

# Segregation in Social Networks based on Acquaintanceship and Trust

Thomas A. DiPrete

Andrew Gelman

Tyler McCormick

Julien Teitler

Tian Zheng

Columbia University

January 4, 2010

This research was supported by National Science Foundation Grant SES-0532231, by the Applied Statistics Center of Columbia University, and by the Russell Sage Foundation. We acknowledge research assistance from Masanao Yajima and from Rozlyn Redd, and thank Rozlyn Redd, Matt Salganik, and Delia Baldassarri for their comments on this paper and at earlier stages of this research project. Previous versions of this paper were presented at the spring, 2008 Population Seminar Series of the Office of Population Research at Princeton, at the seminar series of the Economic and Social Dynamics Research Group at Yahoo, at the Inequality and Social Integration research unit at the WZB-Berlin, at the fall, 2008 sociology colloquium series at Yale, and at the spring 2009 seminar series of the Center for Demography and Ecology at the University of Wisconsin, Madison. We thank Richard Breen and the other participants at these seminars, Nan Lin, Howard Aldrich, and also three anonymous AJS reviewers for their helpful comments and suggestions. Please direct correspondence to Thomas A. DiPrete (tad61@columbia.edu), Russell Sage Foundation, 112 E. 64th St., New York, NY 10065.

## **Abstract**

Using recently collected data from the 2006 General Social Survey, we compare levels of segregation by race and along other dimensions of potential social cleavage in the contemporary United States. Americans are not as isolated as the most extreme recent estimates suggest. However, hopes that “bridging” social capital is more common in broader acquaintanceship networks than in core networks are not supported by the GSS data. Instead, the entire acquaintanceship network is perceived by Americans to be about as segregated as the much smaller network of close ties. People do not always know the religiosity, political ideology, family behaviors, or socioeconomic status of their acquaintances, but perceived social divisions on these dimensions are high and in some cases rival the extent of racial segregation in acquaintanceship networks. The major challenge to social integration today comes less from the risk of social isolation than from the tendency of many Americans to isolate themselves from others who differ on race, political ideology, level of religiosity, and other salient aspects of social identity.

## Introduction

Scholars have long recognized that Americans are socially divided along multiple dimensions. It is generally believed that social interaction is most highly segregated along racial lines, but other forms of segregation have received increased attention in the past decade. Skocpol and Fiorina (1999), for example, contend that patterns of civic engagement have become more polarized by class, while Evans (2003) and Rosenthal (2004) argue that Americans have become more polarized by political ideology. Political conflict between proponents of secular and religiously orthodox values has been especially prominent since the Reagan presidency (Green, 1996; Brooks, 2002). Coupled with this concern about high levels of segregation and polarization in contemporary American society is new evidence that close ties even to people like oneself have diminished in the past twenty years (McPherson et al., 2006).

Given the level of interest in the topic of social integration, it is remarkable how little hard evidence we have about the extent to which Americans have contact with people who differ from themselves on core status and values dimensions. Most studies use indirect measures, or focus exclusively on friendships, the people that one discussed important matters with, or other operationalizations for the set of people to whom one has strong ties. Little is known about how religion, political ideology, or social class structure the broader acquaintanceship networks of Americans. In light of the huge number of studies that focus on residential segregation, it is ironic but true that the same can be said about racial segregation in acquaintanceships. As a consequence, we do not know whether religion, class, or political ideology rival race in shaping everyday patterns of social interaction. We do not know whether Americans have more integrated social networks at their workplace and in voluntary associations than they do in their families or neighborhoods. These questions are the focus of much speculation, but there is little firm knowledge about their answers.

Using recently collected data from the 2006 GSS, we compare levels of segregation by race and across the principal dimensions of potential social cleavage in contemporary America. We study

both the relatively small networks based on trust relationships and the much larger acquaintanceship networks of Americans in order to answer three major questions. First, how socially connected are Americans? Second, to what extent do these connections cross social boundaries defined by race, socioeconomic markers, political ideology, and religiosity? Third, is the expected high level of homophily in core networks offset by greater diversity in the larger group of people that count as acquaintances? Our answers to these questions offer a mix of reassurance and concern to those who value social integration. We find that Americans are not as isolated as suggested by recent estimates obtained from the 2004 GSS (McPherson et al., 2006). However, hopes that “bridging” social capital is more common in broader acquaintanceship networks than in core networks are not supported by the 2006 GSS data. Instead, the entire acquaintanceship network appears to be about as segregated as the much smaller network of close ties. We find that social divisions based on religiosity, political ideology, family behaviors, and socioeconomic standing are high and in some cases rival racial segregation in their intensity. Social polarization rather than social isolation appears to be the greater impediment to social integration in the U.S. today. Our most positive result is the surprising integrative role of the family. The growing heterogeneity of American families, coupled with the difficulty of hiding potentially objectionable statuses from other family members, appear to produce family-based social networks that are less segregated on a number of dimensions than are networks based on workplace, neighborhood, or voluntary associations.

## **Social Integration and Interpersonal Association**

It has long been known that people prefer to associate with others who are similar to themselves, which produces segregation in people’s social networks along a variety of core demographic statuses, including race/ethnicity, age, education and income (Billy et al., 1984; Coleman, 1961; Blau, 1977; McPherson and Smith-Lovin, 1987; McPherson et al., 2001). The homophily principle is so powerful that its existence is taken as a given in the social capital literature. Two other issues, however, are considered to be highly problematic in the contemporary U.S., and arise from the recent

and growing literature on social integration in modern Western societies. One issue concerns absolute levels of social isolation, i.e, the quantitative extent to which people are socially connected to others, including with people like themselves. The second issue concerns relative isolation, i.e. the extent to which people—despite their tendencies toward homophily— have sufficient ties with people who are different from themselves to be exposed in a meaningful sense to a broad spectrum of attitudes, beliefs, and opinions.

Social isolation is theoretically linked in the contemporary literature to the issue of social inclusion or exclusion, which especially in the European context has been closely tied to concerns about social inequality and poverty. Social inclusion is defined by the European Social Fund as the ability “to participate fully in economic, social and cultural life and to enjoy a standard of living and well-being that is considered normal in the society in which they live” (Council of the European Union, 2004). People are included in the “life of the community” (Sen 1992, p. 39) through their social capital as well as through consumption of goods and services made possible by an adequate income. From this characteristically European perspective, social inclusion or exclusion has both a material aspect, which affects standard of living, and a social aspect, which affects level of integration into the broader society. Each of these aspects, moreover, can be conceptualized at the level of the individual or of social groups, and becomes a measure of the level of integration and inequality for the society as a whole.

The American discussion similarly addresses both material and social dimensions. Some of this literature follows Bourdieu (1980) and Coleman (1988) in placing primary emphasis on social capital as an individual-level resource in arenas such as educational attainment, labor markets, business, and politics. Other scholars, notably Putnam (Putnam, 1993, 1996, 2000) and Portes (Portes and Sensenbrenner, 1993; Portes, 1998; Portes, 2000) stress the macro-level characteristics of social capital. Portes has placed primary emphasis on homophilous social capital, particularly within the context of ethnic communities, which he refers to as “bounded solidarity” (Portes, 1998), and which corresponds to what Gittell and Vidal (1998) refer to as “bonding” social capital. As Portes (1998) and Waldinger (1995) have argued, bounded solidarity can be a resource for an

immigrant community, but it also can be a source of deprivation when practiced by more privileged groups (e.g. white ethnic workers in the construction trades) to exclude new ethnic groups from jobs.

Contrasting to “bounded solidarity” or “bonding” social capital is what Gittel and Vidal (1998) called “bridging” social capital, which concerns extra-community ties, and which fosters integration in the larger society through heightened levels of trust (Woolcock, 1998; Fukuyama, 1995; Gambetta, 1988; Putnam, 2000). When trust is low, social isolation is high. High in-group trust (high “bonding” social capital) but low out-group trust (low “bridging” social capital) “bolsters narrow identities” and “may create strong out-group antagonism” according to Putnam (p. 23). In contrast, “bridging” social capital involves connections that “are outward looking and encompass people across diverse social cleavages” (Putnam, 2000). The combination of “bonding” and “bridging” social capital arguably correspond to the condition of “generalized trust” (Putnam, 2000) where one thinks that “people in general can be trusted” because one actually has experience interacting with people who are both similar to and different from oneself (Paxton, 2007).

Prominent scholars claim to have found dis-integrationist trends in American patterns of association. Putnam (2000) provided numerous sources of evidence for declining civic engagement, and concluded his book by arguing that “the evidence from our inquiry shows that this longing is not simply nostalgia or ‘false consciousness.’ Americans are right that the bonds of our communities have withered, and we are right to fear that this transformation has very real costs” (p. 402). Skocpol and Fiorina (1999) reached somewhat similar conclusions, namely that Americans were increasingly detached from the kinds of cross-class membership organizations that had once defined the landscape of voluntary association in America, to be replaced by nominal memberships (what Putnam called “mailing list” memberships) that were primarily defined by the paying of dues rather than actual social interaction.<sup>1</sup>

Other forms of evidence paint a mixed picture. Residential segregation between blacks and whites declined between 1970 and 2000 though not to a large extent and not uniformly, while Asian and Hispanic residential segregation has slightly increased (Massey and Denton, 1993; Ice-

land et al., 2002; Frey and Myers, 2005). Families have become more heterogeneous, and interracial marriages in particular have increased though remain relatively rare (Ellwood and Jencks, 2004; Gullickson, 2006). Meanwhile, abundant evidence has emerged concerning the growing correlation of statuses in American society, a process that Blau (1977) characterized as the “consolidation” of social parameters. This growing correlation opposes the mild integrationist trend that some see in the residential segregation data. In particular, the association between income and family type has increased (Burtless, 1999). The association between wife’s education and husband’s education has increased (Schwartz and Mare, 2005). The association between income and political partisanship has increased (McCarty et al., 2006). Our own calculations from the General Social Surveys have established that the association between being married with children and frequent church attender increased, the association between being married with children and being politically conservative increased, and the association between being a frequent church attender and being politically conservative increased. All other things equal, one would expect that a rising correlation of statuses would imply a lower frequency of “cross-cutting status sets” and “cross-cutting cleavages,” and higher levels of values polarization and conflict (Merton, 1957; Coleman, 1957; Lipset, 1963). Consistent with this expectation is Lee’s (2007) finding that generalized trust has been declining in the U.S. for the past 30 years. Also consistent is the work of Poole and Rosenthal (2000), who documented a growing distance between the political positions of the median Democrat and the median Republican since roughly the middle 1970s. While DiMaggio et al. (1996) found no evidence for a growing values divide as of the middle 1990s, analyses of more current trend data by Evans (2003) show growing evidence that “partisan” Americans (those who label themselves as liberals or conservatives) were becoming polarized around moral issues such as abortion, sexuality, school prayer (see also Mouw and Sobel, 2001; Green, 1996; Brooks, 2002; Frank, 2004; and Baldassarri and Gelman, 2008).

Recent studies suggest that the absolute level of connectedness of Americans depends upon the character of the relationship tie elicited by the survey question. Zheng et al. (2006) obtained a median network size estimate of 610 based on the 1998 McCarty et al. (2001) survey that asked

respondents questions of the form “How many people do you know who [are in group X]?” The 2006 Pew survey instead queried respondents about their strong ties using the prompt

Let’s start with the people you feel [alternatively SOMEWHAT CLOSE TO or VERY CLOSE TO], which might include those you discuss important matters with, regularly keep in touch with, or are there for you when you need help. Thinking about ALL the people who fit this description and who do NOT live with you, how many are. . .

Using these prompts, Boase et al. (2006) found that Americans had a median of 35 somewhat close ties and 15 very close ties.

The 2004 GSS used a different prompt, and reported a much lower level of connectedness (McPherson et al., 2006). In both the 1985 and the 2004 surveys, the GSS interviewer asked:

“From time to time, most people discuss important matters with other people. Looking back over the last six months—who are the people with whom you discussed matters important to you? Just tell me their first names or initials. IF LESS THAN 5 NAMES MENTIONED, PROBE: Anyone else?” (NORC interviewer writes down just the first five names and then asks further questions about these names).

In 1985, the mean respondent reported that he/she had discussed important matters during the past six months with 2.9 individuals out of a maximum of five. In 2004, in contrast, the mean was only 2.1, and one quarter of 2004 respondents (later revised to 22.5% in McPherson et al. (2008)) offered no names in response to this question vs. 10% in 1985 (McPherson et al., 2006). This high estimate has recently been criticized by Fischer (2009), and both Fischer and McPherson et al. apparently now agree that the 22.5% estimate of social isolates is at least partly an artifact of the data collection process in the 2004 GSS (McPherson et al., 2009). Regardless of the correct answer, however, estimates of core network size cannot by themselves reveal the level of social integration achieved through social interaction, because much of this interaction occurs with associates who would not be characterized as strong ties.

## **Strong and Weak Social Ties across “Diverse Social Cleavages”**

Putnam argued in *Bowling Alone* that the “bonding”/“bridging” distinction is “perhaps the most important” dimension along which social capital could vary, but that he could find “no reliable, comprehensive, nationwide measures of social capital that neatly distinguish ‘bridgingness’ and ‘bondingness’,” which caused him to de-emphasize this distinction in his empirical analysis and focus instead on the simpler question of whether social capital in general had declined (Putnam 2000, pp. 22, 23). Despite the large empirical literature on social networks, his conclusion about the state of available evidence remains accurate for two reasons. First, more attention has been paid in homophily studies to some statuses than to others, which leaves gaps in our understanding about potential barriers to social interaction. The second and more fundamental reason is the lack of good data about the structure of complete social networks –including the weak ties as well as the strong ones.

As McPherson et al. (2001) discuss, studies of association range from marriage (Kalmijn, 1998), confidants and friends (Marsden, 1988; Verbrugge, 1977, 1983) to mere contact (Wellman, 1996), knowing about someone (Hampton and Wellman, 2001) or appearing with them in a public place (Mayhew et al., 1995). This literature documents multiple dimensions of homophily, including age, gender, race, and socioeconomic status. However, much of what is known about the level of homophily in social networks concerns close relationships (Moody, 2001), largely because of the methodological difficulty of gathering information about people to whom one has relatively weak ties.

Race is typically identified as the dimension along which social networks are most segregated. Most of the evidence for this assertion comes from the study of close ties of marriage, kinship, and friendship, especially school friendships or core-network designs such as the 1985 and 2004 GSS (Marsden 1988; McPherson et al. 2001). Marsden’s (1987) study of the 1985 GSS questions about core social networks found that only 8% of adults with networks of size two or more reported being tied to someone of a different race. Marsden estimated this frequency as only one-seventh as high as one would expect if people sorted themselves at random. Many studies have similarly found

strong evidence of segregation in racial friendships (e.g., Quillian and Campbell, 2003; Moody, 2001; Mouw and Entwisle, 2006). But, to repeat, these studies are almost always about close ties. Little is known about inter-racial acquaintanceships made at work, in the neighborhood, or in voluntary associations.

Even less is known about ties among Americans with different religious practices or political preferences. McPherson et al. (2001) argued that marriage, friendship, and confiding relations are homophilous with respect to religion, though religious homophily is not typically as strong as race or ethnicity (Laumann, 1973; Marsden, 1988; Fischer, 1982; Louch, 2000). Kalmijn (1998) reported that marital homophily with respect to religion appears to be declining. McPherson et al. (2001) note that some religious groups (e.g., Jews) clearly display homophily in their choice of friends and spouses. In contrast, they conclude from their review of the literature that religion –by which they primarily mean religious denomination– "may not matter much at all" in relationships that are not close. According to McPherson et al. (2001), the main exception concerns fundamentalists and members of sects, for whom religion has become something of a total environment.<sup>2</sup> Similarly, McPherson et al. (2001) report that people form ties based on a similarity of values as well as of social statuses, but the extent to which this generalization covers weak ties outside friendship groups or core social networks is an open question.

Many scholars have offered speculation about the relationship between tie strength and level of homophily. The principle underlying Granovetter's "strength of weak ties" hypothesis was that weak ties provided connections to people who were more occupationally and socioeconomically dissimilar from oneself than did strong ties (Granovetter, 1973; see also Lin, 1999). Putnam similarly argued that close ties were more likely to be with people like oneself, while weak ties were more likely to be with people who are different from oneself. Smith-Lovin (2007), following Blau (1977), argued that homophilous as well as multiplex ties are more likely to be strong ties, while ties among dissimilar others are more likely to be weak. The 2004 GSS data, however, suggested that multiplex ties are uncommon even within core social networks (Smith-Lovin, 2007).

The major challenge for testing these ideas is that relatively little is known about the structure

of weak ties. Research using position generators (Lin et al., 2001) and resource generators (Van der Gaag and Snijders, 2005) has focused more on the specific issue of instrumental ties in the labor market than on the broader question of social integration. So-called “complete network” designs, in which the connections between all members of some relevant subpopulation are collected (e.g. the Newcomb (1961) fraternity study, the Add Health friendship and sexual relationship study (Bearman et al., 2004), or the Nang Rong, Thailand study (Rindfuss et al., 2004)) obviously miss weak ties that link outside the subpopulation under study, and in any case, these designs do not scale well to the world of adult Americans. The 2006 GSS data, therefore, offers the potential to fill an important gap in scientific knowledge about the structure of segregation and homophily in complete social networks.

## **Data and Methods**

The data for this study were collected as a special topical module in the 2006 General Social Survey. The basic design was similar to McCarty et al.’s 1998 and 1999 surveys that employed a “how many X’s do you know?” methodology in order to estimate the distribution of individuals’ network size, and also to estimate the sizes of special subpopulations that tend to be hard to count with standard survey methodologies (McCarty et al., 2001). Our survey differed from the McCarty et al. surveys in its focus on ties to highly salient groups that define important sources of heterogeneity among Americans and potentially important sources of social cleavage. Our survey also differed from McCarty et al. in the type of relationships that we measured and in the several subsets of a person’s full network that our questions pertained to.

We asked about two types of relationships. Our prompt concerning acquaintanceship was as follows:

I’m going to ask you some questions about all the people that you are acquainted with (meaning that you know their name and would stop and talk at least for a moment if you ran into the person on the street or in a shopping mall). Again, please answer the

questions as best you can.

The second type of relationship that we studied concerned trust. Coleman defined trust as the willingness to place intellectual, financial, physical or other resources at the disposal of another party (Coleman, 1990).<sup>3</sup> An individual usually trusts one's friends, but there are other people one may trust who do not qualify as friends, such as kin, or mentors, or people that one has a service or business relationship with. The extent of one's trust relationships may in turn be related to one's level of "generalized trust," i.e., one's belief about the trustworthiness of the average person or of the "benevolence of human nature in general" (Yamagishi and Yamagishi, 1994). Our trust question is about the respondent's specific trust relationships as opposed to generalized trust, and was elicited with the following prompt:

Now I'm going to ask you some questions about people that you trust, for example good friends, people you discuss important matters with, or trust for advice, or trust with money. Some of these questions may seem unusual but they are an important way to help us understand more about social networks in America. Please answer the questions as best you can.

Following the prompts concerning acquaintanceship or trust, the GSS interviewers asked respondents a series of "how many of the people that-you-are-acquainted-with/that you-trust are named [one of a set of names]" in order to estimate the size of the respondent's network (i.e., the network degree).<sup>4</sup> The interviewers then asked about specific ties with people at various socioeconomic levels, people who were members of various race and ethnic groups, people with various religious behaviors, people in various family types, and people with various political orientations.<sup>5</sup> The specific groups that we asked about are listed in Table 1.

[Table 1 about here.]

It is a general property of human interaction that statuses, behaviors, and values which are central to one's own identity may be misperceived or go unnoticed by one's acquaintances. Ego

would generally know the race of people that he is acquainted with, and he may well know the political ideology, religiosity, or family situation of people that he knows well. However, ego might often not know the political ideology, religiosity, or family situation of his acquaintances. If he were to count the number of his associates who are politically liberal, or who are gay, or who attend religious services on a regular basis, he would (necessarily) base his count on his *perceptions* about others. Thus, when two individuals with the same estimated network size report that they know very different numbers of people who are politically conservative, there might in fact be a big difference in the number of political conservatives in their networks, or they could instead have similar networks but very different perceptions about their acquaintances. Regardless of the true level of integration of acquaintanceship networks, perceived integration is important because it describes the social world as experienced by the people who live in it. As Thomas and Thomas (1928) wrote, “If men define situations as real, they are real in their consequences.”

In the McCarty et al. surveys, the groups being asked about were often very small (e.g., women who adopted kids in the past year, or people who committed suicide in the past year), and respondents were asked to list the exact number of individuals they knew in each of these groups. In contrast, our interest encompasses socially prominent groups that typically have a large membership (e.g., people who are unemployed, or people respondent is pretty certain attend religious services rarely or never), and it is either burdensome or infeasible to ask respondents to recall the exact number of people they know in these groups. Consequently, we asked respondents to indicate whether the number of people they knew fell within specific numerical ranges, specifically zero, one, two to five, six to ten, or more than ten.

We asked questions about the number of persons known or trusted in the respondent’s entire social network. In addition, we asked these questions with respect to four specified subnetworks: (1) family, relatives, or in laws, (2) neighbors, (3) people at work or customers or clients, and (4) people from associations, clubs, preschool, school, or places of worship. We asked about each of these subnetworks to establish how segregation with respect to specific groups varied across major “foci of interaction” within a person’s overall (Feld, 1981). These questions also served

two methodological purposes: they reduced response burden by limiting the scope for the recall process, and they created additional response variance concerning the number of ties with persons in the specified social groups.

Our overall sample size was 1371. In order to accomplish the project’s objectives, we subdivided our sample in complex ways. Fifty percent of the sample were asked the questions about acquaintanceship and trust concerning their entire social network. The other fifty percent were divided into four subsamples, and each of these subsamples was asked about ties within three of the four subnetworks listed above. Figure 1 illustrates the sample design. Restrictions on total module length caused us to exclude questions about contact with the same or opposite gender because men and women make up such large shares of the population that it would be difficult, given our methods, to measure variation with accuracy.<sup>6</sup> We also omitted questions about contact with groups defined by age or education in order to focus on the cleavages most salient to the current debate on social integration, namely race/ethnicity, class, religion, political ideology, and family or romantic relationships. The response rate varied by question, from 99% for some of the names and the race questions to 95% for having acquaintances who were unemployed or who owned a second home or were gay, to 92% for knowing people who go to church on a regular basis and 89% for knowing people who never attend church. The lowest response rate (81%) was for knowing people who “you are pretty certain are strongly liberal.” The pattern of missing data for the trust questions was similar to that for the acquaintanceship questions.

[Figure 1 about here.]

Our modeling strategy is described in detail in Appendix A (see also Zheng et al., 2006). We assume that the number of individuals in group  $k$  that are known to individual  $i$  (i.e.,  $y_{ik}$ ) follows a Poisson model, i.e.

$$y_{ik} \sim \text{Poisson}(\lambda_{ik})$$

where  $\lambda_{ik}$  is the expected number of individuals that individual  $i$  knows in group  $k$ . The main task therefore is to model  $\lambda_{ik}$ .

In a world where associations were made at random, it would be straightforward to model  $\lambda_{ik}$ ; for every individual  $i$ , the expected number of people in group  $k$  that she knows would equal the product of the size (degree) of her network multiplied by the fraction of all acquaintanceship ties that involve group  $k$ . For example, if 12% of all acquaintanceship ties involved African-Americans, an individual who know 500 people would be expected to know 60 African-Americans. More formally, let

$a_i$  equal the estimated degree of individual  $i$ 's acquaintanceship network.

$b_k$  equal the proportion of all ties that involve group  $k$ . Then we could write

$$y_{ik} \sim \text{Poisson}(a_i b_k) \tag{1}$$

Model (1) is unrealistic because individuals differ in their propensity to know members of any particular social group. We take this overdispersion into account by allowing the relative propensity of individuals to know members of group  $k$  to differ. We define  $g_{ik}$  as the relative propensity of individual  $i$  to know someone in group  $k$ , where  $g$  is the ratio of the expected number of ties for individual  $i$  to the number of ties he would be expected to have if acquaintanceship ties were made at random, i.e.,

$$g_{ik} = \frac{\lambda_{ik}}{a_i b_k}$$

and we elaborate the basic model such that

$$y_{ik} \sim \text{Poisson}(a_i b_k g_{ik}) \tag{2}$$

We cannot directly estimate the parameters in model (2) because the number of parameters exceeds the number of data points. Instead, we integrate out the  $g_{ik}$  by assuming that it follows a gamma distribution, and thereby obtain the negative binomial model.

$$y_{ik} \sim \text{negative binomial}(\text{mean} = a_i b_k, \text{overdispersion} = \omega_k)$$

where  $\omega_k$  scales the variance of the number of acquaintanceship ties between individuals in the population and members of group  $k$ , i.e.,

$$V(y_{ik}) = \omega_k E(y_{ik})$$

Higher values of  $\omega_k$  imply greater overdispersion. When  $\omega_k$  is unity, the negative binomial model reduces to the Poisson model where the variance equals the mean.

We use overdispersion as our primary measure of network segregation. Segregation, homophily, polarization and overdispersion are related concepts, but they are not exactly the same. DiMaggio et al. (1996) used “polarization” to refer to three aspects of the distribution of public opinion: the extent to which opinions on some issue were opposed, the extent to which attitudes on different issues were correlated (they used the word “constrained”) and the extent to which attitudes were correlated with various social statuses (which they referred to as “consolidation”). Taking opinions one at a time, they measured the level of polarization in terms of the variance of the attitude distribution (they called this “dispersion”) and the shape of the distribution (they measured this in terms of kurtosis, which is related to bimodality).

The related concept of “segregation” is the extent to which people are separated from each other on the basis of specific statuses, such as race, gender, or learning difficulties. The separation is typically defined with respect to some single characteristic of individuals, such as one’s occupation, job, employer, classroom, or the geographic location of one’s residence. It is typically measured in terms of the difference in the distribution of two or more groups with respect to this characteristic (e.g., as the percent of each group that would have to be rearranged in order to equalize the distributions of the groups). High segregation implies unequal or at least different group experiences with respect to the characteristic in question (job, residence, or classroom) and also usually implies lowered rates of contact to the extent that social interaction is structured by geography, employer, classroom etc.

In this paper, we are directly concerned with the level of contact itself rather than the charac-

teristics that may structure contact, and so we use the term “segregation,” which is related to the concept of homophily, i.e. the tendency for people to associate with others who are like themselves on some (or several) particular status or attitude or belief dimension(s). We operationalize network segregation as the extent to which the individual-level variance in the level of contact with a particular social group (“dispersion”) is higher than one would expect under a random mixing model. In theory, high overdispersion could be produced by low homophily (e.g., if people avoided contact with others like themselves), but as a practical matter (and as we have verified for the GSS data), overdispersion is generally produced in large part by homophily. In other words, people tend to know more people who are in the same statuses as themselves than one would expect from an assumption of random interaction.

Additional conceptual insight can be obtained by comparing overdispersion to measures that have been used in the literature on segregation. While the index of dissimilarity is the most well-known measure of segregation, researchers have also conceptualized segregation as a measure of inequality across geographic units (e.g., census tracts) within some larger geographic area (e.g., metropolitan areas) in the proportion of the population that is minority (Massey and Denton, 1988). Our model focuses on individuals, not geographic units, and instead of proportions who are in a particular group, we model the number of ties that involve a specific group. The coefficient of variation (CV), which is a standard measure of inequality (Allison, 1978), equals the standard deviation of some resource divided by its mean. If we conceptualize contact with a specific group as a resource that may be unequally distributed in the population, then it follows that

$$CV_{ik} = \frac{\sqrt{V(y_{ik})}}{E(y_{ik})} = \sqrt{\frac{\omega_k}{a_i b_k}}$$

If we take the ratio of inequality of contact with members of groups  $k$  and  $k'$  for individuals who have the same network size (i.e., the same value of  $a$ ), we obtain

$$\frac{CV_k}{CV_{k'}} = \sqrt{\frac{b_{k'} \omega_k}{b_k \omega_{k'}}$$

In other words, if contact with groups  $k$  and  $k'$  is equally overdispersed, then the inequality of contact with group  $k$  differs from the inequality of contact with group  $k'$  only as a function of the difference in the relative share of ties that involve groups  $k$  and  $k'$ .

A related measure of segregation is the index of exposure that was introduced by Bell (1953), and elaborated by Lieberson (1981) (see also Massey and Denton, 1988). The isolation index (which equals one minus the interaction index) measures the extent to which members of a particular group are exposed only to one another, rather than to the rest of the population. This index was written by Lieberson as

$${}_xP_x^* = \frac{1}{X} \sum_{i=1}^N \frac{x_i^2}{t_i}$$

where  $x_i$  are the number of members of minority group  $x$  in geographic unit  $i$ ,  $t_i$  is the total population in geographic unit  $i$ , and  $X$  is a scaling factor that equals the total number of minority group members across all  $N$  geographic units. If we substitute individuals for geographic units, then  $x_i$  is analogous to the number of ties between an individual and members of group  $x$ , and  $t_i$  becomes the size of individual  $i$ 's network. If we re-express this relationship in terms of expectations from our Poisson model (and refer to group  $x$  as group  $k$ ), we obtain

$$P_k \sim \sum_{i=1}^N \frac{(a_i b_k g_{ik})^2}{a_i} = b_k^2 \sum_{i=1}^N a_i g_{ik}^2$$

If everyone in the population had the same network size, this expression becomes a simple function of the variance of the relative propensities in the population to have ties with members of group  $k$ , which is related to the overdispersion parameter,  $\omega_k$ . Thus, we see that standard measures of segregation are closely related to the concept of overdispersion used in this paper. Residential segregation measures are typically computed for specific geographic areas, for example, metropolitan areas. In this paper, we compute measures of overdispersion across the entire country rather than (for example) for distinct metropolitan areas, but this is a consequence of the nature of our data (a national sample of limited size) rather than of the measure; if sufficient data were available, overdispersion measures for segregation in acquaintanceship or trust networks could also be com-

puted for individuals living within specific metropolitan areas or other geographic areas within the United States. The social networks of these individuals, of course, would generally extend outside the specific geographic area unless the question restricted the network scope to alters living in the same area as ego.

Three further issues need to be briefly summarized. Two of these issues concern the estimation of the size of acquaintanceship or trust networks. In model (2), the predicted  $y_{ik}$  depends on the product of  $a_i$  (the size of ego's network) and  $b_k$  (the proportion of ties that involve group  $k$ ). In order to identify  $a_i$  and  $b_k$  separately, we borrow information about the size of the groups from other sources, such as the fraction of the population with specific names (see Appendix A or McCormick and Zheng (2007) for details).

The second issue concerning the estimation of network size is recall error. Prior research demonstrates that individuals find it easier to count accurately the number of individuals they know from rare groups than from common groups. Put concretely, it is easier to recall the number of females that one knows who are named Bethany than it is to recall the number of males one knows who are named Michael.<sup>7</sup> To ease respondent burden, we used intervals to ask respondents about people they know (zero, one, 2-5, 6-10, or greater than 10), but this does not by itself solve the problem of under-reporting. McCormick and Zheng (2007) show that people tend to over-recall ties involving very rare names and under-recall ties involving common names. We estimated a recall function to transform the known proportion of group  $k$  in the population into an estimate of the fraction of network ties that will be recalled to connect with group  $k$ , and this then gives our estimate of degree size (see Appendix A for further details).

Using external information on the frequency of names along with the recall function works well for estimating the size of acquaintance networks, but it gives estimates of the size of trust networks that in our judgment are too large. The names that we selected for the GSS survey were only a small fraction of 1% of the American population, which means that 0, 1, and 2-5 would be typical responses to the question about how many people of this name one is acquainted with. However, trust networks are much smaller than acquaintanceship networks. It would have required

another set of more common names – each around 1% of the population – to estimate the size of the trust network on the basis of trust of people with a given set of first names. It is also likely that recall problems are much less severe for the relatively small group of people that one trusts than for the larger group of people that one is acquainted with. Consequently, applying the recall function estimated from the acquaintanceship data to the trust network would upwardly bias the estimated number of people that one trusts. An alternative normalization strategy assumes that the proportion of ties involving racial groups equals their collective proportion in the population. We use the latter strategy in our analysis of the network of people that one trusts. Our alternative normalization strategy provides estimates of the size of trust networks that closely approximate estimates obtained in the 2006 Pew Survey (Boase et al., 2006). This similarity suggests that the logic underlying our race-normalization strategy is reasonable (see below including footnote 20 for additional information). In any case, our estimates of overdispersion are not affected by our choice of normalization strategy or by the use of a recall function to estimate the size of the degree network (see Zheng et al. (2006) for further details).<sup>8</sup>

The third issue, which we have already mentioned above, concerns the distinction between observable and hidden statuses. Killworth et al. (2003) refer to the situation where information about one's status is not transmitted with equal probability to all people that one knows as a "transmission effect." Some statuses –most notably skin color– are often (though not always) observable. Other characteristics such as political ideology or sexual orientation are not as readily observed, and it might often be true that a respondent would recall a particular acquaintance but not necessarily know that the acquaintance was politically conservative, gay, in a cohabiting relationship, or someone who goes to church on a regular basis. Sometimes the respondent does not know because the information has low salience for him. In other cases, he may overestimate the extent to which other people that he knows are like himself (McPherson et al., 2001; Goel et al., 2009). Finally, sometimes the information is masked on purpose by acquaintances who think he would be put off by this knowledge (Noelle-Neumann, 1993). Thus, conservatives may hide their ideological orientation from liberals, gays may hide their sexual orientation fact from those who are homophobic,

etc. Generally speaking, we expect these “transmission errors” will make networks appear to be more segregated than they actually are and may contribute to a perception that the U.S. is a more polarized society than it actually is (DiMaggio et al., 1996; Baldassari and Bearman, 2007; Gelman et al., 2008). The fact that our estimates will overstate segregation on certain dimensions is not simple error, however; it instead provides an accurate estimate of the level of segregation and the extent of “bridging social capital” that ego perceives in his network.

## **Results**

### **Acquaintance Networks**

The size of acquaintanceship networks varies substantially in the adult population. Figure 2 shows the distribution of the recall adjusted acquaintanceship network. We estimate that the median person is acquainted with 550 people, with an estimated interquartile range of approximately 400 to 800.<sup>9</sup> Our estimate from the 2006 GSS data is similar to the 610 estimate of the median made by Zheng et al. (2006) based on the 2000/2001 Kilworth and McCarty data (see also Marsden, 2005).<sup>10</sup> As Table 2 shows, the strongest predictors of acquaintanceship degree in our data are education, income, race, immigrant status, and church attendance, a pattern that is consistent with studies that have used other strategies to study social networks (McPherson, 1983; Marsden, 1987; McPherson et al., 2006).<sup>11</sup> Each year of education is associated with an increase of 22 people, or about 3%, in one’s acquaintanceship network. Net of education, income also has a small effect, with each \$10,000 in additional family income predicting an increase of 9 acquaintances. Blacks and U.S. born Hispanics have smaller estimated networks than do whites, though the difference is not statistically significant, net of other covariates in the model. However, members of other races and foreign-born Hispanics have estimated acquaintanceship networks that are 26% smaller than those of white, and respondents who attend church on a weekly basis have 25% larger networks (about 150 people) than do those who rarely or never attend church, net of other covariates. The added network members of frequent church members are presumably the people that they know

from their participation in religious services and other activities at their places of worship. The relatively small networks of the foreign born and of respondents who are neither white, black, nor Hispanic suggests greater social isolation for respondents who migrated to this country and for those who belong to relatively small population groups (cf. Blau, 1977), though our methodology may understate the network size for these latter groups of Americans.<sup>12</sup>

[Figure 2 about here.]

[Table 2 about here.]

Estimated overdispersions in acquaintanceship social networks (medians and interquartile ranges) are presented in the second and third columns of Table 3. The overdispersion parameters provide an estimate of the ratio of the true variance to the variance from the null model of random mixing. In the case of people named Kevin, the estimated median overdispersion is 1.7. So for example, if ego knows 900 people, and if 1% of all people are named Kevin, then ego would be expected to know 9 people named Kevin under the null model with a standard deviation of 3. An overdispersion of 1.7 implies that the standard deviation of the number of Kevins known to people with 900 acquaintances is inflated from 3 upwards only slightly to 3.9 people (i.e., 3 multiplied by the square root of 1.7). In general, the overdispersions for groups defined by names were low, which supports our using these names to estimate the distribution of network degree in the GSS sample. In contrast, overdispersion is much greater for ties with groups defined by or related to class, race, political orientation or religiosity. For example if 5% of social ties involved the unemployed, then a person who knew 500 people would be expected under the assumption of random mixing to know 25 unemployed people with a standard deviation of 5. In fact, we estimate the standard deviation to be 16, implying an approximate 95% confidence interval of 0 to 57, which is wider than the 15-35 confidence interval in a world of random mixing. In other words, the social networks of actual Americans are more heterogeneous than the random mixing model would predict, with some people knowing very few unemployed, while for others more than 10% of their acquaintances are unemployed.

[Table 3 about here.]

The existing literature, which is largely based on information collected about a few close ties, reports that segregation on the basis of race outstrips by far segregation on other social variables. Our data clearly support earlier findings showing a high degree of segregation on the basis of race. Because whites are numerically dominant, we cannot accurately estimate the level of overdispersion of the number of whites one is acquainted with.<sup>13</sup> For blacks and Hispanics, however, our results show overdispersion parameters of about 9 or 10. In a network of 500 acquaintanceships, we would expect at random about 12% black and Hispanic acquaintances, or 60 blacks and Hispanics each out of 500, and a standard deviation of about 8, and so 95% of social networks would have between 44 to 76 of each group. Instead, the estimated standard deviation is on the order of 25, giving a 95% band of about 10-110 for each group.<sup>14</sup>

Another way to illustrate the meaning of overdispersion is to compare our estimated probabilities of being acquainted with (knowing) especially few (or especially many) members of any particular group against the benchmark of random mixing. Table 4 shows the estimated number in a 400 person network (the 25th percentile of estimated network size) that would belong to each of the measured subgroups based on the proportion that each of these groups constitutes of the American population. We then compare the probability of knowing ten or fewer in each of these groups under the assumption of random mixing with the estimates from our model based on the actual patterns of segregation found in the data. The probability of having only 10 or fewer acquaintanceships out of a 400 person network in each of these groups would be extremely small under the assumption of random mixing. In contrast, we estimate the probabilities of having such segregated networks to be actually much larger than the random benchmarks would suggest. For example, eighteen percent (as opposed to 1 in a 1000) would know 10 or fewer unemployed persons, 1/3 would know 10 or fewer Asians, and 17% would know 10 or fewer gay people.<sup>15</sup> To put it another way, segregated networks in terms of each of these social groups is much more common than would be expected if people mixed without regard to the statuses or behaviors that define these groups.

[Table 4 about here.]

McPherson et al. (2001) summarized the sparse knowledge about acquaintanceship networks to the effect that “in relationships of less closeness, religion may not matter much at all.” (p. 426). While this may be true if religion is operationalized as denomination, our results show that perceived segregation by *religiosity* (i.e., the frequency of attendance at places of worship) is at roughly the same level as perceived segregation on the basis of class or race. All three of these variables have overdispersions that are on the order of 10. It is, of course possible that regular church-goers are simply ignorant of the behavior of acquaintances not in their congregations, while those who rarely go to church are simply unaware of the behavior of their church-going acquaintances.<sup>16</sup> We think it unlikely, however, that associational segregation on the basis of perceived religious behaviors would be nearly as high as associational segregation on the basis of race if religious behavior were not an important factor structuring interaction even among acquaintances. We estimate that the chances of knowing no one (or thinking that one doesn’t know anyone) who goes to church regularly, no one who is unemployed, no one who is gay, no one who cohabits, no one who is strongly liberal, or no one who is strongly conservative is always at least 5 times and as much as 11 times higher in American social networks than would be true under random mixing. Our results suggest a polyvalent pattern of segregation in American social networks which challenges the conventional wisdom that “race and ethnicity are clearly the biggest divides in social networks” (McPherson et al., 2001, p. 6).<sup>17</sup>

Table 3 shows that the pattern of segregation varies across subnetworks. Naturally, race and ethnicity are most highly segregated within families, where integration occurs only either through intermarriage, or through members of mixed-race and mixed-ethnic families assuming different racial or ethnic identities. Outside of the family, race and ethnic segregation are generally of comparable size within the neighborhood, voluntary associations, and the workplace, with acquaintances involving blacks being somewhat less overdispersed at work than in neighborhoods. It is, of course, well known that residential segregation in the U.S. tends to be pronounced, and segregation in neighborhood-based acquaintances is therefore not a surprise. It is also well known

that schools, churches, and social organizations are highly segregated by race. The average census tract-level index of black-white dissimilarity in the 50 largest metropolitan areas of the U.S. is .62, while the average tract-level Hispanic-white index of dissimilarity is .48 (Charles, 2003). Our recent knowledge about workplace segregation derives from EEO-1 data on private establishments with 50 or more employees (Robinson et al., 2005). Tomaskovic-Devey et al. (2006) found that American establishments had a mean white-black dissimilarity index of about .35 and a similarly sized white-Hispanic dissimilarity index.<sup>18</sup> However, they argue that this number is an underestimate, first because it excludes establishments that are racially homogeneous, and second because it is based on the highly aggregated EEO nine-category occupational classification. In contrast, Hellerstein and Neumark (2008) ignored occupation and estimated a dissimilarity index for blacks and whites of .19 in a sample of establishments with at least 40 employees based on the 1990 long form census data merged with the Business Register.<sup>19</sup>

Table 3 suggests lower levels of associational segregation involving African-Americans at work than in the neighborhood. Segregation involving Hispanic or Asian neighborhood acquaintances is clearly lower than is segregation involving black acquaintances. Segregation involving Hispanic or Asian acquaintances through voluntary associations is similarly lower than is segregation involving black acquaintances. We do not have the data to disentangle the various associational contexts within which Americans mix, but certainly religious activities play a major role. It is well known that religious congregations are highly segregated by race (Dougherty, 2003; Vischer, 2001), though little hard evidence exists to support the speculation that segregation at church is greater for blacks than for other racial groups. Whatever its cause, these gradients by racial group deserve further investigation.

The second striking pattern in Table 3 is the extent to which “bridging” social capital is more likely to be found within families than in the associations and business organizations that make up the public sphere. There is less overdispersion in knowing the unemployed or people with a second home in the family than at work, within associations, or in neighborhoods. The same is true for prisoners. Acquaintanceship ties with gays are also less segregated within the family

than at work, in associations, or in neighborhoods. This pattern may partly be explained by the fact that American families have become more heterogeneous over time, and therefore it is more likely that people who are dissimilar with respect to class or prison status will be located in the same family than in the past. It is probably also harder to ignore or be misinformed about statuses or behaviors within the family than it is at work, in associations, or even in the neighborhood. In other words, the greater amount of shared information about family members may produce a closer correspondence between the diversity of networks as they really are and as they appear to ego in the family than in more public contexts. Neither explanation diminishes the irony of this striking finding.

The two names “Mark” and “Linda” show a greater level of overdispersion at work and in associations than in neighborhoods or families. We suspect that this pattern reflects greater age-segregation at the workplace or in associations than would be found in neighborhoods or families, where people of different ages are likely to interact with each other. Mark, for example, was the sixth most popular boys name for cohorts born in the 1960s, but ranked 181st in the 1930s and 34th in the 1980s. Meanwhile, Linda ranked 2nd in the 1940s, 317th in the 1920s, and 128th in the 1980s. When names change popularity over time, more highly age-segregated networks will show greater overdispersion than will less age-segregated networks.

Perceived segregation of acquaintances by church attendance or political ideology are about equally segregated in the family, in the neighborhood, and at work. Glaeser and Ward (2006) estimated that the index of dissimilarity by political party at the national level is about .2 when counties are the unit of analysis. This is much lower than standard results for residential segregation at the tract level, but these numbers are not readily comparable. Counties are much bigger than tracts, and county-level racial segregation is doubtless much lower than is tract-level segregation. However, racial segregation *within* counties is very high, while the level of political segregation within counties is an unknown. Religiosity is much more segregated within associations than at work or in the neighborhood, but this is not surprising given that the category of associations includes places of worship. Political ideology is similarly more segregated within voluntary associations than it is

at the workplace or in the neighborhood. Certainly it is not the case that political associations are a central aspect of the associational life of Americans, but people appear to choose associations or choose whom to associate with in associations in order to produce a greater level of perceived ideological segregation than they experience in their neighborhoods or workplaces. The high level of overdispersion by political ideology in voluntary associations that are officially organized on other principles could be a product of consolidation (Blau, 1977), i.e., where one dimension of belief or behavior (e.g., religious belief or religiosity) is highly correlated with another belief (political ideology). It is, of course, also possible that people more readily attribute their beliefs to others in voluntary associations that – members may rightly or wrongly assume – bring together other people with similar beliefs to their own.

## **Trust Networks**

The number of individuals that one trusts is obviously smaller than the number of people that one is acquainted with, but how much smaller? As noted above, McPherson et al., 2008 (see also McPherson et al., 2006; McPherson et al., 2009) found that the mean size of core networks (as measured by the GSS question concerning a list of people one has "discussed important matters with" in the last six months) dropped from 2.9 out of a maximum of 5 in 1985 to 2.1 in 2004, with 22.5% of the sample listing no names at all. Our 2006 GSS trust question differs from the 2004 (and 1985) GSS questions; it broadens the relationship to include friends, and it is closer to the Coleman idea of trust as the willingness to place material resources along with information at the disposal of someone else. For these reasons, it provides an alternative perspective on the level of isolation among contemporary Americans. We computed the proportion of people in our sample who reported that they trusted no one at all in any of the social categories that we asked about (i.e., all the specific names, all the specific occupations, all races, liberals and conservatives, churchgoers and non-churchgoers, the unemployed, those in prison, those with a second house, gays, and cohabiting women). Only 1.4% of the 2006 GSS sample reported that they did not trust any specific person in any of these categories that we queried about, which is very different from

the 2004 GSS. We further computed the proportion of respondents who did not trust anyone in all but one of these categories (we let the excepted category be anything at all). This relaxed criterion only raised the proportion of "extremely low trusters" to 3.1%. It seems that when confronted with specific prompts for specific types of people, Americans are much more likely to report that they trust at least some specific individual than they are to provide the specific name of someone with whom they have discussed "important matters."

Our estimate for the degree distribution of the trust network is displayed in Figure 3. The distribution of trust ties is skewed to the right, with an estimated median of 17 and an estimated interquartile range between 10 and 26.<sup>20</sup> These estimates are much higher than the mean of 2.1 reported out of the 2004 GSS (McPherson et al., 2006) and more in line with the estimates obtained from the 2006 Pew survey (Boase et al., 2006). Our results suggest that the multiple prompts in the 2006 "trust" question wording (good friends, people you discuss important matters with, trust for advice, or trust with money) and the lack of a six-month scope condition in the 2006 question generate a larger, less close network than does the 2004 GSS question wording. At the same time, trust networks as measured by the 2006 GSS question wording are much smaller than acquaintanceship networks; our estimate of the median number of people in the close networks tapped by our trust question is only 3% of our estimate of the median number of people in acquaintanceship networks.<sup>21</sup>

[Figure 3 about here.]

To establish the determinants of the size of the trust network, we first estimated a fractional polynomial regression of the estimated size of the trust network against the estimated size of the acquaintanceship network. Figure 4 shows the estimated relationship between the number known and the predicted number trusted along with a scatterplot of the estimated number trusted against the estimated number known. Among those whose estimated acquaintanceship degree is in the bottom 25% of the distribution, the predicted number trusted moves from about 5 to about 15, with virtually everyone in this quartile trusting fewer than 20 people. In the middle 50% of the distribution, the expected number trusted climbs from about 15 to about 25. In this range, it

becomes more common for people to report that they trust between 20 and 40 people, even though there is a persisting minority of respondents who trust very few individuals. Finally, in the top quartile, the expected number trusted climbs from 25 to over 40. A minority of people assert that they trust over 60 people, while another minority report that they trust very few individuals despite their large acquaintanceship network.

[Figure 4 about here.]

We next regressed the estimated number trusted on a set of covariates, and we report the answers in Table 5. In Model 1, we omit acquaintance degree. The pattern of coefficients in the trust model is similar to that reported earlier for the acquaintanceship model as well as to analyses of other close network data (McPherson 1983; Marsden 1987; McPherson et al. 2006), and reinforces the conclusion that the predictors of social network size are robust across tie strength and across different strategies for measuring social networks. In model 2, we include the estimated size of one's acquaintanceship network as a covariate. Model 2 suggests that education and church attendance mostly affected the number trusted because of their effect on the number of acquaintances, while the effect of other race or foreign born is diminished. Net of estimated degree size, age appears to have a curvilinear relationship with trust: young adults over 25 and people over 65 trust a higher proportion of their acquaintances than do people of other ages. Model 3 includes the generalized trust variable.<sup>22</sup> In the absence of any other covariates except for degree size, generalized trust has a significant effect on the number trusted; those who think that people mostly can be trusted trust an estimated 15% more people (net of estimated degree size) than do people who disagree that most people can be trusted (results available upon request from the authors). In the presence of other covariates, however, the effect of generalized trust on the degree of one's trust network is weakened below the conventional threshold of statistical significance.<sup>23</sup> When generalized trust as well as degree size are controlled, church attendance again becomes a significant predictor of the size of one's trust network; net of other factors, those who attend church weekly or more trust about 20% more people than do those who never go to church. We speculate that these

additional people in the trust network are in fact the people that churchgoers go to church with, but we do not have the data to confirm this.

[Table 5 about here.]

Next, we address the question of overdispersion in trust networks. The fourth and fifth columns of Table 3 show the level of overdispersion in the trust networks and Table 6 illustrates the impact of overdispersion by comparing the probability of trusting no one in our salient groups as compared with the expected outcome under random mixing. As with acquaintanceship networks, overdispersion is highest for racial groups, but church attendance follows closely behind. Under random mixing, only 9% of people would be expected not to know of any specific African-American that they trust. In the actual data, we estimate that 53% of the population knows no African-American that they trust, 52% percent knows no Hispanic that they trust, and nearly 80% knows no Asian that they trust. The effects of overdispersion similarly magnify the likelihood of trusting no one in groups defined by religiosity or political ideology relative to the baseline random mixing model. While only 9% of the population would be expected not to trust a single liberal under random mixing, our actual estimated probability is 40%. We estimate that 29% of the population do not know any specific conservative person that they trust; in a world of random mixing, this number would be only 4%.

[Table 6 about here.]

We elaborate our analysis of racial segregation in trust networks in Table 7 by comparing the actual frequencies of trusting people of other races that we obtained from the GSS. Other studies have reported that it is relatively common for blacks and whites to report significant contact with members of the other race. In a 1989 national survey, 82% of blacks and 66% of whites claimed to have friends of the other race (Sigelman and Welch, 1993). Jackman and Crane (1986) reported results from a 1975 national sample that showed 10% of whites to have a close black friend, another 21% with a black acquaintance, and 25% of blacks with a close white friend.<sup>24</sup> Sigelman

et al. (1996) reported from their 1992 Detroit survey that 43% of blacks and 27% of whites said that they had a good friend of the other race. Marsden's (1987) study of the 1985 GSS social network questions found that only 8% of adults with networks of size two or more reported being tied to someone of a different race.<sup>25</sup> As Table 7 shows, 37% of whites claim to trust 2 or more blacks, and 28% claim to trust 2 or more Hispanics in the 2006 GSS, while a small majority of blacks and a larger majority of people of other races report that they trust two or more whites. Meanwhile, nearly half of American whites report no blacks in their trust networks, and about a third of blacks report no whites in their trust networks. The GSS data suggest greater levels of interracial contact in 2006 than Jackman and Crain found in 1975, but less than Sigelman et al. (1996) found in 1989. The first conclusion we draw from this comparison is that estimates of interracial ties are sensitive to the method of measurement. Our second conclusion is that trust networks in the United States remain highly segregated.

[Table 7 about here.]

As we argued earlier in the paper, little is known about the relative level of segregation of trust networks vs broader acquaintanceship networks. On theoretical grounds, McPherson et al. (2001) predicted that homophily is stronger in what they refer to as "multiplex" relationships, in which people have a relationship along more than one dimension. One corollary of this is that trust networks should be more homophilous than are acquaintanceship networks, because one is likely to have a more elaborated structure of ties involving kinship, marriage, and friendship in addition to more instrumental connections with people that one trusts than with people that are only acquaintances. Similarly, Putnam (2000) conjectured that "bonding" ties tend to be with people like oneself; his question was whether bridging ties would be sufficiently heterophilous to create a socially integrated society. A comparison of the estimated overdispersion in the acquaintanceship and trust results provides a simple test of this conjecture. In fact, our estimated overdispersions are generally smaller for trust networks than for acquaintanceship networks. From the perspective of ego, trust networks show less variation along key status and values dimensions than do acquaintanceship networks.

To some extent, the larger overdispersion estimates for acquaintanceship networks may reflect variation in recall errors among GSS respondents. Thus, one can see from Table 3 that the estimated overdispersion of names is generally larger (by about 25%) in the acquaintanceship network than in the trust network. However, the estimates of overdispersion in acquaintanceship for the substantively interesting groups generally exceed the estimates of overdispersion in trust by more than 25%. More work is needed to understand the impact of recall error on estimates of segregation in social networks, but our estimates suggest that perceived acquaintanceship networks are at least as segregated as perceived trust networks in contemporary American society.

## Discussion

Segregation in American social networks is pervasive across multiple statuses that have been identified as dimensions of potential social cleavage in the popular press and in the academic literature. Other studies have found this to be true in the context of core networks. Our data confirm that segregation is also pervasive in broader acquaintanceship networks as well. Beyond this confirmation, our data support three major conclusions that constitute a mixed message for those concerned about social integration. On the optimistic side, we find that trust networks are larger than the discussion networks estimated with the 2004 GSS and are about the same size as the close networks estimated with the 2006 Pew survey. The typical American is able to identify between 10 and 20 individuals that he trusts. About a quarter of Americans trust fewer than 10 individuals, and these Americans typically have relatively few acquaintances as well. At the other extreme are the small but not insignificant group of Americans who have a large number of acquaintances but trust very few of them. The typical American has a trusting relationship with only about 1/30th of the people that he or she is acquainted with. This may sound low, but building a trusting relationship takes time, and most people may not have enough time in their lives to build more than twenty or so such relationships.

The greater concern, we suggest, lies not with the size of trusting relationships but rather with

the structure of acquaintanceship networks, which are perceived by ego to be as segregated as trust networks. To say that core networks are homophilous is almost a truism. However, the rhythms of modern life often provide the opportunity to interact with others who are different from oneself. This opportunity is of course not a social constant, it depends upon social resources that provide the possibility to choose where one lives, where one works, and which associations one is able to join. Within these constraints, people exercise choices about workplace, place of residence, and about associational participation. People also have at least some control over the people they get to know in these various settings. When social barriers are high, people of different races or with different political views or religious orientations may avoid social interaction to the extent possible or at least may hide social differences from those whom they must work with or see on a regular basis. Structural opportunity mixes with personal preferences to shape the diversity of one's acquaintances, colleagues, coworkers, and associates.

Core networks are different. People are socialized to be like their family members, and they choose their mates and their friends. It is for this reason that one expects homophily to be high in core networks. That acquaintanceship networks are at least as segregated as are core networks has, we suggest, two potentially important implications. The first, which is consistent with concerns raised by Putnam, Skocpol, and others, is that the organizations of American civil society in the American economy do not play a strongly integrative role in contemporary American society. A second potentially important implication is that new forces in American society may provide the basis for increased integration in the "bounded solidarity" group known as the American family. One of these factors is rising rates of interracial marriage, and another is the relatively high rate of instability of both cohabitation and marriage, which increases the rate of repartnering at older ages and thereby lowers marital homogamy (Schwartz and Mare, 2005). The impact of these trends is magnified by the relative difficulty of hiding one's religious orientation, sexual orientation, political orientation, or cohabitation behavior from other family members. It is also harder to ignore or misperceive the statuses, behaviors, and values of family members than it is for the statuses, behaviors, and values of associates and casual acquaintances in the neighborhood, at

work, or in voluntary associations. Growing heterogeneity combines with willing or unwilling transparency to produce a surprising level of integration in family interactions across multiple important social dimensions.

Our third major finding is the large magnitude of the segregation on important socioeconomic, behavioral, and values dimensions. The estimated level of perceived segregation by race in association networks is roughly on par with the level of perceived segregation by religious behavior, employment status, and political ideology. Religion in particular has emerged as a fundamental cleavage in American society at the level of day-to-day interaction. From the perspective of the culture wars that we have seen play out in the American political sphere and the past decade or so, this may not be surprising. However, it is often assumed that the most visible participants in these culture wars are a relatively small number of partisans. Instead, we find that Americans differ greatly in their perceived ties to people from the more secular and the more religious wings of American society. The same is true for political orientation.

Religiosity and political orientation are more difficult to observe than race, and so “objective” levels of segregation on these dimensions are probably not as high as people report in the General Social Survey. But perceptions shape lived experience, and sharp differences in the experienced social worlds of Americans may impede understanding and tolerance for the views and lifestyles of those who are different than oneself. We cannot, of course, measure the extent to which the “objective” and “perceived” acquaintanceship networks differ from each other. Therefore, we cannot know whether the high segregation in acquaintanceship networks comes from structural factors that objectively segregate Americans into different social groups, from self-selection processes, or from a combination of masking and misperception that cause America’s acquaintanceship networks to be more different from one another in terms of experience than in terms of actual fact. Nonetheless, our findings point to trust networks, the rough equivalent of “bonding” social capital, as providing an important complement to weak ties in maintaining social integration in American society. One cannot readily hide behaviors and values in close networks, and this fact, coupled with the growing heterogeneity of American families, suggests that families and the close friends

associated with them are less about “narrow identifies” and “out-group antagonisms” than Putnam feared them to be.

Aside from technical issues concerning measurement and model specification, there are important substantive questions raised by our results. One such issue concerns the extent to which our measured levels of segregation are driven by the objective characteristics of the people that Americans know, and the extent to which they are driven by misperception or masking of behaviors and opinions that Americans think would be disapproved of by their associates. A second important issue concerns trends over time. While our study provides a baseline for the assessment of future trends, our limited comparisons with previous studies provide some grounds for concluding that segregation in association by race may be diminishing or at least is not increasing. We have no firm basis for drawing any similar conclusions concerning segregation by religious behavior, political orientation, sexual orientation or the other variables measured in the 2006 GSS. Future data collections can provide the basis for comparisons with existing data to establish a level of stability and change in segregation of social networks along these dimensions. A final issue concerns the causes and consequences of network segregation. The General Social Survey provides a good platform for collecting descriptive information about social networks and for studying the behavioral correlates of network structure. However, causal estimates involving these network characteristics cannot readily be obtained from these data, and imaginative strategies are needed in order to determine the individual and structural factors that can explain heterogeneity in segregation across individuals and over time. These are important topics for future research.

# Appendix A

## Likelihood Computation

As noted in the Data and Methods Section, we used intervals (zero, one, 2-5, 6-10, or greater than 10) to ask respondents about people they know. The model originally presented in Zheng et al. (2006) was designed for exact counts. Although the general structure of the model remains the same, some computational modifications are necessary to adapt the Zheng et al. (2006) model for interval data.

From the Data and Methods Section, recall that our model takes the form

$$y_{ik} \sim \text{negative binomial}(\text{mean} = a_i b'_k, \text{overdispersion} = \omega_k) \quad (3)$$

where  $a_i$  is the degree of respondent  $i$  and  $b'_k$  is the prevalence of group  $k$ .  $b'_k$  is adjusted using the calibration curve presented in in the next section.

We fit the model in Equation 3 using Bayesian inference. We assume that the log of the respondent degree parameters,  $\log(a_i)$ , follow a normal distribution with mean  $\mu_a$  and standard deviation  $\sigma_a$ . Similarly, the log of the group parameters,  $\log(b_k)$  are assumed to follow normal distributions with mean  $\mu_b$  and standard deviation  $\sigma_b$ . In both cases the hyperparameters are given noninformative uniform priors. The overdispersion parameters,  $\omega_k$ , are assumed to follow independent Uniform(0,1) distributions on the inverse scale. Since overdispersion can fall in the range  $(0, \infty)$  the inverse,  $1/\omega_k$ , is in  $(0, 1)$ . This prior specification performed well for Zheng et al. (2006) and is consistent with observations in McCarty et al. (2001).

The full posterior distribution is then  $p(a, b, \mu_a, \mu_b, \sigma_a, \sigma_b | y)$ . Since our values of  $y_{ik}$  are intervals, we can partition the posterior based on these categories. Say that, given the option, the respondent would report that she/he knows an exact count of  $z_{ik}$  individuals in group  $k$ . Then, let  $\ell$  be the an indicator of the interval that an observation  $z_{ik}$  belongs to. Then, there are  $L$  intervals, one for each level of  $y_{ik}$ , with each interval containing one or more potential values of  $z_{ik}$ . For example,

if a respondent knew three members of group  $k$  ( $z_{ik} = 3$ ) she would report  $y_{ik} = "2 - 5"$ , which corresponds to  $\ell = 3$ , the third interval. For clarity, let  $y_{ik(\ell)}$  be the interval of  $y_{ik}$  that corresponds to level  $\ell$ . Our likelihood is expressed as

$$\begin{aligned}
p(a, b, \mu_a, \mu_b, \sigma_a, \sigma_b | z) &\propto \sum_{\ell=1}^L \sum_{z_{ik} \in y_{ik(\ell)}} \prod_{i=1}^n \prod_{k=1}^K \binom{z_{ik} + \eta_{ik} - 1}{\eta_{ik} - 1} \left(\frac{1}{\omega_k}\right)^{\eta_{ik}} \left(\frac{\eta_{ik} - 1}{\omega_k}\right)^{z_{ik}} \\
&\times \prod_{i=1}^n N(\log(a_i) | \mu_a, \sigma_a^2) \prod_{k=1}^K N(\log(b_k) | \mu_b, \sigma_b^2) \\
&\times \prod_{z_{ik} \in y_{ik(\ell)}}
\end{aligned}$$

where  $\eta_{ik} = e^{a_i b'_k / (\omega_k - 1)}$  and  $\mathbb{I}_{z_{ik} \in y_{ik(\ell)}}$  is an indicator variable taking value one if the observation is in group  $\ell$  and zero otherwise. The final interval (greater than 10) has an unlimited number of possible  $z_{ik}$  values. This is not problematic since we can equivalently perform the computation for  $z_{ik} \in [0, 10]$  and subtract from one. Estimation is then carried out using Markov-chain Monte Carlo (MCMC) in a fashion similar to Zheng et al. (2006).

## Calibration Curve

In this section we give additional details about the motivation, derivation, and application of the calibration curve. Killworth et al. (2003) documents that respondents have difficulty recalling accurately their ties in large subpopulations and proposes several mechanisms to explain the under-recall. One possible explanation is a process that Killworth et al. (2003) calls "dredging," whereby a respondent recalls one-by-one the first  $m$  acquaintances and then estimates for all groups larger than some size  $m$ . This mechanism would, in theory, produce accurate responses for small groups (less than  $m$  acquaintances) but less reliable responses for larger groups where respondents are estimating total group size rather than counting specific acquaintances (McCarty et al., 2001). Though this mechanism seems plausible, there is no specific process for determining  $m$  or modeling how estimating rather than enumerating would impact the overall accuracy of the results. Additionally, both Killworth et al. (2003) and McCarty et al. (2001) point out that the relatively short time given

to answer each question likely creates difficulty for respondents and is confounded with “dredging.”

Like the Zheng et al. (2006) model, our model has a nonidentifiability since the likelihood depends on  $\log(a_i)$  and  $\log(b_k)$  only through their sum. For clarity, let  $a_i = e^{\alpha_i}$  and  $b_i = e^{\beta_k}$ . To identify the  $\alpha$ 's and  $\beta$ 's the model is renormalized by adding a constant to all  $\alpha'_i$ s and subtracting the constant from the  $\beta'_k$ s. One intuitive way of calculating the renormalizing constant is to set

$$\sum e^{\beta_k} = \sum \{\text{population proportion}\}_k. \quad (4)$$

This is equivalent to assuming that the average degree of individuals in these subpopulations equals the average degree of the population. Obviously, this assumption does not apply to all of the subpopulations in our current survey. When restricted to the subpopulations defined by the first names, however, this assumption is fairly reasonable.

[Figure 5 about here.]

The above strategy also requires that the acquaintance ties recorded in the survey reflect the distribution of ties in the social network. However, the survey did not accurately measure the social network but rather the *recalled* social network by the respondents. Figure 5 gives a graphical representation of the distinction between a respondent's actual and recalled networks. For rare groups, the respondents can recall almost all their ties with these groups, indicated by the left side of Figure 5. The number of ties to a large subpopulation  $k$  is under-recalled. This under-recalling is represented in Figure 5 by the increasing discrepancy between the circles corresponding to the recalled and actual respondent network as the size of the alter group increases. The estimated proportion  $e^{\beta_k}$  from data therefore only estimate the proportion of ties involving subpopulation  $k$  in the recalled social network. Consequently,

$$\begin{aligned} \sum e^{\beta_k} &= \sum g(\{\text{population proportion}\}_k) \\ &\leq \sum \{\text{population proportion}\}_k. \end{aligned}$$

Here,  $g(\cdot)$  represents the *recall* function. If the renormalizing constant is computed based on equation (4) and some popular first names, the degrees of the respondents will be underestimated.<sup>26</sup>

Let  $e^{\beta_k}$  be the proportion of ties in the social network that involve individual in subpopulation  $k$ . And let  $e^{\beta'_k}$  denote the proportion of ties in the *recalled* social network that involve subpopulation  $k$ . Assume  $\beta'_k = f(\beta_k)$  and  $f(\cdot)$  is an increasing function.

Based on our observation and also independent discussion by Killworth et al. (2003), we assume that

$$\begin{aligned} f'(x) &\rightarrow 1 \quad \text{as } e^x \rightarrow 0 \quad (x \rightarrow -\infty) \\ &\rightarrow \frac{1}{2} \quad \text{as } e^x \rightarrow 1 \quad (x \rightarrow 0). \end{aligned}$$

To simplify the inference, we assume that  $f(x) = x$  for small populations with proportion as small as  $e^x = e^{c_1}$  ( $c_1 < 0$ ) and  $f'(x)$  decreases as  $x$  increases (at most) to  $\frac{1}{2}$  as  $x$  goes to zero. More specifically, we assume

$$f'(x) = \frac{1}{2} + \frac{1}{2}e^{-c_2(x-c_1)}, c_2 \geq 0, \text{ for } x \geq c_1,$$

where  $c_2$  controls how fast and how close  $f'(x)$  approaches  $\frac{1}{2}$ .

This gives us

$$f(x) = c_1 + \frac{1}{2}(x - c_1) + \frac{1}{2c_2} \left( 1 - e^{-c_2(x-c_1)} \right).$$

In this paper, we use  $c_1 = -7$ , which corresponds to subpopulations that are  $< .1\%$  of the population and  $c_2$  is to be fitted using  $\beta_k$  originally estimated and the population proportions of first names. This is because, as discussed earlier, we assume that in the absence of recall bias,  $\beta_k \approx \{\text{population proportion}\}_k$  on average. Incidentally, we found that an  $c_2$  of approximately one yielded the best fit.

[Figure 6 about here.]

The names used in our current survey represent subpopulations that are much smaller than those used in the survey presented in the McCarty et al. (2001) paper (The data were actually

collected between 1998 and 1999). That earlier survey included the name Michael, which represents 1.8% of the population. Someone whose personal network size is 600 is expected to know  $600 \times .018 \approx 11$  Michaels. Though imaginable, it is difficult to recall 11 Michaels during the limited amount of time of such a survey; therefore, the actual reported count is likely to be much lower. In fact, in the McCarty et al. (2001) data, respondents reported knowing an average of just under 5 Michaels. In contrast, the six names used here represent only 1.4% of the population with the largest names, Karen and Keith, representing about .34% each. Nonetheless, we still observe some under-recalling among respondents, particularly for these two names. Figure 6 shows that, particularly for the larger names, using the calibration curve improves the estimates of the fractional subpopulation size. We intentionally chose names that were less popular than those used in the previous survey, but not so rare that most respondents wouldn't have any contact with members of the subpopulation. This is mirrored with a larger issue discussed in further detail in McCormick et al. (2008) and McCormick and Zheng (2007).

## Notes

<sup>1</sup>Not all scholars agree with Putnam that social capital has declined, including Ladd (1996) and Wuthnow (1998). Costa and Kahn (2003) analyzed trend data on social capital in multiple datasets including the DDB Life Style Surveys, the Current Population Surveys, the General Social Surveys, the National Election Studies and time diary studies conducted at multiple points in time. Costa and Khan reported that some measures of social capital declined over time, while others did not. There was no strong trend in rates of volunteering across the multiple datasets that they studied. GSS data show the strongest trend in membership organizations involved religious organizations. Membership in professional organizations actually rose considerably, while in other nonchurch organizations, membership rates changed very little. Costa and Khan's analysis of time-trend data agrees with Bianchi et al. (2006) in finding declines in socializing time with friends and relatives, though much of this decline appears to involve the frequency of interaction rather than the existence of ties per se.

<sup>2</sup>Wuthnow (2002, 2003) also finds that religious involvement does not have a net effect on having friends with lower status or with higher status people. Ties to higher status people, in contrast, do tend to be higher for those who are members of religious congregations or who have leadership positions in these congregations.

<sup>3</sup>Tilly's recent definition of trust is similar; according to Tilly: "Trust consists of placing valued outcomes at risk to others' malfeasance, mistakes or failures (Tilly, 2005, p. 12).

<sup>4</sup>We used the following names: Karen, Brenda, Kevin, Shawn, Keith, Rachel, Mark, Linda, Jose, and Maria. While the estimated level of overdispersion with these names was relatively low, no names are truly neutral because they vary in frequency by birth cohort and ethnicity, and these "barrier effects" will bias the estimate of degree size (Salganik et al., 2008). To take the most obvious example, the popularity of specific names varies by ethnic group. To determine the size of this bias, we estimated the data alternatively including and deleting the two Hispanic names (Jose and Maria). The results were highly similar. To illustrate, the mean posterior mean of the acquaintanceship networks differed by less than 0.25% when we alternatively included and excluded the Hispanic names, and the estimated acquaintanceship overdispersions varied at most by 4% across the groups analyzed in this paper, which was considerably smaller than the standard errors for these estimates.

<sup>5</sup>The question wording was of the form: "how many are you pretty certain" are gay men or women, or attend religious services on a regular basis, or are strongly liberal etc.

<sup>6</sup> Social networks tend to be relatively gender-integrated, which is another reason for our excluding gender as a potential dimension of segregation (McPherson et al., 2001).

<sup>7</sup>The average person in the McCarty et al. data reported knowing 600 persons (McCormick and Zheng, 2007). Someone with a personal network of 600 would be expected to know about 11 persons named Michael. However, respondents reported knowing an average of just under 5 Michaels.

<sup>8</sup>One reviewer pointed out that if the roughly forty percent of GSS respondents who reported high generalized trust in fact trusted all of their acquaintances while those with low generalized trust trusted few of their respondents, the average size of the trust network would be close to the 220 estimate that we obtained using the names generator plus the recall function. Arguing against this interpretation is the fact that the specific prompts in the trust question imply a behavioral connection, not a willingness to believe that one's acquaintances are trustworthy. Also arguing against this interpretation are the findings of Boase et al. (2006), which are similar to the estimates we obtain using the race groups normalization. Finally, the empirical pattern in the data is not consistent with the interpretation that generalized trusters trust their acquaintances while low trusters do not. If we use the high names-based estimate for the size of trust networks, we find that the median number trusted for those with high generalized trust is much smaller than the median number of acquaintances for this subgroup, and is only moderately larger than the median number trusted for those who respond that people in general can *not* be trusted. We conclude that the race-based normalization provides a more reasonable estimate of the size of trust networks as measured by the prompt used in the 2006 GSS.

<sup>9</sup>The estimates we provide included the Hispanic names in the normalization. As noted above, our estimates differ by a trivial amount if we exclude these names from the estimation procedure.

<sup>10</sup>The Zheng et al estimate of 610 for the Kilworth and McCarty data was larger than McCarty et al's own estimate of 290 (at the mean) McCarty et al. (2001) because Zheng et al used a recall correction, and because Zheng et al normalized using the rarer names from the Kilworth and McCarty data (McCarty et al normalized using common names from the data). Our estimate of 550 is also similar to that obtained by McCormick et al. (2008), who used a more sophisticated approach to take barrier effects into account related to the different distribution of names across birth cohorts.

<sup>11</sup>In supplementary models, we also included measures of the natural logarithm of size of place and dummy variables for region. Size of place does not have a significant effect on the size of acquaintanceship networks, net of the other covariates in the model. Net of other covariates, inhabitants of New England tended to have larger acquaintanceship networks, while those in middle and south Atlantic states tended to have relatively small acquaintanceship networks.

<sup>12</sup>We also estimated models with separate effects for black and Hispanic and with a separate effect for foreign born. The point estimates for black and U.S. born Hispanic were similar. The coefficients for foreign-born Hispanics, other races, U.S. born and other races, foreign born were also similar, and therefore we combined race and foreign-born categories in the more parsimonious model presented here. The more parsimonious model also more clearly shows the differences by race and foreign born that are statistically significant at conventional levels. At the same time, we acknowledge that the acquaintanceship networks of members of other races may be underestimated because the names of their racial/ethnic groups are not represented in the name prompts that we used in the GSS survey. More precise information about these groups would require a more survey with a larger sample size or with an oversample of those

in the other race category.

<sup>13</sup>The highest response category for our questions was "more than 10." Almost everyone knows more than 10 whites, and so we have relatively little information about overdispersion for this group. Because we did not assume a hierarchical model for the overdispersion parameters themselves, the imprecise estimate for the white group does not affect the estimates for the other groups.

<sup>14</sup> If anything, these estimates probably underestimate the actual overdispersion, in that the majority of acquaintances of many blacks may also be black.

<sup>15</sup>The estimated number of people in a 400 person network who belongs to any particular social group is of course greater than the estimated number of people that one would *recall* from a 400 person network. The illustration could equally well have been worked out for the recalled network as for the total network, and the results would be the same, with the caveat for both cases that the overdispersion refers to what ego *knows* about the people in his network rather than what these people know about themselves.

<sup>16</sup>Some of the overdispersion in ties to those who attend church regularly arises from the fact that, as we showed earlier, regular church goers tend to have larger acquaintanceship networks, and these "extra" acquaintances are homophilous with respect to church attendance. In other words, religiosity raises the level of segregation of social networks by making the networks of church goers bigger in a non-random way. A similar process would be at work if the "extra" acquaintances that one has by virtue of being highly educated or well paid tend to be like oneself. Perhaps bankers tend to know incrementally more rich people by the nature of their job, while the incremental acquaintances that doctors have from their medical practices better approximate random mixing. Whatever the process that determines the size and characteristics of networks, the overdispersion parameters express the extent of segregation in these networks as perceived by ego.

<sup>17</sup>We noted earlier that the highest non response rates were for the religiosity and political ideology questions. If non responders to the religiosity and political ideology did not answer the question because they did not know whether any of their acquaintances were in a specific category, their missing answers could be interpreted as not knowing anyone who they were sure fit the description. In such a case, our estimated overdispersions underestimate true overdispersion in ties to people that one perceives as belong to these categories.

<sup>18</sup>Roughly 15% of establishments were missing either blacks or whites and roughly 20% of establishments were missing either Hispanics or whites (Tomaskovic-Devey et al., 2006).

<sup>19</sup>Racial or ethnic segregation by job is conceptually quite different from racial or ethnic acquaintanceship at work, because people potentially interact both vertically (i.e., between superiors and subordinates) and horizontally at the workplace.

<sup>20</sup>If we used the names normalization for trust networks along with recall correction, we would estimate the posterior mean of the median number trusted to be a too-high 220 as opposed to the 17 we estimate when using racial groups

to perform the normalization. As noted above, the names we used to estimate degree are too rare to provide a precise estimate for trust networks, and produce upwardly biased estimates both because of the tendency for people to “over-recall” ties with rare groups and because recall bias is – we argue– not as great for smaller trust networks as for larger acquaintanceship networks. An inspection of Table 7 demonstrates the basis for this much larger estimate. In the raw data, 44% of whites reported that they trusted 10 or fewer specific white people, 67% of blacks reported that they trusted 10 or fewer specific black people, and 76% of those of other races reported that they trusted 10 or fewer specific Hispanic people. With such high proportions of the three major race groups having relatively small trust networks, the estimate of a 220 median seems implausible. We believe the estimate of 17 is closer to the truth though probably a lower-bound on the correct answer. As noted above, our estimate of the degree size has no impact on our estimates of overdispersion in connections with the various population groups contained in our survey data.

<sup>21</sup>The large difference in the estimated size of trust and acquaintanceship networks suggests that respondents correctly reported about *specific* trust relations rather than about generalized trust; about one-third of GSS respondents reported in the abstract that most people can be trusted, which presumably would have included the people that they themselves were acquainted with.

<sup>22</sup>NORC asked the generalized trust question to approximately 2/3 of the GSS sample that was also asked our questions about trust, and so the sample size for model 3 is smaller than for models 1 and 2.

<sup>23</sup>However, respondents with high generalized trust know an estimated 70 more people than do those with low generalized trust. Generalized trust is related to the number one trusts partly through its association with the number one knows.

<sup>24</sup>Jackman and Crain’s data used a “stronger” form of acquaintanceship than used in our data. Their prompt defined acquaintanceship as people that respondents “keep in touch with or get together with occasionally.” It seems likely that many people who would be defined as acquaintances based on knowing their name and stopping on the street to say hello are not people that one keeps in touch with or gets together with occasionally.

<sup>25</sup>Marsden estimated this frequency as only one-seventh as high as one would expect if people sorted themselves at random.

<sup>26</sup>Zheng et al. (2006) observed that the estimated average degree is 384 if using all 12 names to normalize, but 739 when normalizing only on the rarer names.

## References

- Allison, P. (1978). Measures of inequality. *American Sociological Review*, 43:865–880.
- Baldassari, D. and Bearman, P. (2007). Dynamics of political polarization. *American sociological review*, 72:784–811.
- Baldassarri, D. and Gelman, A. (2008). Partisans without constraint: political polarization and trends in american public opinion. *American Journal of Sociology*, 114(408–446).
- Bearman, P., Moody, J., Stovel, K., and Thalji, L. (2004). Social and sexual networks: The national longitudinal study of adolescent health. In Morris, M., editor, *Network Epidemiology: A Handbook for Survey Design and Data Collection*. Oxford University Press, Oxford, UK.
- Bell, W. (1953). Probability Model for the Measurement for Ecological Segregation, A. *Social Forces*, 32:357.
- Bianchi, S. M., Robinson, J. P., and Milkie, M. A. (2006). *Changing Rhythms of American Family Life*. Russell Sage Foundation, New York.
- Billy, J. O., Rodgers, J. L., and Udry, J. (1984). Adolescent sexual behavior and friendship choice. *Social Forces*, 62(3):653–678.
- Blau, P. (1977). *Inequality and Heterogeneity*. The Free Press, New York.
- Boase, J., Horrigan, J., Wellman, B., and Rainie, L. (2006). The strength of internet ties. *Pew Internet and American Life Project*, [www.pewinternet.org/pdfs/PIP\\_internet\\_ties.pdf](http://www.pewinternet.org/pdfs/PIP_internet_ties.pdf).
- Bourdieu, P. (1980). Le capital social. *Actes de la recherche en sciences sociales*, 31(31):2–3.
- Brooks, C. (2002). Religious Influence and the Politics of Family Decline Concern: Trends, Sources, and US Political Behavior. *American Sociological Review*, 67(2):191–211.
- Burtless, G. (1999). Effects of growing wage disparities and changing family composition on the US income distribution. *European Economic Review*, 43(4-6):853–865.

- Charles, C. Z. (2003). The dynamics of racial residential segregation. *Annual Review of Sociology*, 29(1):167–207.
- Coleman, J. (1988). Social Capital in the Creation of Human Capital. *The American Journal of Sociology*, 94:95–120.
- Coleman, J. S. (1957). *Community Conflict*. Free Press, New York.
- Coleman, J. S. (1961). *The adolescent society*. Free Press.
- Coleman, J. S. (1990). *Foundations of social theory*. Harvard University Press, Cambridge, MA.
- Costa, D. L. and Kahn, M. E. (2003). Understanding the american decline in social capital, 1952–1998. *Kyklos International Review for Social Sciences*, 56(1):17–46.
- Council of the European Union (2004). Joint report by the commission and the council on social exclusion. Technical Report 7101/04.
- DiMaggio, P., Evans, J., and Bryson, B. (1996). Have American's Social Attitudes Become More Polarized? *American Journal of Sociology*, 102(3):690–755.
- Dougherty, K. (2003). How monochromatic is church membership? Racial-ethnic diversity in religious community. *Sociology of religion*, 64(1):65–85.
- Ellwood, D. T. and Jencks, C. (2004). The Uneven Spread of Single Parent Families: What Do We Know? Where Do We Look for Answers? In Neckerman, K., editor, *Social Inequality*, pages 3–78. Russell Sage Foundation, New York.
- Evans, J. H. (2003). Have Americans' Attitudes Become More Polarized?-An Update. *Social Science Quarterly*, 84(1):71–90.
- Feld, S. L. (1981). The focused organization of social ties. *American Journal of Sociology*, 86(5):1015.
- Fischer, C. (1982). *To Dwell Among Friends*. University of California Press, Berkeley.

- Fischer, C. (2009). The 2004 GSS Finding of Shrunken Social Networks: An Artifact? *American Sociological Review*, 74(4):657–669.
- Frank, T. (2004). *What's the Matter with Kansas?: How Conservatives Won the Heart of America*. Metropolitan Books, New York.
- Frey, W. H. and Myers, D. (2005). Racial segregation in us metropolitan areas and cities, 1990-2000: Patterns, trends, and explanations. Population Studies Center Research Report 05-573, University of Michigan.
- Fukuyama, F. (1995). *Trust*. Free Press, New York.
- Gambetta, D. (1988). *Trust: making and breaking cooperative relations*. B. Blackwell, New York.
- Gelman, A., Park, D., Shor, B., Bafumi, J., and Cortina, J. (2008). *Red State, Blue State, Rich State, Poor State: Why Americans Vote the Way They Do*. Princeton University Press, Princeton, NJ.
- Gittell, R. J. and Vidal, A. (1998). *Community Organizing: Building Social Capital As a Development Strategy*. Sage Publications, Thousand Oaks, CA.
- Glaeser, E. L. and Ward, B. A. (2006). Myths and Realities of American Political Geography. *Journal of Economic Perspectives*, 20(2):119–144.
- Goel, S., Mason, W. A., and Watts, D. J. (2009). Diversity, deliberation, and awareness of political attitudes in social networks. Technical report, Yahoo Research.
- Granovetter, M. (1973). The strength of weak ties. *American Journal of Sociology*, 78(6):1360–1380.
- Green, J. C. (1996). *Religion and the Culture Wars: Dispatches from the Front*. Rowman & Littlefield, Lanham, MD.

- Gullickson, A. (2006). Education and Black-White Interracial Marriage. *Demography*, 43(4):673–689.
- Hampton, K. and Wellman, B. (2001). Long distance community in the network society: Contact and support beyond netville. *American Behavioral Scientist*, 45(3):477–497.
- Hellerstein, J. and Neumark, D. (2008). Workplace segregation in the united states: Race, ethnicity, and skill. *Review of Economics and Statistics*.
- Iceland, J., Weinberg, D. H., and Steinmetz, E. (2002). *Racial and Ethnic Residential Segregation in the United States: 1980-2000*. U.S. Census Bureau, Bensus Special Report CENSR-3. U.S. Government Printing Office.
- Jackman, M. R. and Crane, M. (1986). "some of my best friends are black...": Interracial friendship and whites' racial attitudes. *Public Opinion Quarterly*, 50:459–486.
- Kalmijn, M. (1998). Intermarriage and homogamy: causes, patterns, trends. *Annual Review of Sociology*, 24:395–421.
- Killworth, P. D., McCarty, C., Bernard, H. R., Johnsen, E. C., Domini, J., and Shelley, G. A. (2003). Two interpretations of reports of knowledge of subpopulation sizes. *Social Networks*, 25(2):141–160.
- Ladd, E. C. (1996). The data just don't show erosion of America's social capital. *The Public Perspective*, 7(4):1–22.
- Laumann, E. O. (1973). *Bonds of Pluralism: The Form and Substance of Urban Social Networks*. Wiley, New York.
- Lee, C.-S. (2007). The institutional origin of social trust: The configuration of social policies and coordination of interests in rich democracies. Unpublished manuscript.
- Lieberson, S. (1981). An asymmetrical approach to segregation. *Ethnic segregation in cities*, pages 61–82.

- Lin, N. (1999). Social networks and status attainment. *Annual Reviews in Sociology*, 25(1):467–487.
- Lin, N., Fu, Y., and Hsung, R. (2001). The Position Generator: Measurement Techniques for Investigations of Social Capital. *Social Capital: Theory and Research*, pages 57–81.
- Lipset, S. (1963). *Political Man: The Social Bases of Politics*. Doubleday, New York.
- Louch, H. (2000). Personal network integration: transitivity and homophily in strong-tie relations. *Social Networks*, 22(1):45–64.
- Marsden, P. (1988). Homogeneity in confiding relations. *Social Networks*, 10:57–76.
- Marsden, P. (2005). Recent developments in network measurement. In Carrington, P., Scott, J., and Wasserman, S., editors, *Models and Methods in Social Network Analysis*, pages 8–30. Cambridge University Press.
- Marsden, P. V. (1987). Core discussion networks of americans. *American Sociological Review*, 52(1):122–131.
- Massey, D. S. and Denton, N. A. (1988). The dimensions of residential segregation. *Social Forces*, 67(2):281–315.
- Massey, D. S. and Denton, N. A. (1993). *American Apartheid: Segregation and the Making of the Underclass*. Harvard University Press, Cambridge, MA.
- Mayhew, B. H., McPherson, M., Rotolo, T., and Smith-Lovin, L. (1995). Sex and race homogeneity in naturally occurring groups. *Social Forces*, 74(1):15–52.
- McCarty, C., Killworth, P. D., Bernard, H., Johnsen, E. C., and Shelley, G. A. (2001). Comparing Two Methods for Estimating Network Size. *Human Organization*, 60(1):28–39.
- McCarty, N., Poole, K. T., and Rosenthal, H. (2006). *Polarized America: The dance of ideology and unequal riches*. The MIT Press, Cambridge, MA.

- McCormick, T., Salganik, M. J., and Zheng, T. (2008). How many people do you know?: Efficiently estimating personal network size. *Working paper*.
- McCormick, T. H. and Zheng, T. (2007). Adjusting for recall bias in "how many x's do you know?" surveys. *JSM Proceedings*.
- McPherson, M. (1983). An ecology of affiliation. *American Sociological Review*, 48(4):519–532.
- McPherson, M. and Smith-Lovin, L. (1987). Homophily in voluntary organizations: status distance and the composition of face-to-face groups. *American Sociological Review*, 52:370–379.
- McPherson, M., Smith-Lovin, L., and Brashears, M. (2009). Models and Marginals: Using Survey Evidence to Study Social Networks. *American Sociological Review*, 74(4):670–681.
- McPherson, M., Smith-Lovin, L., and Brashears, M. E. (2006). Social isolation in america: Changes in core discussion networks over two decades. *American Sociological Review*, 71:353–375.
- McPherson, M., Smith-Lovin, L., and Brashears, M. E. (2008). Erratum: Social isolation in america: Changes in core discussion networks over two decades. *American Sociological Review*, 73:1022.
- McPherson, M., Smith-Lovin, L., and Cook, J. M. (2001). Birds of a feather: Homophily in social networks. *Annual Review of Sociology*, 27:415–444.
- Merton, R. (1957). *Social Theory and Social Structure*. The Free Press of Glencoe, Glencoe, IL.
- Moody, J. (2001). Race, school integration, and friendship segregation in america. *American Journal of Sociology*, 107(3):679–716.
- Mouw, T. and Entwisle, B. (2006). Residential segregation and interracial friendship in schools. *American Journal of Sociology*, 112:394–441.

- Mouw, T. and Sobel, M. E. (2001). Culture Wars and Opinion Polarization: The Case of Abortion 1. *American Journal of Sociology*, 106(4):913–943.
- Newcomb, T. (1961). *The Acquaintanceship Process*. Holt, Rinehart & Winston, New York.
- Noelle-Neumann, E. (1993). *The spiral of silence: Public opinion, our social skin*. University of Chicago Press.
- Paxton, P. (2007). Association memberships and generalized trust: A multilevel model across 31 countries. *Social Forces*, 86:27–76.
- Poole, K. T. and Rosenthal, H. (2000). *Congress: A Political-Economic History of Roll Call Voting*. Oxford University Press, USA.
- Portes, A. (1998). Social Capital: Its Origins and Applications in Modern Sociology. *Annual Review of Sociology*, 24(1):1–24.
- Portes, A. (2000). The Two Meanings of Social Capital. *Sociological Forum*, 15(1):1–12.
- Portes, A. and Sensenbrenner, J. (1993). Embeddedness and Immigration: Notes on the Social Determinants of Economic Action. *American Journal of Sociology*, 98(6):1320.
- Putnam, R. (1996). The strange disappearance of civic America. *The American Prospect*, 24(1):34–48.
- Putnam, R. (2000). *Bowling Alone: The Collapse and Revival of American Community*. Simon and Schuster, New York.
- Putnam, R. D. (1993). The Prosperous Community: Social Capital and Public Life. *The American Prospect*, 13(1):35–42.
- Quillian, L. and Campbell, M. (2003). Beyond black and white: The present and future of multiracial friendship segregation. *American Sociological Review*, 68:540–566.

- Rindfuss, R. R., Jampaklay, A., Entwisle, B., Sawangdee, Y., Faust, K., and Prasartkul, P. (2004). The collection and analysis of social network data in nang rong, thailand. In Morris, M., editor, *Network Epidemiology: A Handbook for Survey Design and Data Collection*, pages 175–200. Oxford University Press.
- Robinson, C. L., Taylor, T., Tomaskovic-Devey, D., Zimmer, C., and Irvin Jr, M. (2005). Studying Race or Ethnic and Sex Segregation at the Establishment Level: Methodological Issues and Substantive Opportunities Using EEO-1 Reports. *Work and Occupations*, 32(1):5.
- Rosenthal, H. (2004). Politics, Public Policy, and Inequality: A Look Back at the Twentieth Century. In Neckerman, K., editor, *Social Inequality*, pages 861–892. Russell Sage Foundation, New York.
- Salganik, M., Zheng, T., and McCormick, T. H. (2008). How many people do you know? efficiently estimating network size. Working Paper, Columbia University.
- Schwartz, C. R. and Mare, R. D. (2005). Trends in Educational Assortative Marriage from 1940 to 2003. *Demography*, 42(4):621–646.
- Sen, A. K. (1992). *Inequality Reexamined*. Harvard University Press, Cambridge, MA.
- Sigelman, L., Blesdoe, T., Welch, S., and Combs, M. W. (1996). Making contact? black-white social interaction in an urban setting. *American Journal of Sociology*, 101:1306–1332.
- Sigelman, L. and Welch, S. (1993). The contact hypothesis revisited: Black-white interaction and positive racial attitudes. *Social Forces*, 71:781–795.
- Skocpol, T. and Fiorina, M. P. (1999). *Civic Engagement in American Democracy*. Brookings Institution Press, Washington, DC.
- Smith-Lovin, L. (2007). The Strength of Weak Identities: Social Structural Sources of Self, Situation and Emotional Experience. *Social Psychology Quarterly*, 70(2):106–124.

- Thomas, W. and Thomas, D. S. (1928). *The child in America; behavior problems and programs*. A A Knopf, New York.
- Tilly, C. (2005). *Trust and Rule*. Cambridge University Press, Cambridge.
- Tomaskovic-Devey, D., Zimmer, C., Stainback, K., Robinson, C., Taylor, T., and McTague, T. (2006). Documenting desegregation: Segregation in American workplaces by race, ethnicity, and sex 1966-2000. *American Sociological Review*, 71:565–88.
- Van der Gaag, M. and Snijders, T. (2005). The Resource Generator: social capital quantification with concrete items. *Social Networks*, 27(1):1–29.
- Verbrugge, L. M. (1977). The structure of adult friendship choices. *Social Forces*, 56:576–597.
- Verbrugge, L. M. (1983). A research note on adult friendship contact: a dyadic perspective. *Social Forces*, 62:78–83.
- Vischer, R. (2001). Racial Segregation in American Churches and Its Implications for School Vouchers. *Florida Law Review*, 53:193.
- Waldinger, R. (1995). The ‘other side’ of embeddedness: a case-study of the interplay of economy and ethnicity. *Ethnic and Racial Studies*, 18(3):555–580.
- Wellman, B. (1996). Are personal communities local? a dumptarian reconsideration. *Social Networks*, 18:347–54.
- Woolcock, M. (1998). Social capital and economic development: Toward a theoretical synthesis and policy framework. *Theory and Society*, 27(2):151–208.
- Wuthnow, R. (1998). *Loose Connections: Joining Together in America’s Fragmented Communities*. Harvard University Press, Cambridge, MA.
- Wuthnow, R. (2002). Religious Involvement and Status-Bridging Social Capital. *Journal for the Scientific Study of Religion*, 41(4):669–684.

- Wuthnow, R. (2003). Overcoming Status Distinctions? Religious Involvement, Social Class, Race, and Ethnicity in Friendship Patterns. *Sociology of Religion*, 64(4):423–442.
- Yamagishi, T. and Yamagishi, M. (1994). Trust and commitment in the united states and japan. *Motivation and Emotion*, 18(2):129–166.
- Zheng, T., Salganik, M., and Gelman, A. (2006). How many people do you know in prison?: Using overdispersion in count data to estimate social structure in networks. *Journal of the American Statistical Association*, 101:409–423.

Table 1: Groups Included in the 2006 GSS Queries about Social Ties based on Acquaintanceship and Trust

How many people are you acquainted with/do you trust who...	Occupations	Social Groups
	Are Police Officers	Are currently unemployed
	Are Lawyers	Own a second home
	Are Social Workers	Are currently in state/federal prison
	Are Janitors or Building Cleaners	Asian or Asian-American
	Are Child Care Workers	Black or African/American
	Are Electricians	Hispanic men or women
	Are currently serving in the Armed Forces	Gay men or women
		Unmarried women living with men in a romantic relationship
		Attend religious services on a regular basis
		Attend religious services rarely or never
		People who are strongly liberal
		People who are strongly conservative

Table 2: Regression of Acquaintanceship Degree on Selected Covariates

	Acquaintanceship Degree		Log of Degree	
	Coefficient	Standard Error	Coefficient	Standard Error
Age 30-64	23.9	45.3	0.047	0.06
Age 66+	-43.8	57.0	-0.039	0.08
highest year of school completed	21.931**	5.3	0.033**	0.01
total family income (in 10K units)	9.524*	4.1	0.015*	0.01
Income is missing	61.6	48.5	0.063	0.07
female	-35.5	29.7	-0.040	0.04
black or Hispanic, U.S. born	-62.9	39.7	-0.079	0.06
Foreign born or other race	-146.9*	63.1	-0.26**	0.09
attend church sometimes	51.6	35.0	0.093	0.05
attend church weekly or more	149.3**	39.3	0.25**	0.05
moderate political views	3.2	43.0	-0.012	0.06
conservative political views	-78.6	51.6	-0.11	0.07
widowed	-30.8	60.5	-0.11	0.08
divorced	53.1	43.0	0.076	0.06
separated	-32.7	94.9	0.042	0.13
never married	5.1	41.9	0.018	0.06
_cons	265.4**	97.2	5.7**	0.13
N	647		647	
R <sup>2</sup>	.12		.14	

Table 3: Estimated Overdispersions for Acquaintanceship and Trust Networks

	Acquaintanceship Subnetworks											
	Acquaintances		Trust		Work		Associations		Neighborhood		Family	
	Median	IQR	Median	IQR	Median	IQR	Median	IQR	Median	IQR	Median	IQR
Kevin	1.7	(0.1)	1.0	(0.1)								
Karen	1.6	(0.2)	1.1	(0.1)								
Shawn/Sean	1.6	(0.2)	1.2	(0.1)								
Brenda	1.5	(0.1)	1.4	(0.2)								
Keith	1.2	(0.1)	1.2	(0.1)								
Rachel	1.5	(0.2)	1.4	(0.2)								
Mark	1.7	(0.2)	1.2	(0.1)	2.1	(0.3)	2.3	(0.4)	1.5	(0.3)	1.0	(0.1)
Linda	1.3	(0.1)	1.2	(0.1)	2.0	(0.3)	2.0	(0.3)	1.5	(0.2)	1.2	(0.1)
Jose	3.4	(0.5)	2.1	(0.3)								
Maria	2.5	(0.3)	2.5	(0.4)								
Asians	8.2	(1.3)	5.6	(1.0)	9.0	(1.8)	6.5	(1.4)	8.1	(2.1)	15.2	(6.4)
Blacks	10.7	(1.7)	6.8	(0.8)	10.7	(2.0)	13.4	(2.7)	14.8	(3.5)	78.6	(33.3)
Hispanics	8.8	(1.3)	7.2	(1.2)	12.8	(2.2)	8.7	(1.4)	10.6	(2.2)	24.6	(8.2)
Whites	44.5	(12.3)	9.9	(1.6)	29.7	(10.7)	44.0	(20.1)	29.5	(10.2)	209.4	(68.8)
Unemployed	10.3	(1.5)	5.3	(0.9)	14.9	(4.4)	10.6	(2.3)	12.5	(3.3)	5.3	(0.9)
Own second homes	4.1	(0.6)	3.0	(0.5)	5.7	(1.2)	8.0	(1.8)	3.6	(0.8)	2.6	(0.4)
In prison	3.7	(0.7)	2.8	(0.9)	4.0	(1.5)	7.6	(3.4)	12.6	(8.0)	3.4	(0.9)
Gay men or women	5.7	(0.8)	4.0	(0.6)	4.5	(0.7)	4.7	(0.9)	4.6	(1.1)	2.5	(0.5)
Women who are cohabiting	6.1	(0.9)	4.0	(0.5)	7.7	(2.0)	14.7	(3.2)	7.7	(1.6)	9.6	(2.2)
Attend church regularly	11.5	(2.2)	7.3	(1.0)	10.9	(2.9)	18.7	(4.9)	9.0	(2.0)	8.2	(1.5)
Attend church rarely/never	11.5	(2.0)	6.2	(0.8)	7.1	(1.7)	7.9	(1.9)	6.2	(1.2)	6.1	(0.9)
Strongly liberal	7.9	(1.3)	5.4	(0.6)	9.4	(2.3)	9.5	(2.6)	7.6	(1.6)	9.0	(1.7)
Strongly conservative	8.3	(1.1)	5.2	(0.7)	5.3	(0.9)	8.4	(1.6)	4.3	(1.0)	4.4	(0.9)

Table 4: Deviation from Random in 400 Person Acquaintanceship Networks

Persons who (are)	Expected Count	Probability of knowing $\leq 10$		
		Random	Estimated	Odds Ratio
Unemployed	24	0.00	0.18	202
Own second homes	24	0.00	0.06	55
In prison	4	1.00	0.91	0
Asians	17	0.05	0.33	9
Blacks	48	0.00	0.01	>1000
Hispanics	52	0.00	0.00	>1000
Whites	291	0.00	0.00	>1000
Gay men or women	20	0.01	0.17	19
Women who are cohabiting	17	0.05	0.28	7
Attend church regularly	125	0.00	0.00	>1000
Attend Church Rarely/Never	168	0.00	0.00	>1000
Strongly liberal	60	0.00	0.00	>1000
Strongly conservative	78	0.00	0.00	>1000

Table 5: Regression of the Logarithm of Estimated Trust Degree on Selected Covariates

	Model I	Model II	Model III
Age 25-34	0.17	0.21*	0.30*
Age 35-44	0.19	0.15	0.16
Age 45-54	0.11	0.11	0.16
Age 55-64	0.13	0.10	0.033
Age 66+	0.23	0.28**	0.23
highest year of school completed	0.031**	0.008	0.002
total family income (in 10K units)	0.01	0.001	-0.003
Income is missing	0.025	-0.025	-0.081
female	-0.078	-0.052	0.006
Black or Hispanic, U.S. born	-0.14*	-0.072	-0.08
Other race or Foreign born	-0.38**	-0.20*	-0.23*
attend church sometimes	0.15*	0.078	0.12*
attend church weekly or more	0.28**	0.096	0.20**
moderate political views	-0.038	-0.019	0.018
conservative political views	-0.049	0.045	0.052
widowed	-0.17	-0.11	-0.098
divorced	0.015	-0.037	-0.019
separated	-0.09	-0.095	0.02
never married	0.065	0.047	0.069
estimated acquaintance degree/100		0.27**	0.29**
(estimated degree/100)**2		-0.013**	-0.015**
(estimated degree/100)**3		0.000**	0.000*
Cannot trust most people			-0.10
Whether one can trust "depends. . ."			-0.12
Intercept	2.2**	1.4**	1.39**
Number of observations	642	642	415

Table 6: Deviation from Random in Median Size (17) Person Trust Network

Persons who (are)	Expected Count	Probability of Trusting No One		
		Random	Estimated	Odds Ratio
Unemployed	1	0.36	0.67	3.6
Own Second Homes	1	0.36	0.57	2.4
In Prison	0	0.84	0.91	1.8
Asians	1	0.49	0.76	3.4
Blacks	2	0.13	0.51	7.0
Hispanics	2	0.11	0.50	8.0
Whites	12	0.00	0.04	>1000
Gay Men or Women	1	0.43	0.67	2.8
Women who are Cohabiting	1	0.24	0.51	3.3
Attend Church Regularly	5	0.00	0.19	47
Attend Church Rarely/Never	7	0.00	0.08	109
Strongly Liberal	3	0.08	0.37	7.0
Strongly Conservative	3	0.04	0.27	9.9

Table 7: Distribution of Trust of Other Races, by Own Race

		Own Race		
Number of		White	Black	Other
Whites trusted	0	3.6	31.0	19.7
	1	3.6	16.1	14.8
	2-5	20.4	35.6	34.4
	6-10	16.7	3.5	11.5
	11+	55.8	13.8	19.7
N		504	87	61
Blacks trusted	0	48.0	13.6	52.4
	1	14.9	4.6	15.9
	2-5	25.0	26.1	23.8
	6-10	7.1	22.7	3.2
	11+	5.0	33.0	4.8
N		504	88	63
Hispanics trusted	0	59.9	64.4	38.7
	1	12.5	11.5	8.1
	2-5	20.2	20.7	17.7
	6-10	3.8	1.2	11.3
	11+	3.8	2.3	24.2
N		506	87	62

Figure 1: GSS Sample Design

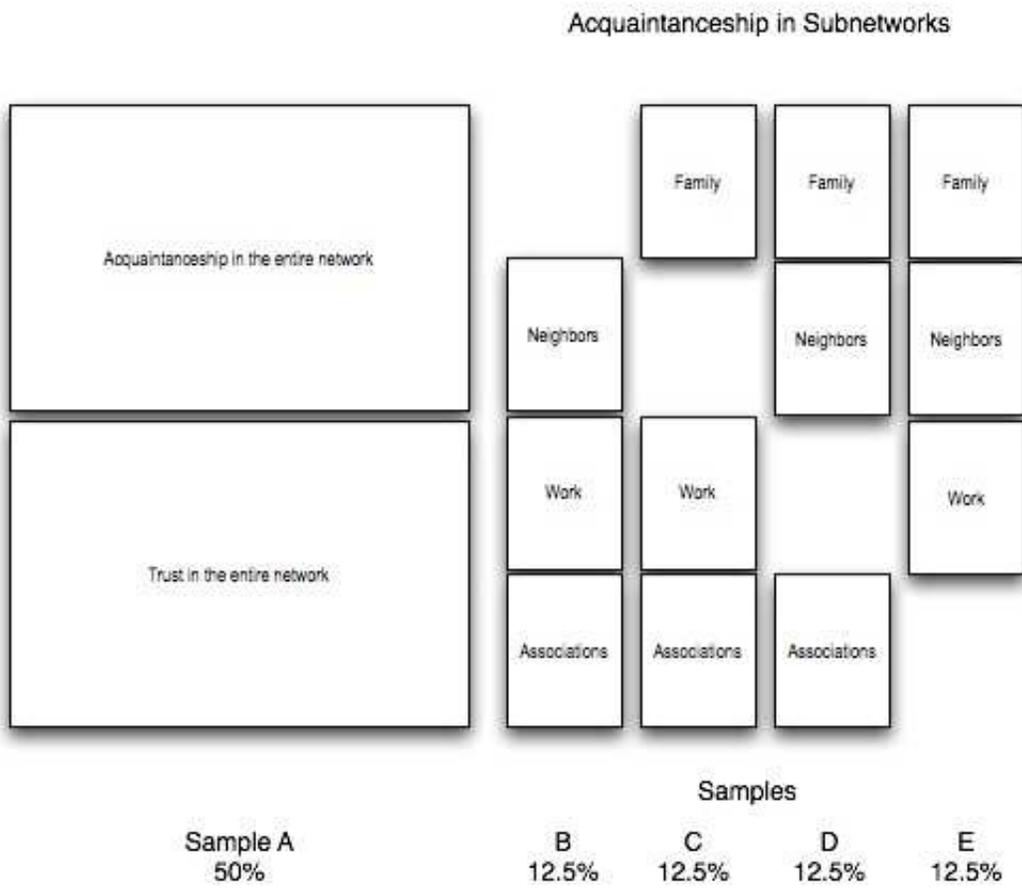


Figure 2: Estimated Distribution of Number of Acquaintanceships

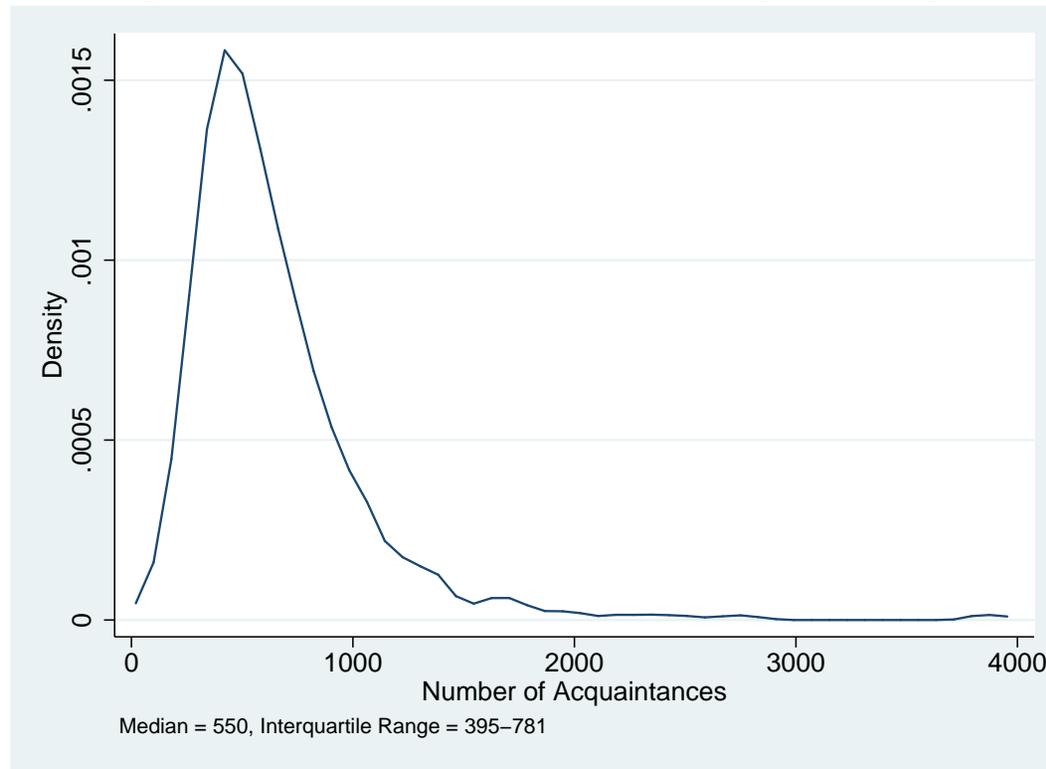


Figure 3: Estimated Distribution of Number Trusted

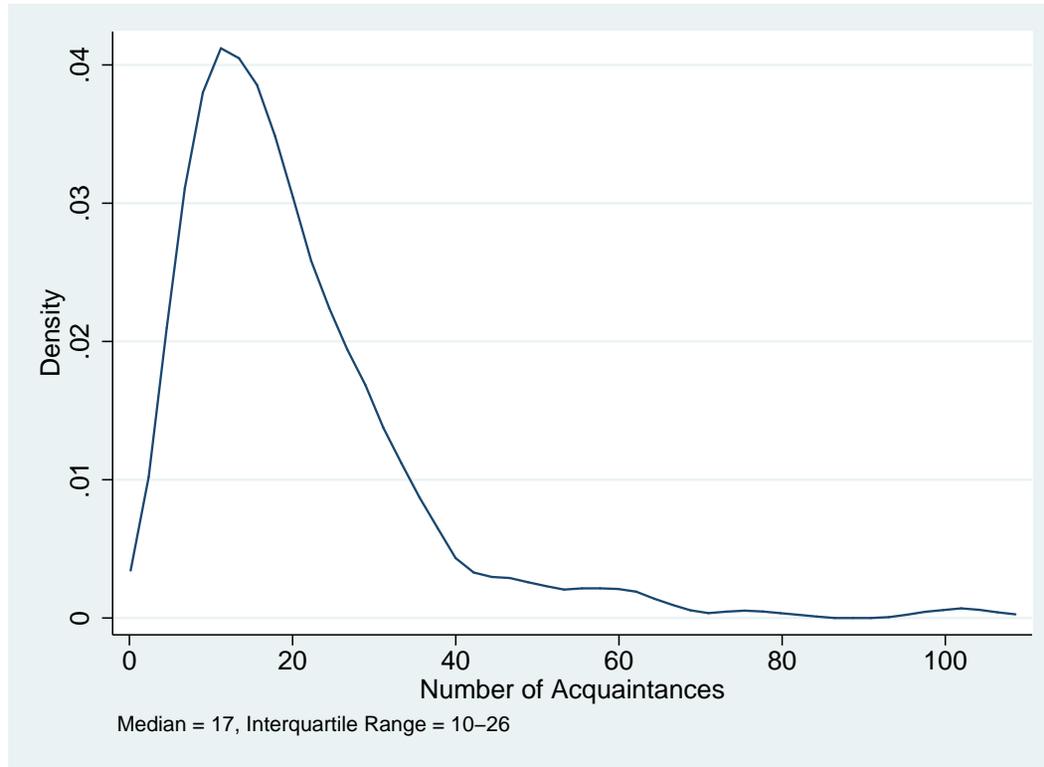
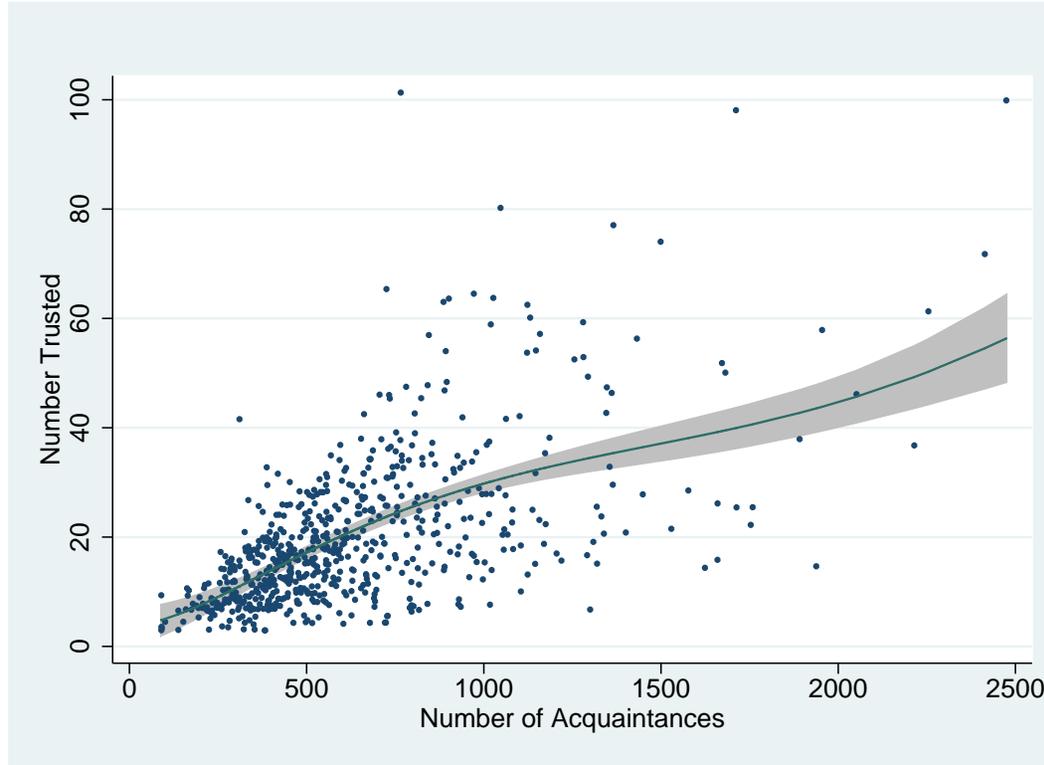


Figure 4: Number Trusted v. Number Known



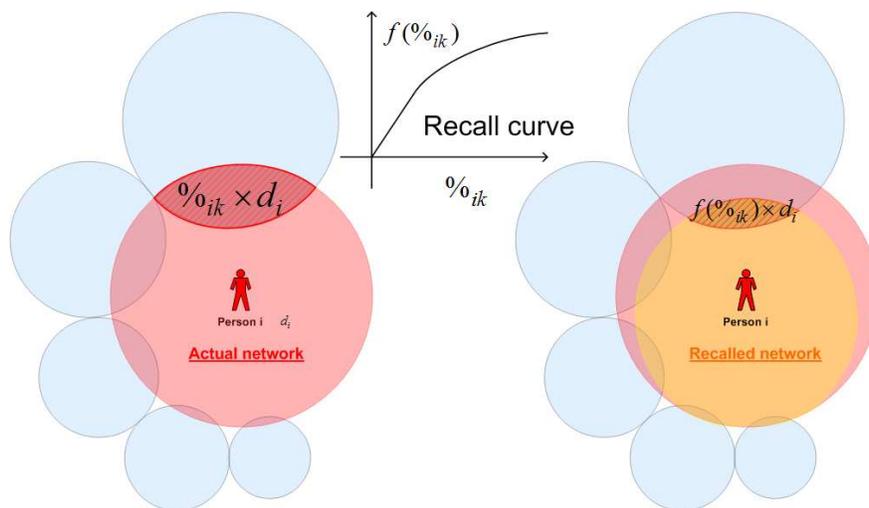


Figure 5: Comparing the actual network to the recalled network for a sample respondent. Outer circles represent potential alter groups and have radii proportional to the size of the alter group. For smaller alter groups, the difference between the recalled network and the actual network is small. As the alter group size increases, however, the amount of recall bias becomes more significant. The calibration curve, illustrated at the top of the diagram, addresses this issue.

### Effects of the Calibration Curve

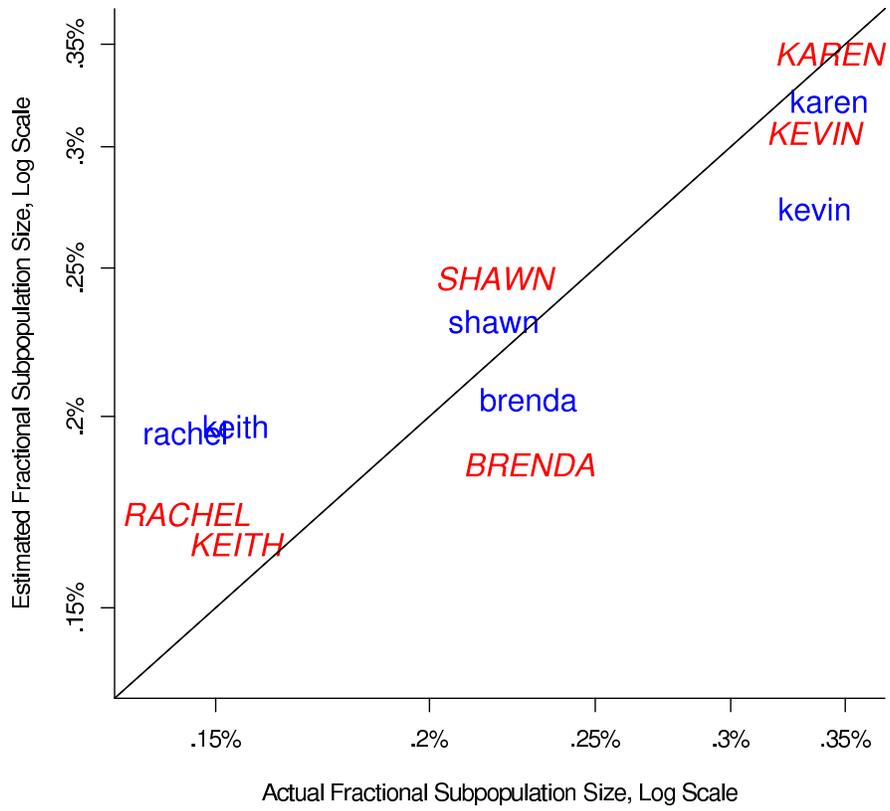


Figure 6: Estimated fractional subpopulation size with and without the calibration curve. The solid line is the  $y = x$  line. Names written in capital letters represent estimates using the calibration curve. Lowercase letters are estimates without the calibration curve.