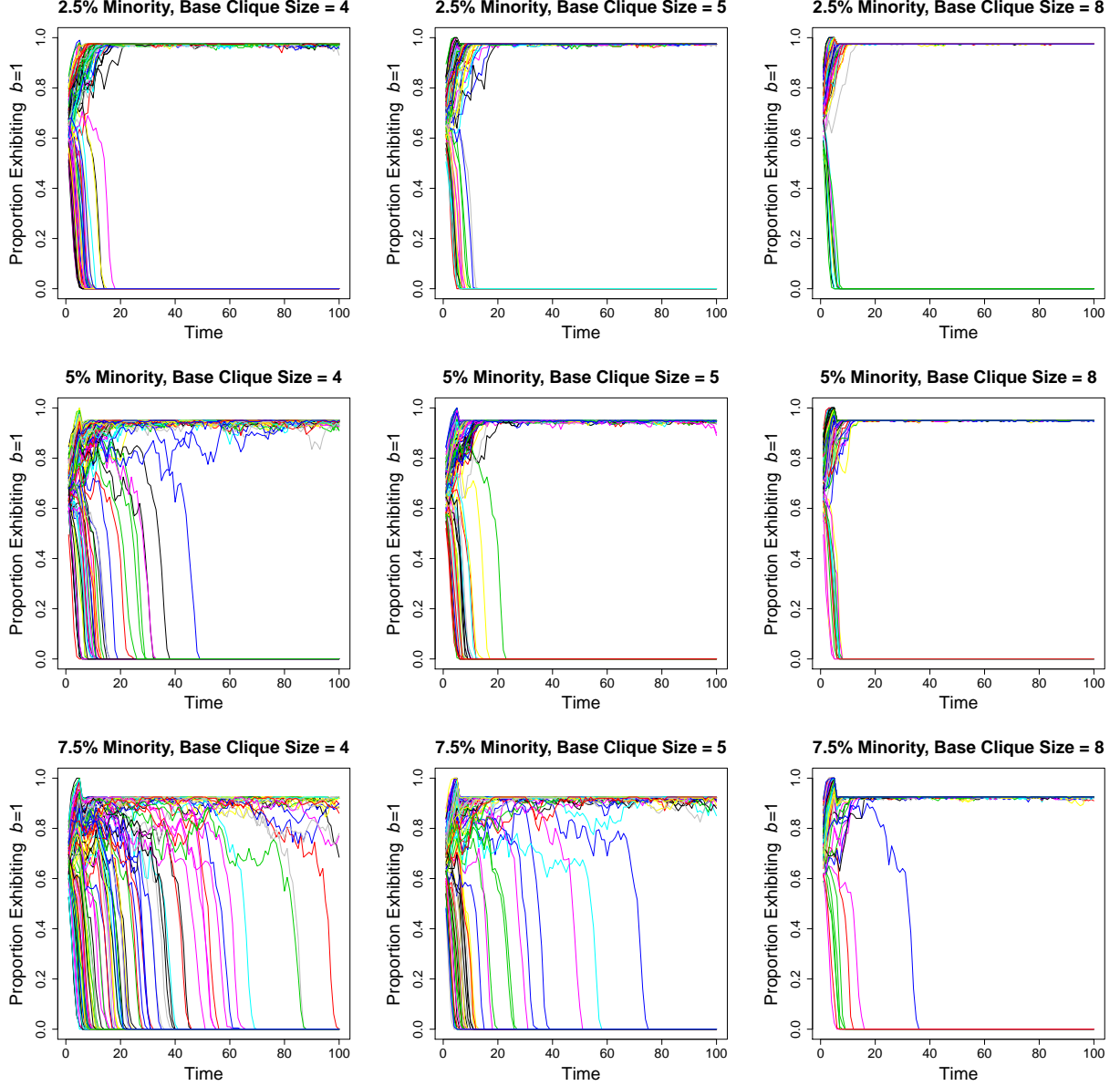


Figure 2: Simulation experiments of first 100 iterations, varying size of minority population and base clique size. Each line depicts average population behavior for one simulation; each panel depicts 100 simulations.

behavior 0 when the base clique size was 4, whereas when the base clique size is increased to 5 or 8, this percentage dropped to 11 and 5, respectively. This provides preliminary evidence that small social groups have greater potential than large groups to provide agents with sufficient reputational incentives to espouse a view of racial equality. To say this another way, large social groups create a great deal more uncertainty and a substantially lower likelihood of blacks making up a significant fraction of agents' neighborhood of direct connections. Greater uncertainty causes agents to more often have to use their past observations of the distribution of behaviors in the

population to formulate estimates of what a neighbor will do in the future. Due to initial biases in the distribution of values, this process of using past behavior to predict future behavior becomes a self-reinforcing dynamic that overwhelmingly leads to the behavior associated with a preference for inequality. And it is important to keep in mind that this is true even with most agents falling in the moderate range with respect to internal preferences.

The same phenomenon holds true when the minority population is defined at 5% or 7.5%. In either case, holding the proportion of the population from the minority group constant and increasing the base clique size leads the simulated populations to settle into equilibrium far more quickly and for the equilibrium selected in the overwhelming majority of cases to be the one in which all agents from the majority group choose behavior 1.

Scanning down columns of Figure 2 paints a similar picture. In each column, as the base clique size is held constant and the size of the minority population is increased, more and more simulated sets of agents select the equilibrium associated with behavior 0. That is, given a constant base clique size, increasing the size of the minority population that is introduced has a substantial impact on the propensity of the society to choose the behavior 0 equilibrium.

Across all simulation experiments, the most interesting cases from the perspective of norm selection are the cases in which populations of agents that had seemingly settled into the equilibrium associated with behavior 1 suddenly found it worthwhile to abandon their previous choice and transition toward the equilibrium associated with behavior 0. Substantively, these represent the cases in which social pressures to act in a way consistent with racial inequality were turned on their head, and began to operate in the opposing direction. These transitions only occurred when either the base clique size took on its lowest value (4) or the percentage of the population from the less powerful group took on its highest value (7.5). The combination of these two features in the bottom left of Figure 2 shows far and away the greatest number of transitions from a norm at $b = 1$ to a norm at $b = 0$. Figure 3 depicts several examples of these norm transitions, focusing on three particular combinations of minority population size and base clique size.

In all of these cases, the most interesting point from a substantive perspective is that the initial movement from a norm of behavior 1 to a norm of behavior 0 is driven by a sudden drop in the social status of some formerly high status individuals. However, remaining in the behavior

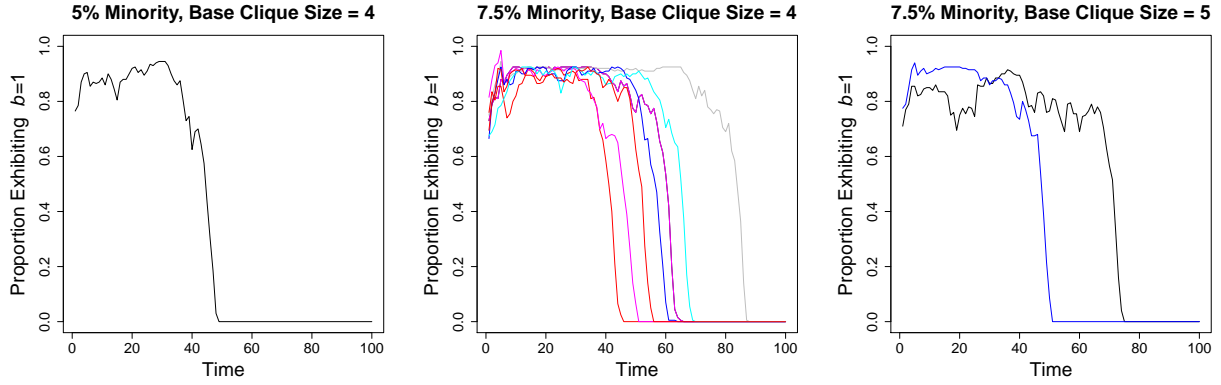


Figure 3: Example average population behavior depicting a transition from a behavior 1 equilibrium to a behavior 0 equilibrium using three parameter combinations.

0 equilibrium is driven by individuals with internal preferences close to 1 actually leading the way in exhibiting behavior 0. That is, given the initial distribution of values, with a substantial number of individuals carrying a preference for inequality, a minority population of no more than 7.5%, and regular rewiring of the network of social connections, the only way that an equilibrium at behavior 0 can be sustained is by groups of individuals with an internal preference for behavior 1 acting concertedly to exhibit the opposite behavior. Much like the results of the Prisoner's Dilemma, though all would be better off in terms of intrinsic utility if they could coordinate their efforts around behavior 0, they nonetheless collectively exhibit behavior 1 due to the previous exhibitions of behaviors combined with uncertainty about future distributions.

Conclusion

Norms of public discourse with respect to the equality of the races have undergone sweeping changes over the course of American history. Whereas it was once a virtual requirement within most white social circles to express a preference for the maintenance of white hierarchy due to the presumption that the races were inherently unequal, exactly the opposite is true today. Regardless of what individuals may truly believe, expressions consistent with the ideas that one race is inherently superior or that the white hierarchy should be maintained are now wholly unacceptable in most white social circles. It is clear that expressions in favor of racial inequality may occur and remain acceptable within self-selected groups of like-minded individuals, but in

the types of casual interactions that make up the bulk of day-to-day social communications for most individuals such expressions may now severely damage one's reputation. The historical trajectory of this change is difficult to summarize in a concise manner, having been subject to wide geographic and temporal variability, but it can be safely claimed that the pace of change in the latter half of the twentieth century was particularly brisk. Coinciding with that brisk pace was a monumental shift in public policies with respect to institutions that structure social interactions.

In this paper, I have attempted to provide a formal account of some of the underlying individual and structural mechanisms that could have plausibly brought about such changes. The model presented above provides a simplified and abstract representation of the changes in the nature of interracial interactions that resulted from policies on the desegregation of formal institutions. The key feature of those interactions is that they are non-homophilic, both interracial and intraracial. When combined with individual preferences for conformity, non-homophilic interactions are expected to carry the potential for individuals forced to express an opinion on a sensitive social issue to potentially put forth a view that is incongruent with internal preferences. If such a phenomenon takes place on a wide scale, the general implication is that norms of public discourse can shift in such a way that perhaps even a majority of individuals that make up a society consistently express opinions that differ from what they truly believe.

The most interesting implication derived from the numerical experiments, and one that was not expected ahead of time, is that a transition from a behavior largely congruent with internal values to one that is largely incongruent is greatly facilitated by assigning individuals to small groups of non-homophilic interactions. That is, holding preferences, network structure, and racial makeup constant, smaller base clique sizes are more likely to lead the population to a transition in norms from the behavior consistent with a preference for inequality to a behavior consistent with a preference for equality. In retrospect, this result makes sense and can be easily explained. Substantively, smaller social groups should be expected to be more intimate, and lead to a greater chance for interracial contacts. Small numbers of individuals in combination with interracial contacts should be more likely to lead those individuals from the dominant group with a preference for maintaining inequality to experience a decrease in social status as a result of lower uncertainty

about the distribution of preferences and the chance of being outnumbered. Formally, smaller numbers of individuals in groups should result in a greater chance of groups of individuals with a preference for equality being put together simply due to small-sample variability and chance. This, in turn, would lead to a higher probability of individual neighborhoods transitioning from an inequality norm to an equality norm, and the reformation of links leading to widespread changes in individuals' estimates of the distribution of preferences in the broader population.

Future research will investigate empirically the validity of the individual-level mechanisms that underly the formal model and further explore the relationship between the size of groups to which individuals are assigned and the likelihood of a norm transition. Additionally, the formal model can be further complicated in a variety of ways that may lead to additional insights with respect to norm transitions. At present, the results of the numerical experiments are promising in that they are plausible and in that they provide a straightforward avenue toward experimental studies to explore their tenability. Further investigation of the assumptions underlying those results is surely needed.

References

- Adams, Henry F. 1912. "Autokinetic Sensations." *Psychological Monographs* 14(59): 1–45.
- Akerlof, George A. 1980. "A Theory of Social Custom, of Which Unemployment May Be One Consequence." *Quarterly Journal of Economics* 94(4): 749–775.
- Asch, S. E. 1951. "Effects of Group Pressure Upon the Modification and Distortion of Judgments." In *Groups, Leadership and Men*, ed. Harold Guetzkow. New York: Russell and Russell pp. 177–190.
- Asch, Solomon E. 1952. *Social Psychology*. New York: Prentice Hall.
- Bandura, Albert. 1977. *Social Learning Theory*. Englewood Cliffs, NJ: Prentice-Hall.
- Bednar, Jenna, Aaron Bramson, Andrea Jones-Rooy, and Scott Page. 2010. "Emergent Cultural Signatures and Persistent Diversity: A Model of Conformity and Consistency." *Rationality and Society* 22(4): 407–444.
- Bernheim, B. Douglas. 1994. "A Theory of Conformity." *Journal of Political Economy* 102(5): 841–877.
- Boyd, Robert, and Peter J. Richerson. 1985. *Culture and the Evolutionary Process*. Chicago: University of Chicago Press.

- Boyd, Robert, and Peter J. Richerson. 2005. *The Origin and Evolution of Cultures*. Oxford: Oxford University Press.
- Clotfelter, Charles T. 2004. *After Brown: The Rise and Retreat of School Desegregation*. Princeton, NJ: Princeton University Press.
- Durkheim, Émile. [1895] 1965. *The Rules of Sociological Method*. New York: Free Press.
- Granovetter, Mark. 1986. "The Micro-Structure of School Desegregation." In *School Desegregation Research*, ed. Jeffrey Prager, Douglas Longshore, and Melvin Seeman. New York: Plenum Press pp. 81–110.
- Henrich, Joe, and Robert Boyd. 1998. "The Evolution of Conformist Transmission and the Emergence of Between-Group Differences." *Evolution and Human Behavior* 19(4): 215–241.
- Henrich, Joseph. 2001. "Cultural Transmission and the Diffusion of Innovations: Adoption Dynamics Indicate That Biased Cultural Transmission Is the Predominate Force in Behavioral Change." *American Anthropologist* 103(4): 992–1013.
- Henrich, Joseph, and Robert Boyd. 2001. "Why People Punish Defectors: Weak Conformist Transmission can Stabilize Costly Enforcement of Norms in Cooperative Dilemmas." *Journal of Theoretical Biology* 208(1): 79–89.
- Jones, Stephen R. G. 1984. *The Economics of Conformism*. Oxford: Blackwell.
- Katz, Elihu. 1957. "The Two-Step Flow of Communication: An Up-To Date Report on an Hypothesis." *Public Opinion Quarterly* 21(1): 61–78.
- Katz, Elihu, and Paul F. Lazarsfeld. 1955. *Personal Influence: The Part Played by People in the Flow of Mass Communications*. New York: Free Press.
- Kiesler, Charles A., and Sara B. Kiesler. 1969. *Conformity*. Reading, MA: Addison-Wesley.
- Kuran, Timur. 1987. "Preference Falsification, Policy Continuity and Collective Conservatism." *Economic Journal* 97(387): 642–665.
- Kuran, Timur. 1989. "Sparks and Prairie Fires: A Theory of Unanticipated Political Revolution." *Public Choice* 61(1): 41–74.
- Kuran, Timur. 1990. "Private and Public Preferences." *Economics and Philosophy* 6(1): 1–26.
- Kuran, Timur. 1991. "Now Out of Never: The Element of Surprise in the East European Revolution of 1989." *World Politics* 44(1): 7–48.
- Kuran, Timur. 1995. *Private Truths, Public Lies: The Social Consequences of Preference Falsification*. Cambridge, MA: Harvard University Press.
- Lazarsfeld, Paul F., Bernard Berelson, and Hazel Gaudet. 1948. *The People's Choice: How the Voter Makes Up His Mind in a Presidential Campaign*. New York: Columbia University Press.
- Mendelberg, Tali. 2001. *The Race Card: Campaign Strategy, Implicit Messages, and the Norm of Equality*. Princeton: Princeton University Press.
- Mutz, Diana C. 1998. *Impersonal Influence: How Perceptions of Mass Collectives Affect Political Attitudes*. Cambridge: Cambridge University Press.

- Noelle-Neumann, Elisabeth. 1974. "The Spiral of Silence: A Theory of Public Opinion." *Journal of Communication* 24(2): 43–51.
- Noelle-Neumann, Elisabeth. 1984. *The Spiral of Silence: Public Opinion—Our Social Skin*. Chicago: University of Chicago Press.
- Orfield, Gary. 1969. *The Reconstruction of Southern Education: The Schools and the 1964 Civil Rights Act*. New York: Wiley-Interscience.
- Orfield, Gary. 1983. *Public School Desegregation in the United States, 1968–1980*. Washington, D.C.: Joint Center for Political Studies.
- Orfield, Gary, and Frank Monfort. 1992. *Status of School Desegregation: The Next Generation*. Cambridge, MA: Metropolitan Opportunity Project, Harvard University.
- Page, Scott E., Leonard M. Sander, and Casey M. Schneider-Mizell. 2007. "Conformity and Dissonance in Generalized Voter Models." *Journal of Statistical Physics* 128(6): 1279–1287.
- Patchen, Martin. 1982. *Black-White Contact in Schools: Its Social and Academic Effects*. West Lafayette, IN: Purdue University Press.
- Sherif, Muzafer. 1935. "A Study of Some Social Factors in Perception." *Archives of Psychology* 27(187).
- Sherif, Muzafer. 1936. *The Psychology of Social Norms*. New York: Harper and Brothers.
- Thelen, Mark H., Stephen J. Dollinger, and Karen D. Kirkland. 1979. "Imitation and Response Uncertainty." *Journal of Genetic Psychology* 135(1): 139–152.