4-2014

Social Distance in the United States: Sex, Race, Religion, Age, and Education Homophily among Confidants, 1985 to 2004

Jeffrey A. Smith
University of Nebraska-Lincoln, jsmith77@unl.edu

Miller McPherson
Duke University, mcphersn@soc.duke.edu

Lynn Smith-Lovin
Duke University, smithlov@soc.duke.edu

Published in *American Sociological Review* 79:3 (2014), pp. 432-456.
doi: 10.1177/0003122414531776

Published online April 24, 2014.

Social Distance in the United States: Sex, Race, Religion, Age, and Education Homophily among Confidants, 1985 to 2004

Jeffrey A. Smith,[1] Miller McPherson,[2] and Lynn Smith-Lovin[2]

1. University of Nebraska-Lincoln
2. Duke University

Corresponding author – Jeffrey A. Smith, Department of Sociology, University of Nebraska-Lincoln, 711 Oldfather Hall, Lincoln, NE 68588-0324, email jsmith77@unl.edu

Abstract

Homophily, the tendency for similar actors to be connected at a higher rate than dissimilar actors, is a pervasive social fact. In this article, we examine changes over a 20-year period in two types of homophily—the actual level of contact between people in different social categories and the level of contact relative to chance. We use data from the 1985 and 2004 General Social Surveys to ask whether the strengths of five social distinctions—sex, race/ethnicity, religious affiliation, age, and education—changed over the past two decades in core discussion networks. Changes in the actual level of homophily are driven by the demographic composition of the United States. As the nation has become more diverse, cross-category contacts in race/ethnicity and religion have increased. After describing the raw homophily rates, we develop a case-control model to assess homophily relative to chance mixing. We find decreasing rates of homophily for gender but stability for race and age, although the young are increasingly isolated from older cohorts outside of the family. We also find some weak evidence for increasing educational and religious homophily. These relational trends may be explained by changes in demographic heterogeneity, institutional segregation, economic inequality, and symbolic boundaries.

Keywords: social distance, social networks, homophily, social structure, social change

Birds of a feather have always flocked together. Building on classic works by Simmel and Park, Borgardus (1925) coined the term "social distance" to indicate whether people in one social category were willing to be closely associated with members of another category. His social distance scale used questions about the acceptability of marriage, entertainment in the home, co-residence in neighborhoods, and other sorts of affiliations. Decades of research that followed have informed our understanding of the cognitive prejudices present in the population (e.g., Hughes and Tuch 2003). With the increasing availability of network data, however, more recent work has increasingly employed actual patterns of interaction to measure social distance (McPherson, Smith-Lovin, and Cook 2001). The question of social distance has thus become more structural, reflecting the social acceptability of affiliation *and* the physical opportunities for interacting.

This article follows the behavioral trend and uses homophily as a summary measure of social distance across time and demographic dimensions. Homophily captures the tendency for similar actors to be socially connected at a higher rate than dissimilar actors; it is one of our best established social facts (Lazarsfeld and Merton 1954; McPherson et al. 2001). Arguably, it is one of our most important. The top two most-cited articles in the *Annual Review of Sociology* both deal with networks, their structure, and their impact on the flow of resources (i.e., Portes 1998 and McPherson et al. 2001, as reported by *Annual Review of Sociology* 2013).

Homophily is important because it measures the salience of sociodemographic features in our social system (Blau and Schwartz 1984; Laumann 1966). A socially unimportant demographic dimension will exhibit low levels of homophily: social boundaries will be porous and individuals will be free to form intimate social ties with members of another group. Homophily can be seen as a behavioral expression of the larger differentiating forces in society—such as demographic availability, institutional segregation, and affective acceptance among categories of people. The size of demographic groups, for example, influences the probability that individuals will come into contact with each other by chance, and potentially overcomes the propensity for in-group association (Blau 1977). Demographic change is limited as an integrating force, however, by residential and occupational segregation, as well as status differences across a population.

Homophily in networks is also important because ideas, resources, and group affiliations flow through networks (McPherson, Popielarz, and Drobnic 1992). Close confidants influence us directly through their supportive interactions (House, Umberson, and Landis 1988; Wellman and Wortley 1990) and indirectly by shaping the kinds of people we become (Smith-Lovin and McPherson 1993). If we are connected mainly to people much like ourselves, we can see a very limited social horizon. Homophily thus plays a key role in reproducing the economic and cultural differences between demographic groups (DiMaggio and Garip 2011). Everything from cultural tastes (Mark 1998) to attitudes (McPherson 2004) to voluntary affiliations (Popielarz and McPherson 1995) become localized in social space to the extent that we surround ourselves with demographically similar others (McPherson 1983). In a real sense, you are who you know.

Given homophily's central importance, it is surprising that we have so little knowledge of whether this fundamental social fact has changed over time. Here, we ask whether the strength of homophily in close personal ties (defined as discussing important matters) has changed over the period 1985 to 2004. We use data from the 1985 and 2004 General Social Surveys (GSS) to examine five important sociodemographic characteristics—sex, race/ethnicity, religion, age, and education—to see if the degree of homophily has changed in U.S. society.

Change in Social Distance and Homophily

Researchers have studied homophily in network ties that range from the closest ties of marriage (Mare 1991; Qian and Lichter 2007), the strong confidant relationships of "discussing important matters" (Marsden 1987, 1988), and the intermediate ties of friendship and trust (Moody 2001; Verbrugge 1977), to the more circumscribed relationships of career support at work (Ibarra 1992, 1995), acquaintance (DiPrete et al. 2010), appearing with others in a public place (Mayhew et al. 1995), mere contact (Wellman 1996), "knowing about" someone (Hampton and Wellman 2001), and even the negative ties of victimization (Sampson 1984).

Most of the information we have about changes over significant periods of time in the network structure of our society come from either the exclusive, close ties of marriage (Kalmijn 1998) or the much weaker ties of co-employment (Tomaskovic-Devey 1993), co-residence (Massey and Denton 1993), or co-matriculation (Shrum, Cheek, and Hunter 1988). This is largely because these relations tend to leave official records that researchers can access. Scholars have been able to track the close but unofficial contacts among people over time only within relatively captive populations like schools, and then only for relatively short periods (Kossinets and Watts 2009; Moody 2001).

The best information comes from studies of marital homogamy (see review in Kalmijn [1998]). Here, we have seen a small decrease in the age gap between males and females in marital unions. Educational homogamy in the United States has increased, as has homogamy on other measures of social status related to workplace and social class (Schwartz 2010), although the findings on educational homogamy do vary somewhat across studies (e.g., Rosenfeld [2008] finds little change). Gender heterophily of spouses has, of course, been complete in a society that only allowed same-sex unions very recently and in limited jurisdictions. Religious homogamy has declined as society is increasingly more structured by education, work, and class than by religious institutions (Fischer and Hout 2006). Racial homogamy in marriage is very high but decreasing (Rosenfeld 2008).

Explaining Changes in Homophily

Our analyses here are the first to address the question of homophily change at a national level for informal, close ties.[1, 2] Past work has linked temporal and contextual variation in homogamy to macro level economic and demographic variables (Blau, Beeker, and Fitzpatrick 1984; Torche 2010). Drawing on this work, we discuss theoretically how homophily is affected by changes in demographic composition, institutional segregation, economic inequality, and symbolic/cultural boundaries. We then describe how our five demographic dimensions should change in social salience, given the observed changes in macro level features.

Demographic Change

Demographic change can influence the *raw*, or absolute, rate of contact between demographic groups. A long line of empirical and theoretical work (Blau 1977; Blau et al. 1984) demonstrates that increased heterogeneity leads to more out-group ties. If there are more Hispanics (or more members of any other minority group), then the opportunity for Whites (the majority) to interact with that minority group increases. Given the changes in the opportunity structure, there should be more cross-group ties.

Demographic change can also influence the salience of demographic dimensions, or the rate of in-group ties relative to chance. The theoretical expectations are more uncertain here, however. As heterogeneity increases, there should be more contact between demographic groups (Blau et al. 1984; Blau and Schwartz 1984). This could decrease the cultural, linguistic, or economic distinctiveness of minority groups. This would, in turn, eventually decrease the salience of that social dimension (Allport 1954). Putnam and Campbell (2012), for example, show that close contact with someone of another religious group makes one not only more positive toward *that* group, but more positive about other religious out-groups as well. In contrast, if increasing contact between groups is conflictual or competitive (Olzak 1992), there is little reason to expect a decrease in homophily (e.g., competition over scarce low-wage work may not increase friendships among competing groups).

For our demographic dimensions, race and religion exhibit the clearest changes in composition. The United States became much more diverse racially and ethnically between 1985 and 2004. New waves of immigration from the Americas and Asia were added to the fairly stable Black population to create the smallest European-American proportion (69 percent) ever by the end of the twentieth century (Fischer and Hout 2006). Religious diversity also increased during our period, fueled by both immigration and differential fertility rates (Fischer and Hout 2006). There was a decrease in the Protestant majority, a relatively stable Catholic population, and an increase in non–Judeo-Christian categories, including individuals affiliating with no religion.

Compositional changes are considerably smaller for sex, age, and education during our period of interest. The expansion of educational attainment was an important feature of the United States during the twentieth century (Fischer and Hout 2006). Most of this change occurred before 1985, however. Similarly, fertility dropped strikingly in the early and middle twentieth century, while life expectancies grew. The shift in age heterogeneity between 1985 and 2004 is comparatively small, and largely a result of cohort succession. If there is any trend, it would point to a small decrease in heterogeneity. Thus, based on demographic pressure alone, we would expect decreasing absolute homophily for race and religion, and little change elsewhere.

Institutional Segregation

Demographic sorting along residential, occupational, and associational lines creates strong barriers to out-group ties and will affect the rate of homophily in a population. People form social ties at work and in voluntary associations (Feld 1981; McPherson 1983). If workplaces and organizations are demographically homogenous (men do this job, women do that job), then individuals will form homophilous social ties (McPherson and Smith-Lovin 1987). Additionally, if people are recruited into jobs and organizations through social ties, and social ties are initially homophilous, then one's pool of friends will be demographically similar and the system is reproduced (McPherson 2004).

Thus, while we expect increases in population heterogeneity to be reflected in higher rates of interconnection among categories, this effect is not definitional. If there is strong physical segregation or occupational sorting, then increasing population diversity may not be reflected in absolute homophily. The potential for institutional influences on homophily relative to chance is even stronger. Weakening institutional segregation should result in decreasing homophily relative to chance, given the changes in demographic heterogeneity.

For example, compositional changes in gender are quite small compared to changes in institutional segregation. Later cohorts of women are more likely to be employed in the labor force, before and during marriage and after childbearing (Fischer and Hout 2006). While men still do less housework than women, their participation has shifted in the direction of more time with both household chores and childcare (Bianchi et al. 2000; Parker and Wang 2013). Similarly, Tomaskovic-Devey and colleagues (2006) find that occupational sex segregation decreased steadily between 1980 and 2003 (see also Marsden 2012). Women and men are thus less institutionally segregated and we expect homophily to have decreased.

Age offers the opposite story: there is little demographic change, but institutional segregation increased over time. The largest structural changes we see for age are in the timing of various life course transitions. Age at marriage and first cohabitation continued to move upward during our period (Fischer and Hout 2006), with more people living as single adults, both before unions and after divorce or death of a spouse. The decline in middle-aged people (30 to 64 years) who were married with children was particularly steep during this period.

Changes in life course patterning could affect the institutional and residential landscape for age. With more people delaying "older" responsibilities of marriage and family formation, the associational, residential, and occupational patterning of the young and middle-aged may differ more starkly over time, leading to an increase in age homophily (i.e., couples with kids have a different association profile than singles without kids). In addition, these institutional changes may create a larger block of "young" people, incorporating those in their 20s and early-30s into one large social group. We would then see a decrease in social distance at the young end of the distribution offset with an increase in social distance between the young and middle-aged (with "young" stretching into the 30s). The overall change in salience is somewhat ambiguous, but we should see a change in the patterning of social distances—with the young increasingly isolated from older Americans.

Race offers a third profile of demographic and institutional change; here, there is a large increase in heterogeneity, but few changes to the larger forces separating demographic groups. Unlike with sex, changes in racial occupational segregation flattened out and changed slowly over the period in question (Tomaskovic-Devey et al. 2006). Similarly, residential segregation changed little for Blacks after 1980 and actually rose for Asians and Hispanics (Logan, Stults, and Farley 2004). We thus expect much smaller changes in homophily (relative to chance) than with age or sex.

Growing Economic Inequality

Recent comparative work points to a kind of isomorphism (Torche 2010) between economic inequality and social relations based on income, education, and other markers of attainment (Schwartz 2010). As inequality increases, status distinctions become starker and mixing patterns reflect the changing social meaning of the demographic dimension (Schwartz and Mare 2005). Increasing economic disparities could lead to residential segregation as well as differences in status and consumption patterns—all of which make it difficult to form and maintain a confiding relationship. As economic inequality increases, we should find an increase in the rate of in-group ties for attainment-based dimensions.

Over our period, we saw increases in income and wealth inequality that were unprecedented among the world's richest democracies (Neckerman 2004). Rates of unionization declined substantially during the period, while returns to education grew significantly. Wage inequality by educational status was significantly greater at the end of the period than at the beginning (Fischer and Hout 2006), as were wealth inequality and consumption disparities. Because education is now more strongly tied to income, and income inequality is growing, the social consequences of not having high education are larger. Interaction between people with different education levels should thus be less likely and the salience of education should increase (Torche 2010).

Homophily on nominal characteristics can also be affected by changes in economic outcomes. For example, men's and women's earning potential converged over time, despite the overall increase in inequality (Leicht 2008). Indeed, a new report by the Pew Research Center shows that a record 40 percent of all households with children have a female as their primary breadwinner now (Wang, Parker, and Taylor 2013). Both single mothers and women who outearn their husbands contribute to this trend; the latter group has increased fourfold (from 4 to 15 percent) in the past 50 years. This suggests that men and women are more likely to be status equals and to treat each other as confidants. We should thus see a decrease in homophily by sex.

Race once again paints a more stagnant picture. Racial income gaps decreased over much of the twentieth century but stalled during the past 20 years (Leicht 2008). This reinforces our expectation that racial homophily will not change relative to chance. If there is any economic trend for religion, we have witnessed a decrease in economic inequality, with Catholics converging on the rest of the population along educational, wealth, and occupational lines (Keister 2003).

Symbolic Boundaries and Attitudes

Homophily may also be affected by symbolic boundaries. If one demographic group views another demographic group as "other" and sees them as incompetent or cold (Fiske

2011), then interaction may be unlikely even if the material conditions and opportunity structure make interaction possible. Similarly, changes in intergroup attitudes, or "us versus them" beliefs, will have only small effects on homophily if there is little change to the occupational, residential, or organizational sorting of demographic groups. One cannot befriend someone they do not meet, even if affective social distance is low. The stronger the connection between macro-level changes and intergroup attitudes, the more likely the structural changes will result in shifts in interaction patterns.

Gender offers a clear example. We have seen large changes to the economic and structural positions occupied by women and men, and attitudes have generally tracked these changes. For example, the GSS questions measuring gender role attitudes show a clear trend in a nontraditional direction, although there was something of a plateau in the mid-1990s (Marsden 2012). Changing attitudes coupled with changing structural patterns point consistently to a decrease in gender homophily.

Race is more ambiguous, with attitudes and structure only sometimes moving in the same direction. Compositionally, the United States has grown racially heterogeneous during our period of interest. If perceptions matter in addition to reality, roughly half of White Americans thought they were already in the minority by 2000 (Alba, Rumbaut, and Marotz 2005). While these incorrect perceptions tend to be linked to negative attitudes toward minorities, attitudes about race and ethnicity have generally shifted in the direction of greater tolerance (Bobo et al. 2012; Firebaugh and Davis 1988). The inclusive trend continued into the 2000s, with willingness to have a close family member marry a person of another race reaching broad acceptance. Some indications remain, however, that Whites are still reluctant to embrace African Americans on an emotional level (Bobo et al. 2012). The movement toward more inclusive attitudes is also undercut by the lack of economic and residential integration.

As with race, the increasing religious diversity of the United States seems to be reflected in a greater tolerance for intimate relationships with other religious groups. Levels of religious intermarriage increased during the period, as did reported acceptance of cross-faith unions (Chaves and Anderson 2012: Fischer and Hout 2006). The number of people who say they have no religion and those who never attend religious services both increased, and the general population (which remains quite religious) has demonstrated somewhat greater acceptance of these nonbelievers (Chaves and Anderson 2012). This would point to decreasing homophily. In contrast, Edgell, Gerteis, and Hartmann (2006) still find a very large divide between believers and nonbelievers, while the political and cultural divide between religious and nonreligious categories may have grown during this period of religious politicization (Hout and Fischer 2002). Hout and Fischer (2002) argue that people who used to identify as Protestant now claim no religion as a reaction against the increasing connection between conservative political ideology and religious affiliation. This would point to increasing homophily, as the cultural distance between Protestants and non-Protestants grew during our period of interest.

Summary of Expectations: Sex, Race, Religion, Age, and Education in the Late Twentieth Century

In short, demographic, institutional, economic, and cultural factors combine in particular ways to increase or decrease the potential for out-group ties. The exact set of macro-level features varies by dimension; the following paragraphs summarize the expected changes in homophily for each demographic dimension.

Sex and gender. The most distinctive feature of sex as a social distinction is that its marginal distribution does not change dramatically over time. This distributional stability has, however, been coupled with institutional desegregation, economic parity, and more liberal attitudes. Given the general trend of gender equality and desegregation, we expect an increase in cross-sex confidant ties.

Race and ethnicity. Increases in racial heterogeneity point to a decrease in absolute homophily. The trends we see for increasing tolerance and lowered affective social distance are not coupled with changes in residential, occupational, or economic outcomes. We thus expect a decrease in absolute homophily, but more muted changes when homophily is measured relative to chance.

Religion. Like race, we expect a decrease in absolute homophily given the country's increasing religious diversity. Our relative-to-chance expectations are less clear: there is economic convergence, political distancing, and only some evidence of affective acceptance. This heterogeneous set of factors suggests little overall change in salience.

Age. We also expect little change in the salience of age. There have been small shifts in age heterogeneity and the institutional effects are ambiguous. There should, however, be changes in the patterning of social distances among age categories. Given the shifts in life course transitions, the young should be increasingly isolated.

Education. Our expectations for education are straightforward: given the growing level of inequality and the increasing returns to education, we expect an increase in the social salience of education.

The Data

The GSS is a face-to-face survey of the noninstitutionalized U.S. adult population (Smith et al. 2013). The 1985 and 2004 surveys used the same questions to generate the names of confidants and identical procedures to probe for additional discussion partners. Survey responses thus provide a very close replication of the same questions and procedures at two points in time, representing the national U.S. populations in 1985 and 2004.[3] These network data have been described elsewhere in considerable detail (Marsden 1987, 1988; McPherson, Smith-Lovin, and Brashears 2006). Here, we give only a brief summary of their characteristics.

The Questions

To generate data on close, core personal ties, the GSS asked respondents about the people with whom they discussed important matters. Specifically, the 1985 and 2004 surveys asked the following question:

> 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?

After asking about the interconnections among the named confidants, the survey then asked about each confidant's demographic characteristics (e.g., race and education) and relationship to the respondent.

Several studies have explicitly compared the GSS question to other types of network measures to see what types of people respondents name. The people most likely to be mentioned in response to the GSS question are strong, close ties who are usually closely connected to others in the network (Marin 2004; Ruan 1998). These studies reinforce our sense that the GSS question elicits the core, frequently accessed interpersonal environments that people use for sociality, advice, and socioemotional support on a regular basis.

We assess shifts in homophily in the absolute sense (Are there more Black-White ties?) and in the sense of their interactional salience (Do Blacks interact with Whites more or less than would be expected by chance?). Our analysis is thus broken into two sections. In the first, we analyze homophily in its raw form,

Table 1. Summary Statistics

	1985		2004	
	Mean	SE	Mean	SE
Race				
Racial Mismatch between Respondent and Confidant***	.047	.006	.098	.010
Racial Mismatch Expected by Chance***	.276	.015	.387	.020
Religion				
Religious Mismatch between Respondent and Confidant**	.241	.010	.290	.014
Religious Mismatch Expected by Chance***	.535	.011	.658	.013
Sex				
Sex Mismatch between Respondent and Confidant*	.403	.008	.433	.011
Sex Mismatch Expected by Chance	.498	.003	.492	.005
Age				
Absolute Age Difference between Respondent and Confidant	11.792	.234	11.150	.283
Absolute Age Difference Expected by Chance**	19.839	.287	18.584	.354
Education				
Absolute Education Difference between Respondent and Confidant	2.115	.049	2.047	.058
Absolute Education Difference Expected by Chance	3.317	.084	3.120	.079

The table includes significance tests comparing the level of homophily in 1985 to the level in 2004. The level of significance is placed next to the name of the statistic. Standard errors are calculated from bootstrap samples for the observed level of homophily, and using complex survey design for the level expected by chance.

* $p < .05$; ** $p < .01$; *** $p < .001$ (two-tailed tests)

with no controls for the marginal distributions of the demographic characteristics. In the second section, we analyze homophily net of the marginals, relative to the chance probability of an in-group tie. In both cases, we discuss two sets of results: one where all confidants are included and one with only non-kin confidants (where kin is defined as any family member).

Results: Absolute Homophily

We begin our results with a simple descriptive table of absolute homophily. Table 1 presents sociodemographic distance between respondents and confidants in 1985 and 2004. We use a dummy variable to capture sociodemographic distance for sex, race/ethnicity, and religion. The dummy variable equals 1 if the respondent and confidant differ on the demographic dimension of interest (e.g., identify with different religions). The table includes the observed rate of mismatching for each categorical demographic dimension; it also includes the rate expected by chance, where we randomly pair respondents together and see if they mismatch on race, sex, or religion. The table measures the absolute difference between a respondent and confidant for our interval variables, age and education.[4]

Table 1 shows a clear decrease in raw homophily rates along racial, religious, and gender lines.[5] Individuals had proportionally more confidants of a different race, religion, and sex in 2004 compared to 1985. For example, the proportion of respondent-confidant pairs of the same religion was .76 in 1985 but only .71 in 2004. We see no significant change for age and education. Non-kin results generally mimic the all-confidant trends, although the gender shifts do not hold for non-kin confidants. In the non-kin results, the pattern of gender ties is stable over time. McPherson and colleagues (2006) note that the increase in cross-gender ties was primarily an outcome of spouses being more likely to mention each other as discussing important matters in 2004 than in 1985. The greater integration across gender lines does not extend to other kin or to non-kin.

In short, there was more contact between individuals with different races, religions, and gender in 2004 than there was in 1985. It is important, however, to keep in mind the magnitude of these shifts. For example, the absolute rate of racial matching decreased from .95 to .90. This is a substantively significant change, but one that still leaves the vast majority of ties within race. It is also important to interpret the results in relation to our macro-level variables, especially compositional change. For example, racial and religious homophily increased due to demographic changes in the population, even though we see few changes in economic inequality or institutional segregation. Gender homophily, in contrast, increased despite no changes in demographic heterogeneity. Age homophily exhibits little change over time, but this too is telling: changes in institutional segregation were not strong enough or consistent enough to affect absolute homophily rates. Educational homophily follows the demographic trends and shows no change over time, even though there was an increase in economic inequality. This suggests that compositional changes often swamp the effect of other macro-level forces on absolute homophily. It is thus useful to control for distributional changes before making any conclusions about the salience of race or religion as organizing social dimensions.

Analytic Strategy: Homophily Relative to Random Mixing

Our analysis of the social salience of characteristics, relative to the demographic opportunity for contact, builds on Marsden (1988). Marsden used log-linear and log-multiplicative methods to describe homophily in the 1985 GSS data. These models allow him to assess levels of homophily net of the impact of the marginals—the sizes of different categories of respondents within the data.

One problem with the log-linear approach is that its parameters are not easily interpretable in terms of the probability of association of people in different social positions. What we need is a model that controls for shifts in the size of social categories over time, while estimating the impact of sociodemographic distance on the probability that two members of the population will have a tie. Another problem with log-linear analyses of homophily is that researchers can typically examine only one or (at most) two dimensions at a time. We know that homophily on one dimension often translates into homophily on a correlated dimension (Kalmijn and Vermunt 2005). However, the cross-classification tables in log-linear analyses develop small or empty cells if more than one or two variables are considered at a time; these cells cause technical problems with the analysis. Therefore, we develop a model that *considers multiple dimensions simultaneously* to assess their contributions to homophily in the larger system, *net of* the other dimensions.

We use a variant of the case-control method to estimate homophily. The case-control method is widely used in medical research to study relatively rare conditions (e.g., a disease state) that are difficult to capture through random sampling (Breslow and Day 1980).[6] The method compares observed cases, which have a condition, to controls, which do not have the condition, on some exposure or preexisting condition of interest (e.g., smoking).[7] The analytic approach is a version of logistic regression applied to data sampled on the dependent variable (Hosmer and Lemeshow 1989; King and Zeng 2001). It produces consistent estimates by combining the cases (those with the condition) with the controls (those without the condition). The logistic regression analysis proceeds as though the entire dataset were sampled under the same regime (cf. Allison 1999a).

The case-control method is a natural fit for ego network data. Rather than take a random sample of dyads, or pairs of people, ego network data capture a rare condition of interest, a confiding relationship between individuals. We compare the cases, pairs with a confidant tie, to the controls, pairs with no confidant tie. The preexisting condition of interest is the demographic distance between people in the dyad.

Sampling Ego Networks: Our Cases

We first show how probability samples of individuals can yield samples of network ties. The ego network approach is a blend of the methods of survey analysis and network analysis. The researcher samples individuals and recovers information on connections among the set of contacts reported by each individual. Probabilistically representative fragments of the entire network are recovered, which are then aggregated statistically to infer characteristics of the whole.[8] Here, we treat the set of actually observed ties between respondents and confidants as a representative sample of the confidant ties that existed among people in the United States in 1985 and 2004.

Our analysis is superior to many epidemiological studies that use the case-control method, because our cases are a probability sample of all instances of confidant ties in the United States at the time of the surveys. A probability sample of cases is considered ideal, but in medical research the case sample is usually a set of available cases (e.g., from a clinic or other medical registry). We do, however, have one problem with the case sample: there is interdependency among the confidant ties generated by the same respondent.

To deal with this problem, we bracket our analysis. In our tables and figures, we report case-control analyses from all reported ties, accepting the interdependency in the sample as a reasonable trade-off against the more complete coverage of ties. In Part B of the online supplement, we present a parallel analysis that eliminates the interdependency problem but has other drawbacks. In this additional analysis, we formed the cases by randomly selecting one confidant from each respondent reporting a tie. The analysis accepts some heterogeneity in the strength of ties (because a tie may be anywhere from the first to the fifth mentioned) and less statistical power to avoid interdependency. Here we emphasize findings that are consistent across the two analyses, and we discuss reasons for divergence when we find it.

Sampling Non-ties: Our Controls

Our control sample was constructed from the set of non-ties among sampled respondents in the GSS. Respondents in the GSS are a probability sample of the non-institutionalized U.S. population, and we can safely assume that two randomly chosen GSS respondents were extremely unlikely to consider each other confidants.[9] We can thus use the non-connections found between randomly paired GSS respondents as our control sample. It is a probability sample of the potential but nonexisting ties among non-institutionalized U.S. residents.

We created the control sample by using the portion of the sample in 1985 and 2004 that reported at least one confidant and constructing non-ties between each of the $[N \times (N - 1)]/2$ pairs of these respondents by year; more technically, respondents with at least one confidant are randomly paired together $[N \times (N - 1)]/2$ times based on population weights. This strategy is the "matched sample" strategy often used in the case-control method (in the sense that it matches the population of the observed cases with the control cases). This also follows traditional log-linear models, which condition the marginals on the outcome of interest. We performed a supplementary analysis in which we constructed the controls using simulated networks (Handcock et al. 2008; Smith 2012). These results are presented in Part C of the online supplement and are very similar to the findings discussed in the manuscript.

Case-Control Analysis

After constructing the case and control samples, we then used them to model the effect of sociodemographic distance on the probability of a network tie. We combined the two samples, case and control, and performed a logistic regression (Hosmer and Lemeshow 1989) with the rare characteristic (presence of a confiding tie between two U.S. residents versus non-ties between randomly paired respondents) regressed on independent variables of interest (the sociodemographic distance between paired individuals) (cf. Allison 1999a). Sociodemographic distance is measured in the same way as in the raw homophily analysis. We regressed the vector of case/control indicators (1 = tie, 0 = no tie) on the observed sociodemographic distances. We repeated the analysis with all confidants and then only nonkin confidants.

Past work clearly indicates that the 2004 data overestimated the number of isolates, or those with no reported confidants (McPherson, Smith-Lovin, and Brashears 2009; Paik and Sanchagrin 2013). A GSS replication of the network module in 2010 suggests there was a decrease in the mean number of confidants (from 3 to 2.5), although no discernible change in the number of people claiming no confidants (Gauthier, Smith, and Smith-Lovin 2013). Given this overinflation of isolates, we ran two robustness checks to complement the reported results. First, we ran the analysis using the entire sample to construct the controls. Here, an individual with zero ties will not inform the cases but will be part of the random pairings in the control dataset; in this way, the "false" isolates will still be part of the controls. We discuss these results where the findings differ with the main tables (see Part D of the online supplement). We also replicated our analysis using the 1985 data and the 2010 network module, which includes data on gender and race for five confidants. The 2010 data include an experiment where individuals were randomly assigned to three survey designs, replicating the questionnaire context of the 1985, 1987, and 2004 GSS surveys. Our supplemental analysis uses the subset that received the 1985 survey design, making the data (GSS 1985 and GSS 2010) directly comparable across time (see Part E of the online supplement).

Significance Tests

The case-control logistic regression yields unbiased coefficients, but standard errors that are based on a potentially inflated *N*. Respondent-respondent pairs are not independent (they are cross-nested), and it is unclear if the "true" *N* used for hypothesis testing should reflect every possible respondent-respondent pairing. For a given year, there are approximately 1,500 respondents but roughly 1 million respondent-respondent pairs. The standard errors may be underestimated if they are calculated with an assumed sample size of 1 million but the true information is considerably smaller.

As a solution to these dependence problems, we calculated the standard errors using a simple bootstrap process.[10] For each iteration, we randomly drew 1,534 respondents from the 1985 sample and 1,467 respondents from the 2004 sample using population weights. The number of drawn respondents reflects the size of the original GSS samples for those years. The amount of information assumed in the analysis is thus parallel to the original data. We then constructed our case-control dataset for each year. The control portion of these datasets reflects random matching (with replacement) among respondents selected in that iteration (with ties).[11] We then ran our logistic regression. After 1,000 iterations, we calculated the standard deviation of the coefficients. We then used these standard errors in traditional statistical tests. The bootstrap standard errors capture all sources of variability in the estimates, while sidestepping any concerns over an inflated *N*.

We want to recover information about the estimated probability of contact between different positions in sociodemographic space. We thus transformed our regression coeffi-

Table 2. Case-Control Logistic Regression Using All Reported Ties, Univariate Analysis

Variable	Intercept	Dimension	Year	Dimension × Year	*N* (dyads)
1. All Ties					
Different Race	–16.877***	–2.033***	–.325***	.267	1,139,161
	(.031)	(.114)	(.042)	(.152)	
Different Religion	–16.703***	–1.287***	–.198***	–.263**	1,139,161
	(.034)	(.056)	(.049)	(.088)	
Different Sex	–17.219***	–.385***	–.501***	.143**	1,139,161
	(.030)	(.031)	(.037)	(.048)	
Age Difference	–16.415***	–.049***	–.421***	–.004	1,139,161
	(.034)	(.002)	(.049)	(.003)	
Education Difference	–16.735***	–.193***	–.405***	–.023	1,139,161
	(.038)	(.010)	(.052)	(.018)	
2. Non-kin Ties					
Different Race	–17.667***	–1.57***	–.402***	.103	442,061
	(.040)	(.127)	(.063)	(.170)	
Different Religion	–17.606***	–.819***	–.267***	–.327**	442,061
	(.047)	(.072)	(.070)	(.120)	
Different Sex	–17.518***	–1.122***	–.544***	.035	442,061
	(.039)	(.063)	(.055)	(.100)	
Age Difference	–16.056***	–.110***	–.377***	–.017*	442,061
	(.049)	(.005)	(.074)	(.008)	
Education Difference	–17.283***	–.236***	–.531***	–.006	442,061
	(.052)	(.016)	(.071)	(.025)	

Each row in Table 2 is a separate logistic regression, with one model for each sociodemographic dimension. Standard errors are in parentheses; they were calculated using bootstrap estimates. Standard errors are equal to the standard deviation of the coefficients across 1,000 iterations and are thus not dependent on the number of dyads. For each iteration, we took a random sample of respondents from each year and re-ran the case-control logistic regression.

* $p < .05$; ** $p < .01$; *** $p < .001$ (two-tailed tests)

cients into probabilities that illustrate the shifts in salience for our social dimensions. To properly estimate the probability of contact, we replaced the sample intercept (which is biased) with an inferred population intercept based on the zero-inflated Poisson models in McPherson and colleagues (2009).[12]

Results: Homophily Relative to Random Association

We begin with the analysis of homophily on one sociodemographic dimension at a time, because most previous literature examines these unidimensional patterns (e.g., Marsden 1987). We regressed the presence/absence of a tie on the measure of sociodemographic distance using three variables: distance on the sociodemographic dimension of interest, year of measurement (0 = 1985, 1 = 2004), and the interaction between year and sociodemographic distance. Each row in Table 2 is a separate logistic regression, with one model for each sociodemographic dimension. The *dimension* variable assesses the impact of sociodemographic distance on the probability of a tie in 1985. The interaction between *year* and *dimension* describes how much homophily on that characteristic changed between 1985 and 2004.[13]

First, note the strikingly consistent negative signs in the *dimension* column in the two panels of Table 2.[14] As we know from Marsden (1987), homophily was strong and ubiquitous on these five dimensions in 1985. All coefficients in panel 1 (indicating the effect of sociodemographic distance on a tie in 1985) are strongly negative. Table 2 also offers a striking shift from the results in the absolute homophily models. In Table 1, we saw a decrease in the rate of in-group ties for religion and race. In the models examining homophily relative to chance, however, we find no decrease in in-group ties for these two dimensions. There may even be a slight increase in religious homophily. Table 2 shows a small increase in religious homophily, while the more conservative analysis that uses a randomly selected confidant shows no change from 1985 to 2004. Although the absolute level of cross-religious confidants increased, it increased at a rate roughly the same (or possibly slower) than that expected by chance association in an increasingly diverse society. The salience of religion within that opportunity structure remained stable or even increased slightly.

Results for race and ethnicity also stand in stark contrast to the absolute homophily trends. The absolute rate of cross-race ties increased, but these changes parallel the population's increasing diversity. We thus see an increase in raw contact but little change in homophily when measured relative to chance.[15] Thus, we have little indication that race and ethnicity are losing their social salience.

This is not to say that a decrease in absolute racial or religious homophily is unimportant. Nothing *had to* change over time. Even with increasing opportunities for interaction, we may not have seen an increase in interracial ties if the economic, residential, or cultural differences were too vast between demographic groups. The fact that we do see increasing cross-race and cross-religion ties with increasing heterogeneity is an important social finding. More people are now close to someone of a different racial or religious background. This may lead to a more socially cohesive country over time, as cross-cutting circles connect formerly disparate parts of the population (Simmel 1955).

These results still suggest, however, that the number of cross-race and cross-religion ties would not have increased without large changes in the population's composition. This makes the changes for gender all the more unique. All three dimensions—race, religion, and gender—exhibit a decrease in absolute homophily (see Table 1). The gender composition in society was quite stable, however, while racial and religious heterogeneity increased dramatically. Gender composition changed very little, but the proportion of cross-gender ties increased. Increasing ties between other-sex confidants, in the absence of a changing sex composition, creates significant positive interaction coefficients in the analysis of all ties. Men and women are increasingly more equal in terms of economic resources and occupational roles. The growing similarities in men's and women's roles place men and women (in particular, spouses) on more equal terms and make it more likely they will see each other as confidants.

As we saw in the absolute homophily models, shifts in gender are centered in kin ties. Cross-sex ties were more common in 2004 than in 1985, but this was driven primarily by spousal ties, as opposed to friendships or work relationships. Married couples of the past were less likely to name each other as confidants. The change is thus in the nature of male-female relationships in the family, and not in cross-gender contact per se.

Multidimensional Analysis

The analysis in Table 2 parallels most analyses of homophily, taking one dimension at a time and assessing its impact on the probability of a confiding tie over the 20-year period. But both classic theorists (Blau 1977) and sophisticated recent research (Kalmijn and Vermunt 2005) call our attention to the fact that homophily on one sociodemographic dimension can create homophily on a correlated dimension. For example, if most Hispanics are Catholic and most African Americans are Protestant, then homophily on race/ethnicity will create homophily on religion (and vice

Table 3. Case-Control Logistic Regression Using All Reported Ties, Multivariate Analysis

	All Ties		Non-kin Ties	
Variables	Model 1	Model 2	Model 3	Model 4
Intercept	−14.456***	−14.519***	−13.855***	−14.047***
	(.048)	(.057)	(.068)	(.078)
Different Race	−1.819***	−1.959***	−1.468***	−1.473***
	(.077)	(.117)	(.096)	(.131)
Different Religion	−1.362***	−1.27***	−.912***	−.818***
	(.044)	(.060)	(.055)	(.074)
Different Sex	−.317***	−.373***	−1.088***	−1.101***
	(.025)	(.033)	(.052)	(.064)
Age Difference	−.049***	−.047***	−.114***	−.107***
	(.002)	(.002)	(.004)	(.005)
Education Difference	−.173***	−.157***	−.208***	−.199***
	(.009)	(.012)	(.013)	(.017)
Different Race × Year		.264		.010
		(.155)		(.171)
Different Religion × Year		−.215*		−.239
		(.092)		(.122)
Different Sex × Year		.144**		.035
		(.05)		(.102)
Age Difference × Year		−.005		−.019*
		(.003)		(.008)
Education Difference × Year		−.044*		−.026
		(.020)		(.026)
Year	−.179***	−.052	−.242***	.022
	(.047)	(.089)	(.068)	(.115)
N (respondents)	3,001	3,001	3,001	3,001
N (dyads)	1,139,161	1,139,161	442,061	442,061
−2 × Log-likelihood	73340.35	73293.46	28589.92	28570.72
AIC	73354.35	73317.46	28603.92	28594.72
BIC (*N* based on dyads)	73437.97	73460.81	28680.92	28726.71

Standard errors are in parentheses; they were calculated using bootstrap estimates. Standard errors are equal to the standard deviation of the coefficients across 1,000 iterations and are thus not dependent on the number of dyads. For each iteration, we took a random sample of respondents from each year and re-ran the case-control logistic regression.

versa). Our case-control method uses logistic regression to estimate coefficients that have a direct relationship to the probability of a tie. We can then enter multiple sociodemographic dimensions simultaneously, and see which aspects of homophily are most central and which are spurious.

Table 3 shows such an analysis. Models 1 and 3 include only the measures of homophily, indicating the independent impact of each dimension on the probability of a confiding tie *net of* the impact of the other dimensions. The even-numbered models add the interaction terms that assess whether homophily (relative to chance) changed between 1985 and 2004.

The homophily parameters in Models 1 and 3 are all negative and strongly significant. Distance on any of these dimensions strongly lowers the probability of a confiding tie, even when we control for the impact of other types of social distinctions. This is a remarkable finding in itself—these demographic dimensions are so strongly related, it is striking that they are all important *even controlling for the others*. Race and religion are stronger determinants in the all-confidant model (i.e., when kin are included) than they are in the non-kin model (i.e., when confidants are outside the family), while sex and age are stronger predictors in the non-kin model than in the all-confidant model.

Conclusions about the presence of trends in our homophily data depend strongly on the statistical criterion one chooses.[16, 17] Using BIC, a model comparison statistic that strongly favors parsimony, we conclude that homophily was stable over the 20-year period. BIC statistics for the models including statistical interactions between demographic dimension and year (Models 2 and 4) are larger than the models that exclude these interactions (Models 1 and 3). Homophily is not only a central feature of our social networks: it is a remarkably stable one. The less conservative AIC criterion, which penalizes models less for free parameters, indicates a slight preference for models that include the statistical interactions. This would provide some weak evidence for shifts in homophily.

The overall picture that emerges is one of very small changes in homophily. There are small reductions in gender homophily within the family, and relative stability for the other demographic dimensions. There may, however, be subtle changes in mixing patterns that do not alter the overall rate of homophily. We organize this more detailed discussion using Marsden's (1988) findings as a benchmark.

Race

Marsden (1988) notes that racial/ethnic divides were the most salient social distinction structuring U.S. confidant relations in 1985. Given the continuing importance of race to social institutions in the United States, we are not surprised to find very strong racial and ethnic homophily continues in the 2004 data. All groups are still much more likely to mention members of their own race/ethnicity as confidants than they are to mention a member of another group. Indeed, the basic pattern of racial homophily does not seem to have shifted over the past two decades. The interaction between year and racial difference is not significant in any model. The analysis using the entire sample to construct the controls points even more strongly to no change in racial homophily, with a smaller interaction coefficient with year.

Some scholars argue that increasing community diversity has led people to draw into their own intimate groups (e.g., Putnam 2007), which might imply increasing homophily. Others see declining overt prejudice (e.g., Bobo and Kluegel 1997; Schuman, Steeh, and Bobo 1988), which might imply decreasing homophily. Either these processes are cancelling each other out, or they are not strongly affecting the actual patterns of close confiding relationships. Racially similar ties are much more likely than cross-race ties (relative to chance), and this is as true today as it was 20 years ago. This pattern suggests that increasing heterogeneity is, at least in the short run, insufficient to change the salience of race, given the strong tendency for physical and occupational segregation in the United States.

The picture is similar when we look at ties between specific racial categories (analyses available from the authors). Here, we use five categories, White (Anglo), Black (African American), Hispanic, Asian, and other, and look at their direct effects and the interactions between this larger set of dummy variables and the *year* variable. Relative to White-White ties (which are only slightly above that expected by chance), Black-Black ties, Hispanic-Hispanic ties, Asian-Asian ties, and other-other ties are significantly above that found by random mixing, and above the level of White-White ties. Mixing rates relative to chance, however, are somewhat lower in 2004 than in 1985.

Looking at cross-racial ties, the major change occurred between Whites and Asians. For the all-confidant models, the interaction between *year* and *Asian-White* is positive and significant (although not for the non-kin model). The Asian category thus moved closer to Whites in social distance, especially in the probability of a kinship-generated tie. But overall, despite some small movements for particular categorical pairings, racial mixing has been relatively stable, with strong in-group biases in both 1985 and 2004. This is reinforced by the 2010 replication experiment, which points to no discernible change in racial homophily and yields a small, nonsignificant negative interaction between *year* and *race difference*.

Religion

After race/ethnicity, Marsden (1988) finds that religion was the next most salient social divide in the close confidant ties of U.S. respondents in 1985. As we noted, strong racial/ethnic divides can create (and be reinforced by) religious homophily. Therefore, our multidimensional analysis is especially interesting for these social distinctions. Like the racial analysis, we begin with a simple difference variable and then look at mixing between more detailed categories—Protestants, Catholics, Jews, other, and none.

Results for the religious difference variables parallel those for race/ethnicity. The effect of religious difference is weaker than race, but still quite strong (–1.362 for religious difference, compared to –1.819 for racial/ethnic difference). In Model 2, we find a significant increase in religious homophily over the time period, although the coefficient is absolutely smaller than in the univariate models (significantly different based on traditional statistical tests). Controlling for other demographic dimensions would thus appear to account for some of the increase in religious homophily over time. More specifically, racial and age homophily more strongly reinforced religious homophily in 2004 than in 1985.[18] Controlling for these demographic differences affected religious homophily more in 2004 than in 1985, and we see a weaker year interaction in the multivariate model than in the univariate one (with race playing the largest reducing role). For the random-confidant model, there was no significant change in religious homophily between 1985 and 2004 after controlling for the other dimensions (see Part B of the online supplement).

We now move to the more complex model, where each categorical pairing has one term in the model and Protestant-Protestant acts as the reference group (analyses available from authors). In general, the within-group parameters are positive and significant. As is typical for smaller minority groups, we find more in-group ties for Jews, none, other, and Catholics relative to chance than for Protestant-Protestant ties relative to chance. Over time, the number of Protestant-Protestant and Catholic-Catholic ties has increased. In contrast, the number of in-group ties for the other and none categories decreased relative to what would be expected by chance (and relative to the baseline).

Overall, the increasing homophily for some categories (Protestant and Catholic) was largely offset by the decreasing homophily for other categories (none and other). The aggregate shifts point to little change in homophily over the past 20 years.

Sex

Gender remains a strong force structuring confidant networks, but with a coefficient of –.317, it is not nearly as strong as racial or religious differences. Furthermore, we find evidence that gender is waning, as spouses increasingly consider each other to be confidants. In Model 2 in Table 3, we see a significant positive coefficient for the interaction term between *sex difference* and *year* (.144). For non-kin ties, we find little change in the strength of gender as it organizes intimate social circles.

Age

Homophily on age is generally quite strong, except for relationships with parents, children, and other generation-linking kinship ties. For non-kin ties, life course patterns and institutional settings (e.g., schools, workplaces, social and sports clubs) tend to generate very age-homophilous networks (Kalmijn and Vermunt 2005). Even small differences in age often produce major differences in interests and institutional environments. Because age is a continuous variable, the coefficient (–.049) in Model 1 appears small. A 10-year age difference, however, makes a larger difference than gender in structuring intimate social circles. A generational difference of 20 to 30 years might have as big an impact as race or religion.

The forces that generate age-similar confidants seem to be quite consistent in 1985 and 2004. The age-year interactions in the all-confidant models are nonsignificant, although there are negative, significant coefficients in the non-kin univariate and multivariate models. This pattern offers evidence of increasing age homophily among non-kin, although this does not hold for the random-confidant univariate model.

Figure 1 offers a more nuanced look at non-kin mixing patterns. Here age is measured as a categorical variable, with categories for 20s, 30s, 40s, 50s, and over 60. The diagonal in 2004 is clearly darker, meaning homophily was stronger in 2004 (except for 40s). This is especially true for the younger age categories: the 20s and 30s both experienced increases of in-group ties (relative to chance). This may be an indication of the growing importance of youth and the delaying of life course events in structuring interaction patterns.

Education

Marsden (1988) reports that educational homophily was the least salient social dimension among those measured in the 1985 confidant networks. In our analysis we treat education, and thus educational distance, as a continuous variable. It is clear from Table 3 that individuals select confidants with similar levels of education, even though education is less salient than dimensions like race or religion (see also Rosenfeld 2008). Using coefficients from Model 1, it takes a roughly eight-year educational difference—the difference between a high school graduate and a graduate degree holder—to equal the average impact of being in different religious groups.

The tendency to choose educationally close confidants increased from 1985 to 2004, although this increase does not hold in the non-kin model. It is important to note that the interaction between educational distance and year is significant in the multivariate model but not in the univariate one. Specifically, we see a larger absolute coefficient for educational distance when controlling for age distance.[19] Results are similar, but weaker (with a non-significant, smaller absolute coefficient), when the controls include the entire sample, rather than respondents with at least one tie.

Figure 2 plots the predicted probability of a tie by difference in education. As educational difference increases, the 2004 lines drop off faster than the 1985 lines. To get a sense of the difference between years, consider the probability of a tie between people who match on all other characteristics but differ on education by four years. For example, the probability of a tie forming between two 40-year-old, White, Catholic females, one with a high school degree and one with a college degree, decreased by roughly 20 percent during the 20-year period.

The increase in kin-based educational homophily is probably due to the changing educational stock in the cohort structure. Whereas ties to parents and grandparents used to connect the more highly educated young cohorts to those with less education, the more consistent educational stock of the population now mutes this integrating aspect of kinship ties. One finds higher levels of educational homophily for kin ties because of the institutional structures that foster homogamous marriage and the intergenerational transmission of educational opportunity. In addition, women's rising educational attainment (now higher than men's attainment, on average) increases the chance that spouses will have similar levels of education. Because spouses are frequently mentioned as confidants, this homogamy increases educational homophily in kin ties.

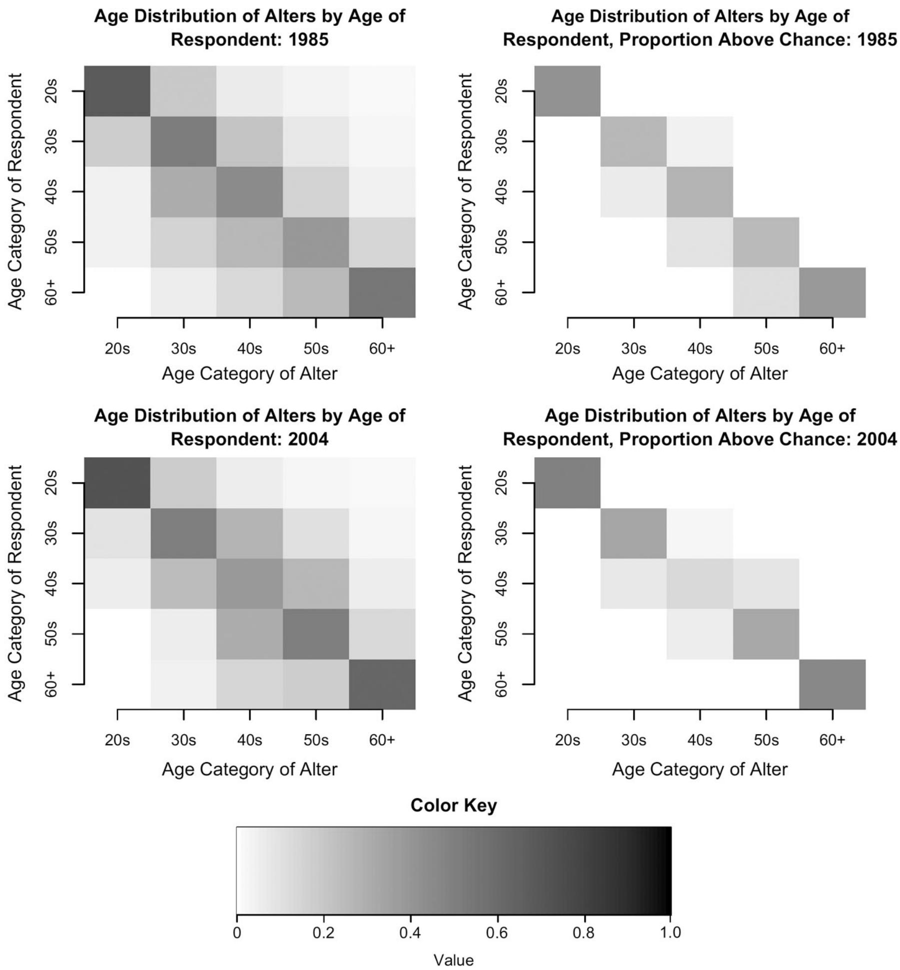

Figure 1. Strength of Homophily on Age: Non-kin Ties Only. Shaded blocks are row conditioned probabilities: the proportion of alters in each age category are calculated separately for the age categories of the respondents. Darker blocks indicate a higher proportion of alters in that age category for that category of respondents. Plots on the right hand subtract the proportion expected by chance from the observed proportion. Values less than 0, or proportions below that expected by chance, are set to 0.

The picture is similar, but more subtle, when education is measured categorically (less than high school, high school, some college, college, graduate degree). Here, homophily changes are driven in large part by changes at the bottom of the distribution—and this is true in both the all-confidants and non-kin models. Individuals with less than a high school degree were more likely to select in-group confidants (relative to chance) in 2004 than in 1985. Similarly, they were more likely to select confidants with high school degrees than to select confidants with higher education. We thus find increased levels of homophily at the lower end of the educational distribution. The increasing impact of educational attainment on income and wealth may mean that educational distinctions are reflected not just in occupational settings, but increasingly in residence and leisure pursuits (Morris, Bernhardt, and Handcock 1994). Individuals without educational capital may be increasingly ghettoized.

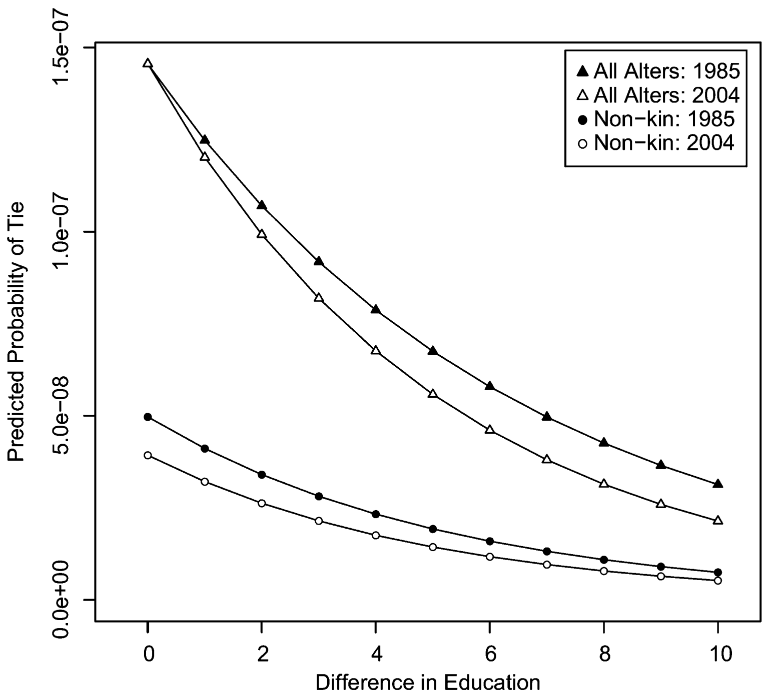

Figure 2. Probability of a Confiding Relationship by Educational Distance. Probabilities are calculated for two people of the same race, opposite sex, same religion, and 15-year age difference.

These results lend some support to our hypothesis of increasing educational homophily, although the evidence is far from uniform and many models show no change. Education may, in fact, increase in social salience as it becomes more important for economic outcomes. This effect is likely reduced, however, due to the large within-education heterogeneity on income and occupation—so that an increase in average returns to education has only a small impact on aggregate educational mixing patterns (Breen and Salazar 2011).

Discussion and Conclusions

The GSS question about discussing important matters captures the close confidants with whom we share problems, joys, and world views. The extent to which these close confidants mirror our own sociodemographic characteristics determines a great deal about how well our society is connected across gender, cohort, class, and religious lines. Having a very homogeneous and comfortable social world brings some benefits, like lowered suicide rates (Ellison, Burr, and McCall 1997) and longer-surviving marriages (Yamaguchi and Kandel 1997). On the other hand, limiting our intimate social horizons to people very much like ourselves can constrain our ability to understand others' world views (McPherson 2004).

In the 1985 GSS data, Marsden (1988) finds network patterns of homophily that follow a straightforward distance imagery: the further away people are demographically, the less likely they are to confide in one another. Deviations from this pattern were rare and generally involved cross-age cohort ties of kinship. In describing our results here, we should first be very clear: the fundamental patterns of homophily that Marsden (1988) finds in the 1985 GSS are still working powerfully in 2004. Homophily is one of the most stable, ubiquitous social facts in our discipline (Blau 1977; McPherson et al. 2001). What we discuss here are variations on that theme; subtle changes that occurred in this key indicator of social structure over the past two decades.

The clearest shifts in homophily occurred for raw mixing patterns. The population is now racially and religiously more heterogeneous, and these compositional changes directly affect the racial and religious makeup of individuals' social networks. Other-religion and other-race confidants were more likely in 2004 than in 1985.

Our case-control models account for the shifting composition of the population, and the results here reveal remarkable stability. If one had to summarize our results in one sentence, it would be "homophily has not changed much." We did, however, find small but important changes in the salience of different demographic dimensions. Race, the most important divide in our society, has not changed much at all (controlling for population composition). Racial distinctions still hold remarkable salience. Religious homophily also exhibited few changes over time. Some models point to an increase in homophily, but this is far from uniform across analyses.

Sex (or more accurately, gender) is the only social dimension on which homophily relative to chance has declined substantially. Nonkin confidants are still very likely to be same-sex, but the increasing reliance on spouses and partners as people with whom we discuss important matters has created important bridges in the gender divide. Gender roles have changed a great deal in our society, with women entering and staying in the labor force, women earning higher salaries, and men participating to a greater extent in childrearing. Spouses are now more similar to one another, and they were more likely to mention each other as confidants in 2004 than in 1985.

Education offers ambiguous results over time. Education may, perhaps, be growing in salience as it becomes increasingly important in determining employment, income, residence, and leisure activities (Fischer and Hout 2006). Individuals on the lower end of the educational spectrum appear to be increasingly socially isolated, but this conclusion is undercut because the increase in homophily does not hold across models (i.e., the univariate models, the non-kin model, and the raw homophily models). If anything, the change is concentrated within families, where spouses are now more likely to have equivalent education. People marry later and women are more structurally similar to their husbands in education, occupation, and income (Taylor et al. 2010). Due to cohort succession, parents and children are also now more likely to be similar in education than they were 20 years ago.

Age changes were concentrated primarily among the young, who appear to be somewhat more cloistered in their generational institutions in recent years. This pattern points to the growing importance of delayed life course transitions. Overall, however, we find no changes to the strength of age homophily.

Theoretically, these results point to stark differences between raw and relative-to-chance measures of homophily. Changes in raw homophily followed compositional changes quite closely. Out-group ties increased when demographic heterogeneity increased. Of course, we could have seen a divided world of increasing heterogeneity but few cross-race and cross-religion ties. The fact that this did not happen is substantively and theoretically important. Theoretically, the systematic changes for race and religion are a testament to the enduring strength of the arguments developed by Blau (1977). Substantively, interacting with demographically different individuals can broaden one's intellectual horizon. This could lower prejudices and the perceived "otherness" of different demographic groups, leading to a possible decrease in racial or religious salience in the future. The changes after controlling for population composition were more subtle. Changes in institutional and affective salience do not seem to follow compositional shifts immediately. Gender is the only social divide in our society eroding to any noticeable extent, and this is occurring primarily within the nuclear family. Our new ability to estimate multidimensional models that control for other sources of homophily allows us to say with unusual certainty that these bases of social distinction are stable and not spurious. They are organized by our institutional structures and do not change quickly.

Our results offer guidance for future work on the changing salience of demographic dimensions. We highlight three general themes. First, changes in homophily are bound up in the changing shape of multiple distributions: the distribution of the demographic dimension itself, as well as the distribution of other entities (e.g., income and geographic residence) across and within demographic categories. Absolute homophily responds directly to the opportunity structure created by the composition of the population, but the pace of residential and occupational segregation may limit changes in demographic salience.

Second, changes in salience may significantly lag other social change. The large changes in gender occurred in the decades before the period under study, but we see changes in confidant behavior in the 1985 to 2004 period. Composition-driven mixing by race and religion might have similar effects over time. Inequality by education is increasing now, but its full effects may not be felt until later. Similarly, the salience of social dimensions may be concentrated in certain parts of social space that make it difficult to see in more global analyses. The (non-kin) young appear to be getting more insular; the poorly educated appear to be getting more isolated. These effects may be real but hidden by the larger population groups in the middle of the distribution.

Third, we are struck by the remarkable stability of the homophily results. The fact that homophily is so stable in our data indicates it is an unusually robust social fact. It is also one of our most important. Homophily simultaneously reflects and reproduces the social order. It has been and will continue to be a sturdy foundation from which to build sociological theories.

Acknowledgments – The authors would like to thank Mark Chaves, Robin Gauthier, and Francois Nielson for comments on earlier drafts of this article. The authors acknowledge the support of two grants from the National Science Foundation, SES 0347699 and BCS 0527671, to the second and third authors.

Notes

1. Researchers seldom assess network connections in ways that can be generalized to a well-defined population at even one point in time (for exceptions, see Burt 1984; Fischer 1982; Marsden 1987, 1988; Verbrugge 1977).
2. More precisely, our analysis focuses on what McPherson and colleagues (2001) call inbreeding homophily. Inbreeding homophily is the rate of in-group ties above that expected from the demographic composition of the *entire* population. This aggregate measure of homophily implicitly captures both induced homophily (individuals are sorted into foci based on demographic characteristics) and choice homophily (individuals create social ties within locations with individuals who are demographically more similar).
3. While the question was identical in both survey years, we can never be sure that the context and meaning of an item is exactly the same over a 20-year period. Some relevant differences include (1) the survey was administered using a CAPI format in 2004, but using a paper questionnaire in 1985; (2) the questions preceding the item involved religion in 1985, whereas in 2004 they asked about voluntary associations; and (3) a larger proportion of the 2004 survey respondents were interviewed by phone rather than face-to-face.
4. Education is measured differently for respondents and confidants. For respondents, education is measured as both years of education and degree earned (we use years of education in the models). For confidants, education is measured as broad categories corresponding to years and degree (specifically, 1 to 6 years, 7 to 9 years, 10 to 12, high school, some college, associate degree, college, and graduate degree). We code confidant years of education as the mean number of years associated with their educational category.
5. Table S1 in Part A of the online supplement (http://asr.sagepub.com/supplemental) offers a more formal model of absolute homophily change. The models predict year (2004 versus 1985) as a function of demographic distance in the respondent-confidant pairs. The model is conditioned on all demographic dimensions simultaneously and yields the same general conclusions as Table 1.

6. This method is essentially an answer to two problems: (1) the inability to do a true experiment, and (2) the enormous sample size necessary to study a rare condition through a normal probability sample.
7. The controls can be explicitly matched to the cases, or they can be drawn in a probability sample from the same population that creates the cases. Here, we match the controls to the cases in the sense that we use GSS respondents who reported at least one confidant when constructing the controls.
8. Of course, there are limitations to what can be studied with ego networks. Because the ego network is only a tiny part of the global network, the kinds of large-scale structural properties that can be studied are limited (but see Smith 2012). Probably the most important measurement constraint is that relying on individuals to recover information about their network contacts introduces many issues of memory and context.
9. The calculated probability of this event (i.e., two GSS individuals discussing important matters with each other) is of the order of $p < .001$.
10. It is difficult to run a multilevel model to account for dependencies in these data. The controls are cross-nested across all respondents, while the respondent-confidant pairs are nested within respondents. This makes it difficult to specify the dependence structure in a traditional generalized mixed model; we find a bootstrapping approach to be more appropriate and straightforward.
11. The number of control dyads is held fixed across years and is equal to (N2004 x N2004 – 1)/2 where N2004 is the number of people in 2004 with ties. This makes it possible to directly estimate the *year* coefficient in the model. 12. Specifically, the true intercept represents the probability of two people in the population, N~200 million, discussing important matters if they share the exact same position in Blau space. We first calculated the probability of two randomly chosen individuals in the population having a tie. The average number of ties per person was estimated from the zero-inflated Poisson model discussed in McPherson and colleagues (2009). The total number of people was taken from census data. We used these pieces of information to calculate the unconditioned population level density. Because our model has covariates, we then altered this baseline intercept to take into account these extra variables.
13. Hypothesis tests are difficult to interpret on interaction terms in logistic regression due to the possibility of unobserved heterogeneity across groups (year in our case) (Allison 1999b). We use the approach described by Williams (2009) to test the robustness of our results. We ran the models under different model assumptions but found little differences from the more standard logistic regression. We report the standard model results here.
14. Researchers interested in replicating our results can use the public access data files for the 1985, 2004, and 2010 GSS. We will also make available, upon request, the R code to run the case-control logistic regression.
15. The decrease in racial homophily is stronger when we take into account differential degree (see Part C of the online supplement). Here, individuals are selected proportional to degree when constructing the controls. Despite these results, the overall rate of homophily for racial ties does not appear to have changed much over time. The only models that ever show a significant decrease are those that control for differential degree—and even these results offer inconsistent findings. Additionally, we would caution any researcher in interpreting these results given the known problems in the 2004 data for degree.
16. AIC and BIC were calculated from the case-control models using the original GSS sample. The fit statistics were then averaged over 100 runs using the original GSS sample, as there was some (minor) variation from run to run in construction of the controls.
17. We ran a series of regressions looking for the best fit (models available upon request); for the all-confidant multivariate results, the best model was the one including all coefficients (using AIC but not BIC).
18. This becomes clear when adding a single demographic dimension at a time to the religion univariate model in Table 2; here we can see how the *religious* x *year* coefficient is affected by each dimension on its own.
19. There is a considerable amount of educational heterogeneity across cohorts. By controlling for age distance, the comparison of educational distance across years is made comparable—individuals at the same "cohort distance," and thus the same kind of educational heterogeneity, are compared.

References

Alba, Richard, Rubén G. Rumbaut, and Karen Marotz. 2005. "A Distorted Nation: Perceptions of Racial/ Ethnic Group Sizes and Attitudes toward Immigrants and Other Minorities." *Social Forces* 84:901–919.

Allison, Paul D. 1999a. *Logistic Regression Using the SAS System: Theory and Application*. Cary, NC: SAS Institute Inc. and John Wiley and Sons.

Allison, Paul D. 1999b. "Comparing Logit and Probit Coefficients across Groups." *Sociological Methods and Research* 28:186–208.

Allport, George W. 1954. *The Nature of Prejudice*. Cambridge, MA: Perseus Books.

Annual Review of Sociology. 2013. "Most Cited Reviews." Retrieved November 1, 2013 (http://www. annualreviews.org/action/showMostCitedArticles?t opArticles Type=sinceInception&journalCode=soc).

Bianchi, Suzanne M., Melissa A. Milkie, Liana C. Sayer, and John P. Robinson. 2000. "Is Anyone Doing the Housework? Trends in the Gender Division of Household Labor." *Social Forces* 79:191–228.

Blau, Peter M. 1977. *Inequality and Heterogeneity: A Primitive Theory of Social Structure*. New York: Free Press.

Blau, Peter M., Carolyn Beeker, and Kevin M. Fitzpatrick. 1984. "Intersecting Social Affiliations and Intermarriage." *Social Forces* 62:585–606.

Blau, Peter M. and Joseph E. Schwartz. 1984. *Cross-Cutting Social Circles: Testing a Macrosociological Theory of Intergroup Relations*. New York: Academic Press.

Bobo, Lawrence D., Camille Z. Charles, Maria Krysan, and Alicia D. Simmons. 2012. "The Real Record on Racial Attitudes." Pp. 38–83 in *Social Trends in the United States: Evidence from the General Social Survey since 1972,* edited by P. V. Marsden. Princeton, NJ: Princeton University Press.

Bobo, Lawrence D. and James R. Kluegel. 1997. "Status, Ideology and Dimensions of Whites' Racial Beliefs and Attitudes: Progress and Stagnation." Pp. 93–120 in *Racial Attitudes in the 1990s: Continuity and Change,* edited by S. A. Tuch and J. K. Martin. Greenwood, CT: Praeger.

Borgardus, Emory S. 1925. "Measuring Social Distance." *Journal of Applied Sociology* 9:299–308.

Breen, Richard and Leire Salazar. 2011. "Educational Assortative Mating and Earnings Inequality in the United States." *American Journal of Sociology* 117:808–843.

Breslow, N. E. and N. E. Day. 1980. *Statistical Methods in Cancer Research,* Vol. 1, *The Analysis of Case- Control Studies*. Lyon, France: IARC Scientific Publications.

Burt, Ronald S. 1984. "Network Items and the General Social Survey." *Social Networks* 6:293–339.

Chaves, Mark and Shawna Anderson. 2012. "Continuity and Change in American Religion, 1972–2008." Pp. 212–39 in *Social Trends in the American Life: Findings from the General Social Survey since 1972,* edited by P. V. Marsden. Princeton, NJ: Princeton University Press.

DiMaggio, Paul and Filiz Garip. 2011. "How Network Externalities Can Exacerbate Intergroup Inequality." *American Journal of Sociology* 116:1887–1933.

DiPrete, Thomas A., Andrew Gelman, Tyler McCormick, Julien Teitler, and Tian Zheng. 2010. "Segregation in Social Networks based on Acquaintanceship and Trust." *American Journal of Sociology* 116:1234–83.

Edgell, Penny, Joseph Gerteis, and Douglas Hartmann. 2006. "Atheists as 'Other': Moral Boundaries and Cultural Membership in American Society." *American Sociological Review* 71:211–34.

Ellison, Christopher G., Jeffrey A. Burr, and Patricia L. McCall. 1997. "Religious Homogeneity and Metropolitan Suicide Rates." *Social Forces* 76:273–99.

Feld, Scott L. 1981. "The Focused Organization of Social Ties." *American Journal of Sociology* 86:1015–35.

Firebaugh, Glenn and Kenneth E. Davis. 1988. "Trends in Antiblack Prejudice, 1972–1984: Region and Cohort Effects Trends in Antiblack Prejudice, 1972–1984." *American Journal of Sociology* 94:251–72.

Fischer, Claude S. 1982. *To Dwell among Friends: Personal Networks in Town and City*. Chicago: University of Chicago Press.

Fischer, Claude S. and Michael Hout. 2006. *Century of Difference: How America Changed in the Last One Hundred Years*. New York: Russell Sage. Fiske, Susan T. 2011. *Envy Up, Scorn Down: How Status Divides Us*. New York: Russell Sage Foundation.

Gauthier, G. Robin, Jeffrey A. Smith, and Lynn Smith-Lovin. 2013. "Trends in the Network Size of Close Confidants: Necessary Caution and Available Information." Unpublished manuscript. Duke University, Durham, NC.

Hampton, Keith and Barry Wellman. 2001. "Long Distance Community in the Network Society: Contact and Support beyond Netville." *American Behavioral Scientist* 45:476–95.

Handcock, Mark S., Steven M. Goodreau, David R. Hunter, Carter T. Butts, and Martina Morris. 2008. "ERGM: A Package to Fit, Simulate and Diagnose Exponential-Family Models for Networks." *Journal of Statistical Software* 24:1–29.

Hosmer, David W. and Stanley Lemeshow. 1989. *Applied Logistic Regression*. New York: John Wiley & Sons.

House, James S., Debra Umberson, and Karl R. Landis. 1988. "Structures and Processes of Social Support." *Annual Review of Sociology* 14:293–318.

Hout, Michael and Claude S. Fischer. 2002. "Explaining the Rise of Americans with No Religious Affiliation: Generations and Politics." *American Sociological Review* 67:165–90.

Hughes, Michael and Steven A. Tuch. 2003. "Gender Differences in Whites' Racial Attitudes: Are Women's Attitudes Really More Favorable?" *Social Psychology Quarterly* 66:384–401.

Ibarra, Herminia. 1992. "Homophily and Differential Returns: Sex Differences in Network Structure Access in an Advertising Firm." *Administrative Science Quarterly* 37:422–47.

Ibarra, Herminia. 1995. "Race, Opportunity and Diversity of Social Circles in Managerial Managers' Networks." *Academy of Management Journal* 38:673–703.

Kalmijn, Matthijs. 1998. "Intermarriage and Homogamy: Causes, Effects and Trends." *Annual Review of Sociology* 24:395–420.

Kalmijn, Matthijs and Jergen K. Vermunt. 2005. "Multidimensional Homogeneity of Social Networks: A Multilevel Analysis of the Role of Age and Marital Status." *Social Networks* 29:25–43.

Keister, Lisa A. 2003. "Religion and Wealth: The Role of Religious Affiliation and Participation in Early Adult Asset Accumulation." *Social Forces* 82:175–207.

King, Gary and Langche Zeng. 2001. "Logistic Regression in Rare Events Data." *Political Analysis* 9:137– 63.

Kossinets, Gueorgi and Duncan J. Watts. 2009. "Origins of Homophily in an Evolving Social Network." *American Journal of Sociology* 115:405–450.

Laumann, Edward O. 1966. *Prestige and Association in an Urban Community*. Indianapolis: Bobbs-Merrill.

Lazarsfeld, Paul and Robert K. Merton. 1954. "Friendship as a Social Process: A Substantive and Methodological Analysis." Pp. 18–66 in *Freedom and Control in Modern Society,* edited by T. A. Morroe Berger and C. H. Page. New York: Van Nostrand.

Leicht, Kevin T. 2008. "Broken Down by Race and Gender? Sociological Explanations of New Sources of Earnings Inequality." *Annual Review of Sociology* 34:237–55.

Logan, John R., Brian J. Stults, and Reynolds Farley. 2004. "Segregation of Minorities in the Metropolis: Two Decades of Change." *Demography* 41:1–22.

Mare, Robert D. 1991. "Five Decades of Educational Assortative Mating." *American Sociological Review* 56:15–32.

Marin, Alexandra. 2004. "Are Respondents More Likely to List Alters with Certain Characteristics? Implications for Name Generator Data." *Social Networks* 26:289–307.

Mark, Noah. 1998. "Beyond Individual Differences: Social Differentiation from First Principles." *American Sociological Review* 63:309–330.

Marsden, Peter V. 1987. "Core Discussion Networks of Americans." *American Sociological Review* 52: 122–31.

Marsden, Peter V. 1988. "Homogeneity in Confiding Relations." *Social Networks* 10:57–76.

Marsden, Peter V., ed. 2012. *Social Trends in American Life: Findings from the General Social Survey since 1972*. Princeton, NJ: Princeton University Press.

Massey, Douglas S. and Nancy Denton. 1993. *American Apartheid: Segregation and the Making of the Underclass*. Cambridge, MA: Harvard University Press.

Mayhew, Bruce H., J. Miller McPherson, Thomas Rotolo, and Lynn Smith-Lovin. 1995. "Sex and Race Homogeneity in Naturally Occurring Groups." *Social Forces* 74:15–52.

McPherson, Miller. 1983. "An Ecology of Affiliation." *American Sociological Review* 48:519–32.

McPherson, Miller. 2004. "A Blau Space Primer: Prolegomenon to an Ecology of Affiliation." *Industrial and Corporate Change* 13:263–80.

McPherson, Miller, Pamela A. Popielarz, and Sonja Drobnic. 1992. "Social Networks and Organizational Dynamics." *American Sociological Review* 57:153–70.

McPherson, Miller and Lynn Smith-Lovin. 1987. "Homophily in Voluntary Organizations: Status Distance and the Composition of Face-to-Face Groups." *American Sociological Review* 52:370–79.

McPherson, Miller, Lynn Smith-Lovin, and Matthew E. Brashears. 2006. "Social Isolation in America: Changes in Core Discussion Networks over Two Decades." *American Sociological Review* 71:353–75.

McPherson, Miller, Lynn Smith-Lovin, and Matthew E. Brashears. 2009. "Models and Marginals: Using Survey Methods to Study Social Networks." *American Sociological Review* 74:670–81.

McPherson, Miller, Lynn Smith-Lovin, and James M. Cook. 2001. "Birds of a Feather: Homophily in Social Networks." *Annual Review of Sociology* 27:415–44.

Moody, James. 2001. "Race, School Integration, and Friendship Segregation in America." *American Journal of Sociology* 107:679–716.

Morris, Martina, Annette D. Bernhardt, and Mark S. Handcock. 1994. "Economic Inequality: New Methods for New Trends." *American Sociological Review* 59:205–219.

Neckerman, Kathryn M. 2004. *Social Inequality*. New York: Russell Sage.

Olzak, Susan. 1992. *The Dynamics of Ethnic Competition and Conflict*. Stanford, CA: Stanford University Press.

Paik, Anthony and Kenneth Sanchagrin. 2013. "Social Isolation in America: An Artifact." *American Sociological Review* 78:339–60.

Parker, Kim and Wendy Wang. 2013. "Modern Parenthood: Roles of Moms and Dads Converge as They Balance Work and Family." Pew Research Center, Social & Demographic Trends.

Popielarz, Pamela and Miller McPherson. 1995. "On the Edge or in Between: Niche Position, Niche Overlap, and the Duration of Voluntary Association Memberships." *American Journal of Sociology* 101:698–720.

Portes, Alejandro. 1998. "Social Capital: Its Origins and Applications in Modern Sociology." *Annual Review of Sociology* 24:1–24.

Putnam, Robert D. 2007. "*E Pluribus Unum*: Diversity and Community in the Twenty-first Century." *Scandinavian Political Studies* 30:137–74.

Putnam, Robert D. and David E. Campbell. 2012. *American Grace: How Religion Divides and Unites Us*. New York: Simon and Schuster.

Qian, Zhenchao and Daniel T. Lichter. 2007. "Social Boundaries and Marital Assimilation: Interpreting Trends in Racial and Ethnic Intermarriage." *American Sociological Review* 72:68–94.

Rosenfeld, Michael J. 2008. "Racial, Educational and Religious Endogamy in the United States: A Comparative Historical Perspective." *Social Forces* 87:1–31.

Ruan, Danching. 1998. "The Content of the General Social Survey Discussion Networks: An Exploration of the General Social Survey Discussion Name Generator in a Chinese Context." *Social Networks* 20:247–64.

Sampson, Robert J. 1984. "Group Size, Heterogeneity, and Intergroup Conflict: A Test of Blau's Inequality and Heterogeneity." *Social Forces* 62:618–39.

Schuman, Howard, Charlotte Steeh, and Lawrence Bobo. 1988. *Racial Attitudes in America: Trends and Interpretations*. Boston: Harvard University Press.

Schwartz, Christine R. 2010. "Earnings Inequality and the Changing Association between Spouses' Earnings." *American Journal of Sociology* 115:1524–57.

Schwartz, Christine R. and Robert D. Mare. 2005. "Trends in Educational Assortative Marriage from 1940 to 2003." *Demography* 42:621–46.

Shrum, Wesley, Neil H. Cheek Jr., and Sandra MacD. Hunter. 1988. "Friendship in School: Gender and Racial Homophily." *Sociology of Education* 61: 227–39.

Simmel, Georg. 1955. *Conflict and the Web of Group Affiliations*. New York: Free Press. Smith, Jeffrey A. 2012. "Macrostructure from Survey Data: Generating Whole Systems from Ego Networks." *Sociological Methodology* 42:155–205.

Smith, Tom W., Peter Marsden, Michael Hout, and Jibum Kim. 2013. *General Social Surveys, 1972–2012* [machine-readable data file]. Principal Investigator, Tom W. Smith; Co-Principal Investigator, Peter V. Marsden; Co-Principal Investigator, Michael Hout. Sponsored by National Science Foundation, NORC ed. Chicago: National Opinion Research Center [producer]; Storrs, CT: The Roper Center for Public Opinion Research, University of Connecticut [distributor].

Smith-Lovin, Lynn and Miller McPherson. 1993. "You Are Who You Know: A Network Perspective on Gender." Pp. 223–41 in *Theory on Gender/Feminism on Theory,* edited by P. England. New York: Aldine.

Taylor, Paul, Richard Fry, D'Vera Cohn, Wendy Wang, Gabriel Velasco, and Daniel Dockterman. 2010. "Women, Men and the New Economics of Marriage." Pew Research Center, Social & Demographic Trends.

Tomaskovic-Devey, Donald. 1993. "The Gender and Race Composition of Jobs and the Male/Female, White/Black Pay Gaps." *Social Forces* 72:45–76.

Tomaskovic-Devey, Donald, Kevin Stainback, Tiffany Taylor, Catherine Zimmer, Corre Robinson, and Tricia McTague. 2006. "Documenting Desegregation: Segregation in American Workplaces by Race, Ethnicity, and Sex, 1966–2003." *American Sociological Review* 71:565–88.

Torche, Florencia. 2010. "Educational Assortative Mating and Economic Inequality: A Comparative Analysis of Three Latin American Countries." *Demography* 47:481–502.

Verbrugge, Lois M. 1977. "The Structure of Adult Friendship Choices." *Social Forces* 56:576–97.

Wang, Wendy, Kim Parker, and Paul Taylor. 2013. "Breadwinner Moms: Mothers Are the Sole or Primary Provider in Four-in-Ten Households with Children; Public Conflicted about the Growing Trend." Pew Research Center, Social & Demographic Trends.

Wellman, Barry. 1996. "Are Personal Communities Local? A Dumptarian Reconsideration." *Social Networks* 18:347–54.

Wellman, Barry and Scot Wortley. 1990. "Different Strokes from Different Folks: Community Ties and Social Support." *American Journal of Sociology* 96:558–88.

Williams, Richard. 2009. "Using Heterogeneous Choice Models to Compare Logit and Probit Coefficients across Groups." *Sociological Methods and Research* 37:531–59.

Yamaguchi, Kazuo and Denise B. Kandel. 1997. "The Influence of Spouses' Behavior and Marital Dissolution on Marijuana Use: Causation or Selection." *Journal of Marriage and the Family* 59:22–36.

Ω Ω Ω Ω

Jeffrey A. Smith is an Assistant Professor in the Department of Sociology at the University of Nebraska-Lincoln. His research interests broadly include networks, quantitative methodology, and stratification. He has done methodological work on network sampling and missing data, and more substantive work on social distance, status, and the relational meaning of demographic categories.

Miller McPherson is Professor Emeritus of Sociology at Duke University and the University of Arizona. He developed an ecological evolutionary model of affiliation. He has applied this model to a number of empirical cases, including occupations, organizations, and attitudes. His paper in the 2001 *Annual Review of Sociology*, "Birds of a Feather," is the second most cited chapter in the history of that series. Current projects include a test of his ecological theory with nationally representative data from the Niches and Networks project, funded by the Human and Social Dynamics Initiative at the National Science Foundation.

Lynn Smith-Lovin is Robert L. Wilson Professor of Arts and Sciences in the Department of Sociology (with secondary appointments in Psychology and Neuroscience and in Women's Studies) at Duke University. She received the 2006 Cooley-Mead Award for lifetime achievement in social psychology from the American Sociological Association Section on Social Psychology and the 2005 Lifetime Achievement Award from the ASA Section on Sociology of Emotions. Her current projects involve (1) an experimental study of justice, identity, and emotion (funded by the National Science Foundation); (2) research with Miller McPherson on an ecological theory of identity (also funded by NSF); and (3) a study of event processing in Arabic (funded by the Office of Naval Research).

Part A. Alternative Model of Absolute Homophily Change

This supplement presents a formal test of absolute homophily change. The raw homophily results are presented as a simple logistic regression. Units in the analysis are the respondent-confidant pairs from 1985 and 2004. The dependent variable is year, equal to 1 if the dyad is from 2004, and 0 if the dyad is from 1985. The independent variables are the sociodemographic distances between the respondent and the named confidant. The models thus predict the probability of a respondent-confidant pair appearing in 1985 versus 2004, as a function of sociodemographic distance. A positive coefficient suggests demographically distant confidants are more likely to exist in 2004 than in 1985. The model does not take into account the sociodemographic distance expected by chance, and simply captures the demographic similarity between confidants over time. The model differs from Table 1 in the main text because it conditions the change in one dimension on changes in another. It also provides results using all confidants and only non-kin confidants. The respondent-confidant pairs are nested within respondents, and we take these dependencies into account when calculating the standard errors. Specifically, we adjust for the complex survey design of the data when running the glm.

Table S1. Logistic Regression Predicting Year as a Function of Demographic Distance

	All Ties	Non-kin Ties
Intercept	–.545***	–.603***
	(.082)	(.110)
Different Race	.730***	.636**
	(.171)	(.195)
Different Religion	.202*	.188
	(.087)	(.119)
Different Sex	.128*	–.031
	(.056)	(.113)
Age Difference	–.002	–.009
	(.003)	(.007)
Education Difference	–.015	.003
	(.017)	(.027)
N	6,515	2,806

Note: Standard errors are in parentheses. Units in the analysis are respondent-confidant pairs nested within respondents. The estimation routine accounts for the dependence in the data when producing the standard errors.
*$p < .05$; **$p < .01$; ***$p < .001$ (two-tailed tests).

Part B. Case-Control Results Using Randomly Selected Confidant

Table S2. Case-Control Logistic Regression for Randomly Selected Confidant, Univariate Analysis

Variable	Intercept	Dimension	Year	Dimension x Year	*N* (dyads)
All Ties					
Different Race	–16.894***	–1.966***	–.132**	.289	1,134,804
	(.029)	(.143)	(.043)	(.192)	
Different Religion	–16.652***	–1.389***	.000	–.245	1,134,804
	(.035)	(.092)	(.049)	(.129)	
Different Sex	–17.276***	–.265***	–.349***	.233*	1,134,804
	(.038)	(.072)	(.058)	(.112)	
Age Difference	–16.330***	–.053***	–.219**	–.005	1,134,804
	(.050)	(.004)	(.076)	(.006)	
Education Difference	–16.749***	–.188***	–.178**	–.037	1,134,804
	(.046)	(.017)	(.062)	(.026)	
Non-kin Ties					
Different Race	–17.630***	–1.711***	–.264***	.176	440,756
	(.057)	(.152)	(.062)	(.216)	
Different Religion	–17.523***	–.971***	–.119	–.328*	440,756
	(.064)	(.097)	(.078)	(.149)	
Different Sex	–17.647***	–.861***	–.446***	.167	440,756
	(.060)	(.096)	(.065)	(.152)	
Age Difference	–16.380***	–.091***	–.210*	–.020	440,756
	(.072)	(.006)	(.100)	(.011)	
Education Difference	–17.340***	–.219***	–.371***	–.017	440,756
	(.071)	(.022)	(.084)	(.036)	

Note: Standard errors are in parentheses; they were calculated using bootstrap estimates. Standard errors are equal to the standard deviation of the coefficients across 1,000 iterations and are thus not dependent on the number of dyads. For each iteration, we took a random sample of respondents from each year and reran the case-control logistic regression.

**p* < .05; ***p* < .01; ****p* < .001 (two-tailed tests).

Table S3. Case-Control Logistic Regression for Randomly Selected Confidant, Multivariate Analysis

	All Ties		Non-kin Ties	
Variables	Model 1	Model 2	Model 3	Model 4
Intercept	–14.395***	–14.470***	–14.226***	–14.423***
	(.065)	(.080)	(.086)	(.103)
Different Race	–1.724***	–1.893***	–1.558***	–1.608***
	(.096)	(.146)	(.111)	(.155)
Different Religion	–1.468***	–1.374***	–1.078***	–.969***
	(.065)	(.092)	(.073)	(.099)
Different Sex	–.148**	–.252***	–.776***	–.844***
	(.057)	(.073)	(.077)	(.097)
Age Difference	–.054***	–.052***	–.097***	–.088***
	(.003)	(.004)	(.005)	(.006)
Education Difference	–.175***	–.151***	–.197***	–.183***
	(.014)	(.019)	(.019)	(.023)
Different Race x Year		.292		.089
		(.196)		(.219)
Different Religion x Year		–.198		–.254
		(.131)		(.153)
Different Sex x Year		.236*		.167
		(.115)		(.155)
Age Difference x Year		–.005		–.021*
		(.006)		(.011)
Education Difference x Year		–.059*		–.038
		(.027)		(.037)
Year	.026	.127	–.081	.184
	(.046)	(.119)	(.064)	(.151)
N (respondents)	3001	3001	3001	3001
N (dyads)	1,134,804	1,134,804	440,756	440,756
–2 x Log-likelihood	29260.14	29229.34	16482.39	16460.78
AIC	29274.14	29253.34	16496.39	16484.78
BIC (*N* based on dyads)	29357.73	29396.64	16573.36	16616.73

Note: Standard errors are in parentheses; they were calculated using bootstrap estimates. Standard errors are equal to the standard deviation of the coefficients across 1,000 iterations and are thus not dependent on the number of dyads. For each iteration, we took a random sample of respondents from each year and reran the case-control logistic regression.

*$p < .05$; **$p < .01$; ***$p < .001$ (two-tailed tests).

Part C. Results of Network Simulation Approach to Testing Homophily Change

This supplement presents the results of an alternative test to our case-control method. Here, we replicated the analysis using simulated networks as a means of constructing the “by chance” comparisons. In the main analysis, we generated our chance expectations by randomly pairing respondents in the GSS together for each year. We assumed the probability of randomly pairing two people together follows a binomial distribution with probability based on the population weights. Construction of the controls, and thus chance expectations, is only constrained on the distribution of demographic characteristics in the population. It is implicitly not constrained on (1) the volume of ties; (2) the degree distribution (i.e., ties per person); and (3) differential degree (i.e., some groups have more ties than others).

We reconsider those assumptions in this supplementary analysis. We made particular choices in measuring chance expectations; we could have made alternative choices. It is important to consider how our results would have differed under different assumptions. Such choices are easier to represent through network simulation, where one generates random networks and uses that to calculate chance expectations for homophily. The construction of the controls in the article is a particular version of this. Specifically, you can think of the random pairing process as creating a baseline network with N x (N – 1)/2 ties. The network is conditioned on the demographic composition of the population, and everyone has the same number of ties. Here, we extend the analysis to constrain the “simulated” network on edges (or volume), degree distribution, and differential degree.

Analytic Strategy

We began by taking a bootstrap sample of respondents in the GSS for 1985 and 2004. We drew the same number of respondents as in the original sample. We then generated networks for 1985 and 2004 using ERGM (exponential random graph models); specifically using the statnet package in R (Handcock et al. 2008). We began by generating networks constrained on the empirically observed degree distribution (i.e., NUMGIVEN in the data). This also implicitly constrains the baseline network on total volume. We then seeded the network with the sampled respondents. The demographic characteristics of the sampled respondents were mapped onto nodes in the network with the same degree as the respondent (see Smith 2012). This maintains the correlation between demographic characteristics and degree. Thus, highly educated people in the simulated network will have high degree if the sampled respondents with high degree are highly educated. This seeding process also ensures that the generated network will reflect the demographic composition in the data. Thus, the simulation will generate a network that represents random mixing in the population, given the degree distribution, differential degree, and the demographic composition of the population. We repeated this process for both 1985 and 2004. In each case, the simulated networks are size 10,000 (it is impossible to simulate a network of the true size, 200 million or so).

We took a sample of ego networks from the simulated network the same size as the original GSS sample for that year—thus mimicking the true sampling process. We then took all *ij* pairs from the ego networks drawn from the simulated network and calculated the demographic distance between *i* and *j* (e.g., racial or religious matching). We compared the demographic distance in the observed ego network data to the demographic distance from the simulated network, capturing chance expectations.

For race, religion, and sex (the categorical variables), we compared the odds of a tie matching demographically in the observed data to the odds of a tie matching in the simulated network. We report how many times the 1985 ratio ($\log(\text{odds}_{\text{observed}}/\text{odds}_{\text{chance}})$) is larger than the 2004 ratio, indicating a decrease in in-group bias (relative to chance). This analysis mirrors a simple CUG (conditional uniform graph) test, and we report results for 1,000 bootstrap samples. We also report a second, alternative summary measure, based on the ratio of frequency counts: log((# Observed Ties Matching)/(# Observed Ties Mismatching)). This ratio is calculated net of chance expectations, based on the simulated network, and compared across 1985 and 2004. We again report how many times the 1985 ratio is larger than the 2004 ratio.

For the continuous measures, age and education, we calculated the ratio: $\log(\text{\# Ties Observed}_{\text{Distance=x}}/\text{\#Ties Chance}_{\text{Distance=x}})$.The ratio compares the number of ties in the observed ego networks to the number of ties in the simulated network at a given education or age distance, x. We then see how much an increase in demographic distance lowers the ratio of observed to chance frequency counts. Larger decreases, on average, mean stronger effects of increasing demographic distance. Formally, we focus on the marginal (or average) effect of increasing demographic distance by calculating the ratio as demographic distance increases by 1 and then averaging over those marginal effects. Again, we report how many times the 1985 ratio is larger than the 2004 ratio. Absolutely larger 1985 values mean homophily decreased.

Results

The results presented here mirror the results reported in the main text. There is a significant decrease in gender homophily, where the odds ratios are larger in 1985 than in 2004. The age and education results show no statistically discernible differences across years (although both lean toward an increase in homophily, as in Table 2 in the main text). Religion shows a significant increase in homophily using one summary measure but not the other, mirroring results in the main text, which show a significant increase under some specifications but not others. The religion results remain inconsistent, while pointing to a possible increase in homophily. The only major difference is with race. Here the results indicate a possible decrease in homophily, although the odds ratio results are more inconsistent across samples. The race interaction is, however, never significant in the results reported in the main text.

The racial differences result from the conditioning on differential degree. Because Whites have more ties on average than non-Whites, White-White ties are more frequent in the controls when degree is allowed to vary across demographic groups. More generally, homophily will appear weaker (relative to chance expectations) when degree differences are taken into account. This process is somewhat more exaggerated in 2004 than in 1985. This means that more of the racial matching can be explained by degree differences in 2004; or, once one “controls” for the differences in degree by demographic group, there is a larger decrease in in-group bias for race.

Looking over all the evidence, we do not believe there has been a decrease in racial homophily relative to chance. This is the only set of results that show a decrease in racial homophily and they are highly conditioned, controlling for both differential degree and the degree distribution. Additionally, the results are particularly dependent on the degree information in the data, and we know the 2004 degree information is problematic (given the inflated number of isolates) (see also note 15 in the main text).

Table S4. Comparing Observed Homophily to Homophily in Simulated Networks, 1985 to 2004

	Homophily Measured by Odds Ratio	Homophily Measured by Frequency Ratios	Homophily Measured by Frequency Ratios: Continuous Version
	Number of Samples with Homophily Increase: 1985 > 2004		
Race	74	1	NA
Religion	999	834	NA
Sex	1	1	NA
Age	NA	NA	940
Education	NA	NA	728

Note: Values correspond to the number of bootstrap samples where there is an increase in homophily. We had a total of 1,000 samples, so a value above 975 is strong evidence for an increase in homophily. A count below 25 is strong evidence for a decrease in homophily.

Part D. Case-Control Results Including Isolates in Controls

Table S5. Case-Control Logistic Regression Using All Reported Ties, Including Isolates in Controls, Univariate Analysis

Variable	Intercept	Dimension	Year	Dimension x Year	*N* (dyads)
All Ties					
Different Race	–16.792***	–2.105***	–0.286***	.191	1,130,856
	(.023)	(.115)	(.040)	(.145)	
Different Religion	–16.726***	–1.273***	–.195***	–.272**	1,130,856
	(.029)	(.055)	(.048)	(.084)	
Different Sex	–17.209***	–0.388***	–.500***	.139**	1,130,856
	(.020)	(.029)	(.034)	(.044)	
Age Difference	–16.404***	–.051***	–.438***	–.003	1,130,856
	(.025)	(.002)	(.047)	(.003)	
Education Difference	–16.735***	–.199***	–.438***	–.013	1,130,856
	(.030)	(.009)	(.053)	(.017)	

Note: Standard errors are in parentheses; they were calculated using bootstrap estimates. Standard errors are equal to the standard deviation of the coefficients across 1,000 iterations and are thus not dependent on the number of dyads. For each iteration, we took a random sample of respondents from each year and reran the case-control logistic regression.

**p* < .05; ***p* < .01; ****p* < .001 (two-tailed tests).

Table S6. Case-Control Logistic Regression Using All Reported Ties, Including Isolates in Controls, Multivariate Analysis

	All Ties	
Variables	Model 1	Model 2
Intercept	–14.376***	–14.444***
	(.043)	(.050)
Different Race	–1.941***	–2.038***
	(.076)	(.119)
Different Religion	–1.355***	–1.263***
	(.043)	(.057)
Different Sex	–.325***	–.380***
	(.024)	(.030)
Age Difference	–.051***	–.049***
	(.002)	(.002)
Education Difference	–.173***	–.161***
	(.008)	(.011)
Different Race x Year		.181
		(.152)
Different Religion x Year		–.215*
		(.089)
Different Sex x Year		.142**
		(.047)
Age Difference x Year		–.005
		(.003)
Education Difference x Year		–.034
		(.019)
Year	–.158***	–.045
	(.046)	(.086)
N (respondents)	3,001	3,001
N (dyads)	1,130,856	1,130,856
–2 x Log-likelihood	72634.862	72597.424
AIC	72648.862	72621.424
BIC (*N* based on dyads)	72732.431	72764.686

Note: Standard errors are in parentheses; they were calculated using bootstrap estimates. Standard errors are equal to the standard deviation of the coefficients across 1,000 iterations and are thus not dependent on the number of dyads. For each iteration, we took a random sample of respondents from each year and reran the case-control logistic regression.
**p* < .05; ***p* < .01; ****p* < .001 (two-tailed tests).

Part E. Case-Control Results Using Data from 2010 GSS Survey Experiment

This supplement presents results of an alternative analysis of homophily change. In the main text, the analysis compares homophily rates from the 1985 GSS to the 2004 GSS. Past work shows the 2004 data contained a disproportionate number of isolates, or individuals claiming no close confidants (Paik and Sanchagrin 2013). The GSS embedded an experiment in the 2010 survey to undercover the source and magnitude of this bias. Individuals were asked the same ego network questions as in previous years, but were randomly assigned to three survey conditions: one mimicking the 1985 survey (where the network questions came earlier in the survey); one mimicking the 2004 data (where the network questions came later in the survey, after a battery of voluntary association questions); and one that mimicked neither the 1985 nor 2004 survey.

We exploit this experiment as a way of validating our results on an independently collected dataset. Here, we reran the analysis using the 2010 data. We limited the sample to individuals who received the 1985 survey design. This analysis does not use the 2004 data. The 2010 data are limited by small sample size and scant demographic information (the survey only asks about race and gender), but it still offers an ideal robustness check for the main results—the 2010 data are directly comparable to the 1985 data in terms of survey design.

Table S7 presents results for race and gender. The general findings are the same as with the 2004 data: there is no change in racial homophily but a decrease in gender homophily. The *racial homophily* x *year* coefficient is smaller than with the 2004 data, but our overall conclusions are not affected by the overinflation of isolates found in the 2004 data.

Table S7. Case-Control Logistic Regression Using All Reported Ties, Using 2010 GSS Instead of 2004 GSS, Multivariate Analysis

	All Ties	
Variables	Model 1	Model 2
Different Race	–1.953***	–1.934***
	(.082)	(.098)
Different Sex	–.395***	–.437***
	(.028)	(.032)
Different Race x Year		–.062
		(.175)
Different Sex x Year		.220**
		(.068)
	(.052)	(.070)
N (dyads)	169,970	169,970
–2 x Log-likelihood	44160.067	44150.019
AIC	44168.067	44162.019
BIC	44208.24	44222.279

Note: Standard errors are in parentheses; they were calculated using bootstrap estimates. Standard errors are equal to the standard deviation of the coefficients across 1,000 iterations and are thus not dependent on the number of dyads. For each iteration, we took a random sample of respondents from each year and reran the case-control logistic regression.
$^*p < .05$; $^{**}p < .01$; $^{***}p < .001$ (two-tailed tests).

References

Handcock, Mark S., Steven M. Goodreau, David R. Hunter, Carter T. Butts, and Martina Morris. 2008. “ERGM: A Package to Fit, Simulate and Diagnose Exponential-Family Models for Networks.” *Journal of Statistical Software* 24:1–29.

Paik, Anthony and Kenneth Sanchagrin. 2013. “Social Isolation in America: An Artifact.” *American Sociological Review* 78:339–60.

Smith, Jeffrey A. 2012. “Macrostructure from Survey Data: Generating Whole Systems from Ego Networks.” *Sociological Methodology* 42:155–205.