ELSEVIER

Social Networks 22 (2000) 45–64

SOCIAL NETWORKS

www.elsevier.com/locate/socnet

Personal network integration: transitivity and homophily in strong-tie relations

Hugh Louch *

Princeton University, Princeton, NJ 08544, USA

Abstract

This article combines studies of transitivity and homophily in an empirical analysis of personal network integration. Using a national sample of individual's personal networks, the paper reveals that transitivity explains a majority of cases of network integration with two important caveats: (1) recent work on networks and social structure points to important structural constraints on personal networks that shape their formation, and (2) homophily (e.g. sex or race) and choice homophily (e.g. religion) improve the likelihood of integration in personal networks. The results indicate that the second finding interacts with relationship length and the availability of focal points to organize individual contact. © 2000 Elsevier Science B.V. All rights reserved.

1. Introduction

Sociologists have long been interested in the homogeneity of groups (Blau, 1977) and homophily of dyadic relations (McPherson and Smith-Lovin, 1987). Consistently, they find that personal relations are more homophilous than chance would suggest (Marsden, 1988), though, this is in good part a function of the relative sizes of different sub-groups in a population (Blau, 1977; Blau et al., 1982; Feld, 1982; Blum, 1985; McPherson and Smith-Lovin, 1987). Less common are studies that look inside the personal networks of individuals and extend the macro perspective of structural theorists like Blau (1977) to the micro level. Those that do typically examine dyadic friendship choices (McPherson and Smith-Lovin, 1987; Marsden, 1988) or constrained situations like organizations (McPherson and Smith-Lovin, 1987).

While several studies demonstrate that the variety of dyadic relations depends greatly on the available diversity of reference groups (Blau, 1977; Feld, 1981; McPherson and Smith-Lovin, 1987), we know little about the relationship between dyads and networks (Parks et al., 1983). For example, while we know that marital partner choice is greatly

* Tel.: +1-415-291-4432.

E-mail address: Louch@ppic.org (H. Louch).

PII: S0378-8733(00)00015-0

constrained by the diversity of the eligible pool (Blau et al., 1982), we know comparatively little about how well individuals' spouses are integrated into their wider networks (cf. Parks et al., 1983), and practically, nothing about how homophily affects this integration.

Coming from the opposite direction, network analysts have long debated the issue of network integration using the language of balance theory and transitivity (Heider, 1946; Cartwright and Harary, 1956; Newcomb, 1968; Holland and Leinhardt, 1970, 1971). Similar to its mathematical cousin, transitivity posits that if a chooses b as a friend ($a \rightarrow b$) and b chooses c as a friend ($b \rightarrow c$), then a will choose c as a friend ($a \rightarrow c$) (Holland and Leinhardt, 1970, 1971). Empirical studies have relatively and consistently found that the principle of transitivity applies in about 70–80% of all cases across a variety of small group situations (Davis, 1970, 1979; Robinson and Balkwell, 1995).

While some have regrets about the failure of transitivity as a deterministic structural principle (Davis, 1979), its usefulness as a tool for social analysis should be well established. For several reasons, however, this tool does not have wide use. First, the application of transitivity to small group analysis has hampered its wider use in sociological research. The original Davis/Holland/Leinhardt sociograms were contained groups, though some were decently large (Davis, 1970). Other studies have sampled groups that are heavily bounded by organizational structures, such as undergraduates (Price et al., 1966; Crano and Cooper, 1973; Robinson and Balkwell, 1995), schoolchildren (Leinhardt, 1972; Tuma and Hallinan, 1979; Hallinan and Hutchins, 1980), or work groups (Bishop, 1979; Segal, 1979). Davis (1979) noted this slide from structural theory to microanalysis, decrying the fact that their theoretical models seemed to ''best fit with junior high school students'' (p. 61).

Second, transitivity theorists have been largely uninterested in the relationship between the structure and content of the relations they studied (Parks et al., 1983). When relational content has appeared, it has been treated tentatively (Hallinan and Hutchins, 1980). Few studies have asked if transitivity in unbounded personal networks is affected by the attributes of the members of the network. Are friendships that are made across social boundaries (e.g. race, gender, or age) more likely to be kept separate (i.e. intransitive) than those that are more similar?

The paper addresses the relationship between homophily and transitivity in light of three mechanisms of friendship formation. First, I consider the number of associates a given dyad shares, an extension of the transitive principle. Second, I consider time-based constraints that limit an individual's ability to manipulate his or her networks. And finally, I consider a set of constraints that derive from what Feld (1981) refers to as the ''focused organization of social ties.'' Since most friendships form around specific situations, I control for shared roles and group memberships that provide settings for friendship formation. I analyze these issues using data on strong-tie networks from the 1985 General Social Survey (GSS).

2. Integration and transitivity: defining the network

I define network integration in terms of connection between pairs of individuals (dyads) within a network. When all individuals in a network are connected, integration is complete. At the level of the network, this is a measure of density. Since we want to understand the relationship between an individual's attributes and his/her integration into the network, however, density is a poor measure. Instead, I examine sets of triads from a relatively bounded personal network; an individual is integrated into the network when triads are complete, that is, transitive.

Above, I defined transitivity for asymmetric relations: if $a \rightarrow b$ and $b \rightarrow c$, then $c \rightarrow a$. This is the classic definition of transitivity set out by the Holland/Leinhardt/Davis studies (Davis, 1970). Here, however, I have referred to a less sophisticated version where all relations are symmetric (i.e. if $a \leftrightarrow b$ and $b \leftrightarrow c$, then $a \leftrightarrow c$), a subset of those identified by Davis (1970). [1] While a more complete analysis would deal with all triad types, data limitations resulting from the attempt to generalize transitivity to a wider variety of real-world situations necessitates starting with the symmetric case first. In this paper, I will use the word transitive to refer to the empirical situation where mutual connection exists between three individuals in a network. When I refer to the principle or theory of transitivity, I will do so explicitly.

This paper attempts to cut a middle ground between the two past types of network studies: small world experiments that provide interesting tests of the degree of connection between random pairs of individuals (Milgram, 1967; Travers and Milgram, 1969; Bernard and Killworth, 1979), and past studies of transitivity that have been limited to relatively small, contained groups where everyone is likely to associate in some way with everyone else (Bishop, 1979; Segal, 1979; Tuma and Hallinan, 1979; Hallinan and Hutchins, 1980). Neither of these two options accurately gets at real personal networks that are drawn from specific, but varied sources, and not just single organizations or far-flung acquaintances.

I further limit my analysis by examining only strong-tie networks. While most people belong to relatively expansive personal networks, many of an individual's weakly tied connections have little cause to associate with the other members of that individual's network. As Granovetter (1973) noted, the very strength of weak ties is the fact that they are not integrated into networks, but tend to bridge them. These ties are interesting when studying information flow, as in job search behavior (Granovetter, 1994), but are less interesting for network integration, since we expect them, by their very nature, to be poorly integrated. An appropriate middle ground studies relatively complete personal networks, where all individuals in the network have a potential basis to associate with one another, but are not necessarily brought into contact, as they would be in an organization. This facilitates studying the way individuals attempt to manipulate their social networks within the constraints that affect their everyday lives.

[1] Davis (1970) identifies four symmetric triad types, three of which are transitive (3-0-0, 1-0-2 and 0-0-3), while the other is intransitive (2-0-1). Here, I only deal with the first of the three transitive types, as the other two are not possible to find in ego-based data and are relatively uninteresting — one is completely empty, and the other, two-thirds empty, presenting little reason for individuals contained within to associate in real-world settings. In contained small groups, however, these are of much greater interest.

3. Shared associates, structural constraints, and homophily

I present four sets of independent variables to explain variation in the incidence of integration in personal networks. The first is simply an extension of the principle of symmetric transtivity, though it takes a somewhat different form. In its base form, the theory of symmetric transitivity argues that the connection between pairs of individuals is a function of their connection to a third individual. I extend this framework by examining multiple possible connections instead of a single possible connection. That is, the greater number of associates that two individuals have in common, the more likely a connection between them. This also keeps sight of the network better than when a triad is examined in seclusion.

The second and third sets of predictors examine constraints on the transitive principle. Since many relations that should be transitive are not (Davis, 1979; Robinson and Balkwell, 1995), it is important to consider constraints on the pure transitive model. Two types of constraints are relevant: time-based, or strength-of-tie constraints; and shared roles, or focus constraints (Feld, 1981). These constraints expand the framework of transitivity to take into consideration the limits that present choices about relationships create for future choices (Zeggelink, 1994).

The final set of predictors examine the relative importance of homophily as a predictor of triadic closure. Given transitivity and structural constraints, are individuals who are similar more likely to be connected than those who are different? This forms the core question that the paper addresses.

3.1. Transitivity for triads and groups

In the symmetric theory of transitivity, a connection between a given dyad is a function of the ties between the two other dyads in the triad:

$$t_{ab} = f(t_{xa}, t_{xb}) \tag{1}$$

where $t_{ab} = 1$ if the two alters, *a* and *b*, are connected, and zero otherwise. Heider (1946) argued that balance (from which transitivity was derived) was the result of a fundamental human need to maintain cognitive consistency, and that intransitive triads are a source of psychological distress. Other more recent treatments view transitivity in terms of more structural factors, such as the strength of ties (Granovetter, 1973), structural limitations on individual agency (Feld, 1981; Zeggelink, 1994), and peer influence (Hallinan and Hutchins, 1980).

While transitivity sometimes fails to predict connections with complete accuracy (Davis, 1979; Robinson and Balkwell, 1995), it does help explain the nature of connections between individuals. It is not unreasonable to expect that individuals will introduce their friends to one another. Thus, mutually shared associates will make a connection between two individuals more likely. To fully understand this argument, we need to think of transitivity more in terms of a larger group, and not just a triad. If one mutual connection increases the probability of a tie between two individuals, multiple

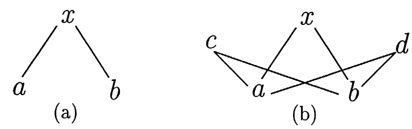

Fig. 1. Transitivity for a triad and a group.

mutual connections should increase that probability even more. This idea is diagramed in Fig. 1.

Panel (a) in Fig. 1 is the classic case of transitivity. A tie between x and a and between x and b increases the likelihood of a tie between a and b — a transitive triad. Panel (b) complicates things slightly by adding two additional actors — c and d. These additional ties ($c \leftrightarrow a$, $c \leftrightarrow b$, $d \leftrightarrow a$, $e \leftrightarrow b$) further reinforce the likelihood of a tie between a and b.

Eq. 1 can be redefined as follows:

$$t_{ab} = f\left(t_{x_m a}, t_{x_m b}\right), \tag{2}$$

where the extra subscript m ($m = 1, \ldots, M$) notes each of the individuals to whom either a or b is connected. This makes the tie between a and b dependent on ties to multiple individuals. In addition to restricting the study to cases where the dyad shares a single associate (where $t_{xa} = 1$ and $t_{xb} = 1$), I can measure the extent of closure for all the other triads in a network. This suggests the following hypothesis:

Hypothesis 1. *The greater number of shared associates for two individuals, a and b, the greater the likelihood of connection between them.*

3.2. Strength-of-tie constraints

Recent attempts to understand transitivity examine constraints to the transitive principle. My argument follows from two types of constraints. The first examines the strength of ties. Granovetter (1973) argues that because sustaining relationships requires energy, limits on time and emotional resources will reduce the strength of the bond between individuals. The strength of ties is interesting because weakly tied individuals tend to bridge different networks and are therefore less integrated into any particular network (Granovetter, 1973). Thus, transitivity is likely to vary by the strength of ties between individuals in the network.

Two mechanisms affect transitivity through strength-of-tie constraints. First, the frequency of contact between an individual and his/her alters will increase the probability of contact between the alters. Since individuals have limited time, they are limited in the number of strong ties they can have. An individual with several strong ties will likely spend time with them at the same time (Parks et al., 1983). Granovetter (1973) also notes that strong ties are unlikely to bridge different networks. Social relationships that are well-established and long-lasting tend to settle more often into transitive patterns (Hallinan and Hutchins, 1980). This suggests two measures of constraint: (1) regularity of contact between an individual x and two alters a and b; and (2) the length of time x has known both a and b.

> **Hypothesis 2(a).** *Increasing the strength of ties between an individual x and two alters a and b — measured by frequency of contact and length of time known — increases the probability of connection between the alters.*

Network size also constrains available time, and thereby, affects the probability of connection between pairs of alters in a triad. This constraint follows logically from the previous discussion. Large networks will create additional time pressures that individuals will have to manage, assuming a fixed level of time devoted to network contact. They can do this by spending less time with people or by spending more time with certain individuals simultaneously. By controlling for the former (Hypothesis 2a), network size will most likely measure the effects of limited time, increasing the probability of triadic closure in larger networks.

> **Hypothesis 2(b).** *Increasing network size for an individual x will increase the probability of triadic closure for each pair of alters, a and b, in the network.*

3.3. Focus constraints

The second structural limit arises from the patterned nature of relationship formation. Feld (1981) argues that most ties develop around specific foci, including ‘‘persons, places, social positions, activities, and groups’’ (p. 1018). As individuals make choices about who to be friends with, they build a friendship structure which acts on future choices they might make (Zeggelink, 1994). I already noted that in order to examine integration, there must be at least a single common center — in this case, a specific person with whom an alter pair commonly associates. This framework, however, can be extended to deal with multiple overlapping roles and group memberships.

There are two ways to measure the effects of focus constraints on network integration. The first measure directly taps shared roles. When an individual selects multiple friends from a group, these friends are highly likely to be friends with one another. On the other hand, if an individual chooses friends from widely different arenas of social life — one is a neighbor, one is a coworker, one is a relative, and none share any status — they are less likely to know one another.

Another way to think of this is as follows. Groups of kin, neighbors or coworkers are likely to know one another. The extent to which a specific individual’s (x’s) personal network overlaps with a particular role or group (i.e. all of x’s friends are his/her coworkers) the greater the likelihood of connection for every other pair of individuals in the network (because they are all neighbors or coworkers).

The second measure is more indirect. Foci are likely to be time-contingent. After a focus dissolves, however, the relationships that developed from it may well continue. When a person moves out of a neighborhood or leaves a job, they often maintain ties to people who remain. Individuals commonly maintain relationships with their kin even when distance separates them. Thus, we will expect that relations made at similar points in time will be more likely to have revolved around similar foci at some point in time, and these people will likely be in contact as well. These two measures of focus constraints suggest a third hypothesis.

> **Hypothesis 3(a).** *When both members of the alter pair share a specific role or group memberships (neighbor, coworker, group member, or kin) with the network ego, contact between alter pair is more likely than when they do not.*

While shared roles and group memberships should increase the incidence of transitive triads, we might also think that alters who occupy different statuses, or niches, in an individual's personal network will decrease it. That is, in Feld's (1981) terms, people participate in multiple foci — work, family, etc. — and often times, connections are not made across these groups. Similarly, friends who hold a higher status in an individual's mind may be actively kept separate from those that hold similar statuses. This suggests an extension of Hypothesis 3a.

> **Hypothesis 3(b).** *When the alter pair are drawn from widely different groups or roles (such as kin and coworker), or a status differential exists between them, contact between alter pair is less likely.*

While these constraints help prevent or ensure the transitivity of triads, the constraints are structural, deriving from the limits on the social organization of human activity (Feld, 1981). Section 3.4 investigates more deeply the content of those relations through the effect of homophily on triadic closure.

3.4. Homophily and transitivity

Pure transitive theory is inattentive to the choices that individuals make when forming their personal networks. The constraints discussed above suggest several important limits on when transitivity in real world situations will and will not occur. Although the limits on time and behavior apply regardless of individual attitudes and preferences, they do not deal directly with the way social boundaries affect network formation.

Marsden (1988) notes that personal networks tend to be homophilous. Further, Fischer (1982) and Marsden (1987) note that more dense networks tend to be homogeneous — sets of alter pairs are homophilous. As suggested above, density relates directly to network integration. As density increases, the number of transitive triads in the network also increases. Thus, there are some empirical grounds for associating increased homophily with transitivity.

Density and transitivity are not identical, however. Density is a group phenomena, while transitivity relates to specific sets of individuals (triads) in the network. Put another way, just because more homophilous networks tend to be more dense does not mean that individuals who are different are network isolates and those that are the same are well connected. How well these individuals are integrated into networks is an empirical question.

The best evidence that homophily affects density comes from two sources. The first, mentioned above, documents the tendency for individuals to have friends similar to themselves (Fischer, 1982; Marsden, 1988). A second source examines preferences towards homophily within a specific organization (McPherson and Smith-Lovin, 1987; McPherson et al., 1992; Ibarra, 1992). However, none of these studies adequately control for structural constraints and shared associates.

Marsden (1988) identifies several dimensions of homophily which are important: race, education, age and religion. Additionally, he finds limited evidence for the importance of gender. While these results analyze dyads instead of triads, they suggest several useful dimensions for analyzing the effects of homophily.

> **Hypothesis 4.** *The racial, age, educational, gender, and religious homophily of an alter pair (ab) increases the probability of a connection between them, controlling for the effects of shared associates and structural constraints.*

4. Data and methods

The data for this paper come from the 1985 (General Social Survey) GSS. The GSS is a national probability sample of the english-speaking non-institutionalized population of the United States. In 1985, the GSS included a special module on respondent's personal networks (details can be found in Burt (1984)). Respondents were asked to name all the individuals, or alters, they ''spoke to about important matters in the last 6 months''. Of these, interviewers elicited information on the top five, including how long each alter was known; frequency of contact with each of the alters; attributes of each alter — race, gender, education, age, and religion; and whether or not each pair of alters knew each other. Respondents were free to mention fewer than five.

These data are useful to examine the issue of transitivity because the respondent-alter ties are, in at least one sense, strong ties. Emotional intensity and intimacy are two of the four criteria that Granovetter, (1973) identify for the strength of a tie. Discussing important matters is a far more intense and intimate bond than, for example, casual conversation.

The unit of analysis for this paper is a triad where one member is always the respondent and the other two are each set of alter pairs drawn from the respondent's personal network. This increases the sample size for the models from 1534 to, potentially, 15,340. If each individual had five alters, there would be 10 combination of alters per respondent and 15,340 units of analysis. However, because not every respondent named five alters, and there was some missing data, the total number of alter pairs is 5017. Descriptions of the data are given in Table 1, as well as the distribution of the dependent variable for each of the independent variables.

Table 1
Descriptive statistics for independent variables

		Frequency (N)	Closed triads (%)
Total		5017	81
Shared associates	…zero	911	50
	…one	1499	79
	…two	1273	88
	…three	1334	97
Respondent talks to	…both rarely	312	76
	…one often	1252	71
	…both often	3453	85
Respondent has known	…neither long	558	74
	…one long	1474	66
	…both long	2985	90
Network size	…4 or less	015	84
	…5 or more	3002	79
Alters are kin	…yes	1070	95
	…no	3947	77
Alters are neighbors	…yes	133	89
	…no	4884	81
Alters are coworkers	…yes	368	90
	…no	4649	80
Alters are group members	…yes	4431	80
	…no	586	88
One alter is kin	…yes	2974	82
	…no	2043	78
One alter is advisor	…yes	4211	82
	…no	806	75
Racial homophily	…same	4808	82
	…different	209	62
Gender homophily	…same	2547	78
	…different	2470	84
Educational homophily	…same	2475	84
	…different	2542	78
Age homophily	…within 5 years	1749	82
	…more than 5 years	3268	80
Religious homophily	…same	3567	84
	…different	1450	73

4.1. *Measures*

The dependent variable for this analysis, transitivity (t), is measured by whether or not each pair of alters has any tie. Respondents were asked if the two alters were ‘‘especially close’’, ‘‘neither close nor strangers’’, or ‘‘total strangers’’. This variable is measured 1 if the alter pair is either ‘‘especially close’’ or ‘‘neither close nor strangers’’. 0 otherwise. Both alters always have a tie to a third individual, the respondent, which forms the triad. In the GSS data, 79% of the triads are transitive (see Table 1).

To measure shared associates, I counted the number of ties that the two alters shared in the network (excluding the tie all alter pairs share to the respondent). If *a* and *b both* had a tie to another of the respondent's alters (*c*, *d*, or *e*), they receive an additional point on this scale. The scale ranges from 0 to 3. Since respondents gave information for at most five alters (and each alter pair involves two of those), there can only be three possible shared ties for any given alter pair.[2] I expect that the number of shared associates will be positively related to the probability of connection between alter pairs.

Several independent variables test the strength-of-tie hypothesis. Frequency of contact between the respondent and the alter pair is a three-category variable: both infrequently (less than once a week); one frequently (at least once a week); and both frequently. Both infrequently is the reference category. Increasing contact should have a positive effect on the probability of transitivity.

The length of time the respondent has known the alter pair also takes three categories: both a short time (less than 6 years); one a long time (more than 6 years); and both a long time. Both a short time is the reference category. The length of time known should also have a positive effect on the probability of transitivity.

Network size is the reported network size that the respondent gave to the question of how many people they spoke to about important matters. This variable has a top code of six, for six or more individuals. However, only 84 of 1534 (5.4%) individuals mentioned six or more alters. Network size is expected to have a positive effect on transitivity.

Six variables measure the focus constraints hypotheses. The first four measure shared roles of the respondent and the alter pair: one each for coworkers, neighbors, and shared group membership; and the fourth, for kin. The variables take a value one if the alter pair both share the particular role (coworker, neighbor, group member, or kin) with the respondent, and a zero otherwise. The fifth variable captures situations where alters are drawn from different groups. It takes a value of one when one alter is kin and the other is not. The final variable measures status differential, taking a value of one when one alter is considered an ''advisor'' and the other is not. Alter pairs who are coworkers, neighbors, group members, or kin will more likely be connected than those who are not. Alters who are drawn from different groups or statuses are less likely to be connected.

To test the homophily hypothesis, I created a variable for each trait — race, gender, age, education, and religion. For race, gender, and religion, if the pair of alters had the same value for the trait (e.g. both women, both white, or both protestant), they received a 1, and a 0 otherwise. I grouped education into categories (less than high school, high school degree, college degree, graduate degree), and then constructed a measure for each alter pair as above. I measured age homophily by the absolute difference between the pair of alters' ages. I expect homophily to increase the probability of transitivity.

[2] This variable reaches different maxima depending on the size of the network that the respondent named. While a percentage of total named alters would eliminate this problem, it also makes the variable substantively less interesting. For example, it would render one mutual tie in a very small network equivalent to several mutual ties in a much larger network. Since the number of shared associates is the theoretical mechanism that affects transitivity, the count is a more sensible measure than the percentage.

4.2. Model specification

Since the dependent variable is binary, it is common to use a logistic regression model. This model is drawn from the generic generalized linear model of the form:

$$E(y_i) = p_i + e_i, \tag{3}$$

for all triads $(i = 1, \ldots, T)$ where $E(y_i)$ is the expected value of the dependent variable, p_i is the probability of the outcome, and e_i is a residual, or error term.

The standard logistic regression transforms p_i yields,

$$\operatorname{logit} p_i = \mathbf{X}_i \boldsymbol{b}, \tag{4}$$

where $\mathbf{X}_i$ is a matrix of independent variables including a constant and $\boldsymbol{b}$ is a vector of coefficients, one for each column in $\mathbf{X}$. A central assumption of this model is that each unit of analysis is sampled independently from every other unit; that each e_i is uncorrelated with every other unit. However, because these data are clustered by respondents, this conventional logistic regression model is not entirely appropriate.

Clustered data can best be accommodated by hierarchical models (Bryk and Braudenrush, 1994). When data are sampled from multiple respondents within a larger entity (for example, a school or a personal network), a hierarchical model can adjust for dependence among units from the same cluster. Results will tend to be more conservative than in the usual regression model that assumes independence among units.

For triad i $(i = 1, \ldots, T_j)$ of respondent j $(j = 1, \ldots, N)$, the hierarchical generalized linear model takes the form:

$$E(y_{ij}) = p_{ij} + e_{ij}, \tag{5}$$

and the standard logistic transformation takes the form:

$$\operatorname{logit}(p_{ij}) = X_{ij} b + a_j. \tag{6}$$

Here, the X_{ij}'s do not include a constant. Instead, we pick up the constant through a_j, the random effect. The j subscript indicates that a varies across respondents, and not triads. Rather than one constant, there is a separate constant for each personal network. This random effect captures differences in the probability of ties that are to be associated with individual respondents. In other words, it effectively decomposes the error term (e) into a standard unit residual and a respondent residual, controlling for the potentially misleading effects of clustering.

For each of the regression models I present a psuedo-R^2 as a goodness-of-fit test. The pseudo-R^2 is a rough measure of the relationship between variance of the fitted values from the regression and the variance in observed values of triadic closure. It is created by dividing the variance of the fitted values by the sum of the mean of the observed values minus the squared mean of the observed values.

5. Findings

5.1. Shared associates

The first model in Table 2 presents a hierarchical logistic regression that analyzes the simultaneous effects of all variables on the likelihood of a connection existing between

Table 2
Logistic regression results — likelihood of integration
Note: Numbers in parentheses are standard errors.

	Model 1	Model 2	Model 3	Model 4	Model 5
Intercept	0.423 (0.383)	0.378 (0.430)	0.850 (0.512)	−3.825* (1.300)	1.632 (0.996)
Shared associates	0.365* (0.057)	0.240* (0.061)	0.761* (0.087)	−2.747* (0.146)	0.789* (0.153)
Strength of tie constraints					
Talk to one often	0.192 (0.151)	0.292 (0.161)	−0.019 (0.249)	0.347 (0.246)	−0.590 (0.549)
...both often	1.024* (0.165)	1.098* (0.179)	0.869* (0.258)	2.924* (0.338)	0.059 (0.558)
Known one long	−0.272* (0.125)	−0.119 (0.146)	−0.336* (0.156)	–	–
...both long	1.249* (0.154)	1.461* (0.176)	0.770* (0.187)	–	–
Network size	−0.234* (0.064)	−0.220* (0.072)	−0.340* (0.077)	1.373* (0.276)	−0.425* (0.153)
Focus constraints					
Both alters kin	0.854* (0.204)	0.839* (0.217)	–	5.928* (0.583)	–
...neighbors	1.249* (0.373)	–	1.002* (0.408)	2.987* (0.693)	0.259 (0.610)
...coworkers	1.425* (0.200)	–	1.289* (0.243)	8.916* (1.135)	2.510* (0.634)
...group members	0.829* (0.174)	–	0.953* (0.233)	8.301* (1.036)	0.722* (0.364)
One alter kin	−0.245* (0.097)	−0.293* (0.106)	–	0.457* (0.201)	−0.311 (0.446)
...advisor	−0.349* (0.106)	−0.434* (0.120)	−0.270 (0.157)	−1.874* (0.248)	−0.223 (0.298)
Homophily					
Race	0.419* (0.181)	0.380 (0.210)	0.421 (0.254)	2.540* (0.402)	0.655 (0.440)
Gender	−0.447* (0.072)	−0.488* (0.078)	−0.570* (0.121)	−0.676* (0.132)	−0.203 (0.225)
Education	0.299* (0.078)	0.359* (0.086)	0.319* (0.122)	0.883* (0.146)	−0.189 (0.239)
Age	−0.029* (0.010)	−0.025* (0.011)	−0.020 (0.018)	−0.036 (0.020)	−0.060 (0.033)
Age^2	0.001* (0.000)	0.001* (0.000)	0.000 (0.001)	0.001* (0.001)	0.001 (0.001)
Religion	0.375* (0.084)	0.337* (0.091)	0.060 (0.129)	0.132 (0.151)	0.408 (0.250)
N	5017	4068	1904	2985	558
Pseudo-R^2	0.366	0.218	0.307	0.347	0.366

$^*p < 0.05$.

the alter pair. This model indicates that the greater number of additional connections that the two specific alters share increases the probability that they will be connected to one another. Each additional shared tie increases the integration by nearly 150%.

Transitivity is clearly a strong structural principle. If *a* and *b* are connected to multiple other individuals in the network, these connections reinforce the pressure towards transitive triads that the respondents (x) strong ties to *a* and *b* create. When many of the other triads in a network are transitive, then *a* and *b* are deeply embedded into the structure of the network, and we expect a tie between them. If anything, these results under-estimate the effects of transitivity, since potential other individuals exist who are connected to both *a* and *b*, but who were unknown to the respondent *x*.

Given the importance of transitivity, two questions remain: (1) how much do constraints further affect transitivity; and (2) given both shared associates and constraints, what is the additional importance of homophily for transitivity. These two questions are addressed in turn.

5.2. Strength-of-tie constraints

The estimated effects of the strength-of-tie constraints on the probability of each triad being transitive are largely as predicted, though there are interesting exceptions. First, when both alters are spoken to frequently or have been known a long time, connection between them is much more likely (178% and 248%, respectively). However, regular contact with only one alter is insignificant, and having known only one alter for a long time actually reduces the probability of connection (by almost 25%). Furthermore, network size has a distinctly negative effect.

The negative effect of network size can be partly explained by its strong correlation with the shared associates variable ($r = 0.70$). Since the data are ego-based, a larger network creates greater possibilities for the alter pair to share additional associates. In smaller networks, we are probably more likely to miss individuals who are connected to the respondent, but not mentioned. Fitting a model without shared associates, but with network size, generates a significant positive result for network size. However, there is some indication that network size may proxy a reduced multiplexity of relationships. Respondents may divide up their time with different individuals to achieve separate purposes.

It is notable that the effect for length of acquaintance remains non-linear from the summary statistics (Table 1) to the regression model (Table 2). This supports the focus constraints mechanism — friends who have been known similar amounts of time were likely made in similar contexts (note the effect of neighbors, coworkers and group members, discussed below). Thus, friends made at different times are less likely to have contact with one another than those made at similar times, even when the latter-set friendships were developed relatively recently. This idea is tested more formally below. At the same time, it is clear that alter involvement in the respondent’s life plays a strong role in shaping networks. The effects of regularity of contact and length of time known are among the strongest. Friends who are known a long time just know more about a respondent’s group of friends and are more likely to have met each of them.

5.3. Focus constraints

Shared roles and group memberships also significantly increase the probability of alter connection. When both alters in the pair are kin, neighbors, coworkers or members of the same group or organization, the likelihood of connection between alter pairs increases substantially. The effects are strongest for coworkers (316% more likely) and neighbors (249% more likely), though they are not small for kin and group members either (each over twice as likely). The range of strengths probably reflects something of the nature of the boundaries implied by the different measures. For most people, work is a very bounded environment, while the definition of a neighborhood or a voluntary group is more fluid (Hunter, 1980). Thus, two people a respondent calls neighbor will, on average, be from a more dispersed area than two individuals who share a single workspace, making contact less likely. The fact that the kin effect is not as strong as the others is in many ways remarkable, but explicable. Since kin alters are more likely to be the same race and religion, and to have been known similar lengths of time, this effect may well be diluted. I return to this idea below.

We have strong evidence that structural constraints on time and potential for contact help predict when triads are transitive. Understanding the nature of individual contact always necessitates paying attention to constraints on action.

5.4. Homophily

Given the control for shared group ties and the constraints of model 2, few structural theorists would argue that homophily on any dimension would affect the propensity towards transitivity. Yet model 4 provides strong support for the homophily hypothesis. Excepting gender, the effects are as predicted. Since gender presents a contrary case, I will address it after the effects of the other traits.

Race, education, and religion have relatively similar effects. When alters share one of these attributes, the likelihood of contact increases significantly (52% more likely for race, 35% for education, and 45% for religion). In the internal dynamics of a personal network, however, race, education and religion are very different sorts of boundaries, one ascribed, one achieved, and the third largely a matter of choice. Most alter pairs are racially homophilous but educationally heterophilous, with religion between the other two (Table 1). Yet, connection between alter pairs is significantly associated with all three variables.

The absolute distance between alters in age decreases the likelihood of connection between them as predicted. This effect, however, is non-linear and begins to reverse itself as the distance becomes quite large (more than 50 years difference in ages). There are a number of possible explanations for this non-linearity. First, some of the alters named may be mentors or bosses, and therefore, span a large age gap. Second, some individuals have mixed age friendship circles and they will offset the strong positive trend. Finally, the non-linearity may pick up the existence of in-laws. Since spouse is not counted in the kin variable, a spouse's parents will be substantially older than the spouse and not be counted as kin (though in some sense they are).

One unexpected effect for these results is for gender. Alters of the same gender are actually significantly less likely to form connections than those who are different (over 35% less likely). Marriage, however, entirely explains this peculiar effect. I fit a model (not shown) which includes two additional variables: first, a measure of whether or not one of the two alters is the respondent's spouse; and second, an interaction between the spouse variable and gender homophily. In this model, gender homophily is not significant, but both spouse and the interaction are. The effect of spouse is positive — spouses are highly likely to know an individual's friends, potentially having been drawn from the same group. This effect is mitigated by a strong negative effect of the interaction between spouse and gender homophily. That is, when married respondent's have friends who are of a different gender, it is less likely that those friends will also be friends with the respondent's spouse.[3]

In this final model, gender and alter connection are only weakly related. Dissimilar alters are as likely to know one another as similar ones. Another way of thinking of this is that same sex alters are not kept apart any more than opposite sex alters. Marriage serves to make alter–alter contact more likely, but less so if they are of the same sex, especially for men.

The overall importance of homophily holds even after controlling for both the structural constraint produced by transitivity and constraints on time and location. Non-homogeneous relations are less likely to be integrated than homogeneous ones. Race, education, age and religion are strong cleavages across which contact between individuals becomes less likely. Gender is not nearly as important, except for married respondents.

5.5. Using stronger constraints

The above models provide substantial support for the importance of homophily above and beyond the effects of transitivity and constraints. However, several stronger constraints exist, which might provide a better test of the effects of homophily. Two types of stronger constraints are of interest here: (1) the contingency of the effects of homophily on focus constraints; and (2) the contingency of the effects of homophily on the strength-of-tie constraints.

5.5.1. Role constraints

Friendship formation in tightly bounded circles is different than in those that are less constrained. In the former, there may be pre-existing barriers to entry that keep certain individuals out. Neighborhoods and workplaces are relatively formalized entities, which often act to prevent access to entire groups of individuals on the basis of single attributes

[3] Fitting the model separately for male and female respondents indicates a strong effect for men, but not for women. One interpretation of this is that women are more likely to introduce opposite sex friends to their husbands than men are to their wives. I am grateful to Steven J. Tepper for suggesting that I pursue this line of analysis.

— e.g. race and education. As a result, the homophily variables will operate differently depending on the level of the shared group membership variables. The group membership variables may vary systematically with the homophily variables by constraining the pool of potential associates (Feld, 1981, 1982). The second model of Table 2 controls for the shared roles of the alter pair. The sample for this model contains only those cases in which the alter pair is neither neighbors, nor coworkers, nor group members.

As model 2 indicates, however, most of the effects of homophily remain important after controlling for shared roles. In each case, the direction and magnitude of the effects are consistent from model 1. Educational, religious and kinship similarities all increase the probability of transitive triads, while an increasing age gap has the opposite effect. Again, the effects of gender appear to work in the opposite direction. Controlling for spouse as above produces similar results. Notably, the race effect is non-significant, though it approaches statistical significance ($p < 0.10$). Finally, the effects of shared group ties and constraints produce similar results as in the previous models.

Model 3 presents a similar test for non-kin relations. The dimensions of homophily will likely vary systematically with kin and non-kin relations. As mentioned above, kin relations tend to more similar than friends racially and religiously, though more varied on gender and age. In this model, only those alter pairs where neither alter is the respondent's kin are contained in the sample. Here, there are some interesting differences in comparison with model 1. First, the gender effect appears more pronounced for only non-kin pairs. Second, both racial and religious homophily are non-significant, though in the appropriate directions. Racial homophily, once again, approaches close to the level of statistical significance ($p < 0.10$), but religion does not. The results seem to indicate that racial and religious boundaries are especially strong when they coincide with various foci, strongly supporting the work of Feld (1981). Age also becomes insignificant, though fitting the model without the squared term makes the regular term significant. Since kin are much more likely to be older, examining only non-kin is likely to contract the difference in ages (which it does by about 10 years, on average).

5.5.2. Strength of tie constraints

The other stronger constraint of interest is the conditionality of homophily on the strength of the tie. Because homophily is not simply an issue of individual choice, but also the result of strong societal pressures, relationships will likely be different over a longer period of time. That is, it may be more difficult for individuals who are different to remain in contact with one another over a longer period of time. Models 4 and 5 (Table 2) provide substantial support for this notion.

Model 4 contains only those alter pairs where the respondent has known both alters for a long time (more than 6 years). For longer-term friends, the importance of shared ties reverses itself (94% less likely). These seem to be well-established friendships to whom many of the structural principles of friendship *formation* no longer apply. Constraints, however, become more important — witness the immense coefficients for coworkers, group members, and kin.

The effect of homophily on transitivity changes as well. Racial similarity is much more important for contact for longer-term friends (almost 13 times more likely) than it was for everyone grouped together. Education, as well, has a larger effect. The other effects, however, remain consistent with the full model (model 1), though religion is insignificant.

For shorter-term friends (model 5), the effects are in the opposite direction. Here, none of the homophily variables significantly predict contact. This suggests the dynamic nature of network formation. Friends made recently are not only less bound into networks (see Table 2), but many of the pressures, which bind them in work much differently than for longer-term friends.

5.5.3. Summary

Using stronger constraints creates predictions which are, in most cases, similar to those obtained from earlier models. That the effects of homophily remain significant suggests that the effects of structure are neither prior nor necessarily more important than the effects of homophily. Only when considering the length of time the respondent has known the alters are there significant differences from the full model (model 1). Thus, across several dimensions, the homophily of the alter pair significantly affects the probability of transitivity. Excepting gender, alter pairs who are similar are considerably more likely to know one another than those who are different.

6. Discussion

The results presented here help synthesize two strands of analysis in the network literature: homophily and transitivity. While we know that ties tend to be transitive, especially in small group settings, transitivity has not been extended to less constrained situations. Similarly, while we know that individual's ties tend to be homophilous, we do not know the implications of the cases when they are not. We do not know how well different individuals are integrated into personal networks.

The results suggest a somewhat less strict definition of transitivity — as a continuum instead of a principle — and the importance of three mechanisms — strength-of-tie constraints, focus constraints, and homophily — that shape the transitive principle. The paper synthesizes several lines of previous work on network formation — the strength of ties perspective developed by Granovetter (1973); the network foci theory developed by Feld (1981); and the work on dyadic homophily (e.g. Marsden, 1988) — and brings them to bear on an important issue in social networks — the integration of individuals into personal networks.

The results are important because they address an important gap in social-structural reasoning. Structural analysis of inter-group relations has examined the question of inter-group contact at the macro and micro level, but rarely asked what happens after individual choices are made. The availability of individuals who are different acts as a structural limitation when choosing what type of persons to be friend with (Blau, 1977). Further, individual choices often do affect later ones (Zeggelink, 1994). Yet, individuals more often keep friends who are different apart; individuals who are dissimilar act more often as network bridges than those who are similar. It is important to keep in mind, however, that there is much individual agency in this formulation. While structural properties are considered, the focus of the paper is on choices that individuals make, an issue which is often absent in network analysis (Emirbayer and Goodwin, 1994).

The greatest limitation of the paper is in the quality of the data. First, the paper only examines one type of tie. One aspect of the strength of tie is its multiplexity, which refers to the number of overlapping types of relations that a particular tie consists of (Fischer, 1982). This paper examines only relatively strong friendship ties, and it would be valuable to examine others, and to examine several in combination. Different types of ties will most likely be generated in different sorts of situations (Feld, 1981), thus, making the analysis of multiple types of ties interesting.

A second data problem is that these data are ego-based — a single respondent gives all the information about each of the alters. This creates problems in terms of data quality. This means that other unnamed alters might potentially connect the alter to the respondent; that while integration is homophilous for triads, it might not be so when two-step ties are considered. At the same time, the small personal networks reported in this paper are likely the core of an individual's support network. They will have great influence over decisions that an individual makes, which makes integration at this level substantively more interesting. And, it is interesting in and of itself that certain boundaries, such as race, appear to be salient in the respondent's perceptions of their personal networks.

Third, the data are cross-sectional, yet I analyze boundaries, which are shaped over time. Jacobson (1985) notes that boundaries are likely to be extremely varied over time; that individual's friendship choices at one point in time will be quite different when they make the same choices again later. Because the data are on a random sample of Americans, however, at any given point in time, we can expect social networks to look as they do in this paper. However, over-time effects will play a role with changes in macro factors, such as race and gender relations.

Finally, several of the variables are potentially endogenous to the dependent variable. When two individuals are connected, the respondent will be more likely to spend similar amounts of time with them, or to have met them at about the same time. This will be especially true when the two alters are married, and therefore, quite likely to spend time with their friends together. The effects of the frequency of contact variable, and perhaps, the length of time known variable, will likely be inflated as a result.

More data collection is required. Data collected on several localities simultaneously where each of the individuals was questioned about their relationships to all others would be quite useful to examine the question in greater depth. However, because the expense of such data collection might be prohibitive, smaller-scale efforts, which examine multiple different kinds of ties and ask about multiple types of ties (see Fischer (1982) for a good example of this strategy) would also be useful.

The paper has clear implications for work in several areas of sociology. The results have bearings on studies of race and class, indicating that race and education work as independent, though complimentary, boundaries. Further, the comparable similarity of the non-kin results to the full model (kin and non-kin) has implications for the network studies of the family. Constraints and boundaries seem to affect kin relations in similar ways as non-kin relations.

References

Bernard, H.R., Killworth, P.D., 1979. A review of small world literature. Sociological Symposium 28, 87–100.

Bishop, J.M., 1979. Transitivity in work-relevant and sentiment-based sociograms. Pacific Sociological Review 22, 185–200.

Blau, P., 1977. Heterogeneity and Inequality: Towards a Primitive Theory of Social Structure. Free Press, New York.

Blau, P., Blum, T., Schwartz, H., 1982. Heterogeneity and intermarriage. American Sociological Review 47, 45–62.

Blum, T.C., 1985. Structural constraints on interpersonal relations: a test of Blau's macrosociological theory. American Journal of Sociology 91, 511–521.

Bryk, A.S., Braudenrush, S.W., 1994. Hierarchical Linear Models: Applications and Data Analysis Methods. Sage, Newbury Park.

Burt, R.S., 1984. Network items and the general social survey. Social Networks 6, 293–339.

Cartwright, D., Harary, F., 1956. Structural balance: a generalization of Heider's theory. Psychological Review 63, 277–293.

Crano, W.D., Cooper, R.E., 1973. Examination of Newcomb's extension of structural balance theory. Journal of Personality and Social Psychology 27, 344–353.

Davis, J.A., 1970. Clustering and hierarchy in interpersonal relations. American Sociological Review 35, 843–851.

Davis, J.A., 1979. The Davis/Holland/Leinhardt studies: an overview. In: Holland, P.W., Leinhardt, S. (Eds.), Perspectives on Social Network Research. Academic Press, New York, pp. 51–62, Chapter 4.

Emirbayer, M., Goodwin, J., 1994. Network analysis, culture, and the problem of agency. American Journal of Sociology 99, 1411–1454.

Feld, S.L., 1981. The focused organization of social ties. American Journal of Sociology 86, 1015–1035.

Feld, S.L., 1982. Social structural determinants of similarity among associates. American Sociological Review 47, 797–801.

Fischer, C.S., 1982. To Dwell Among Friends: Personal Networks in Town and City. University of Chicago Press, Chicago.

Granovetter, M., 1973. The strength of weak ties. American Journal of Sociology 78, 1360–1379.

Granovetter, M., 1994. A Job: A Study of Contacts and Careers. 2nd edn. Cambridge Univ. Press, Cambridge.

Hallinan, M.T., Hutchins, E.E., 1980. Structural effects on dyadic change. Social Forces 59, 225–245.

Heider, F., 1946. Attitudes and cognitive organization. Journal of Psychology 21, 107–112.

Holland, P.W., Leinhardt, S., 1970. A method for detecting structure in sociometric data. American Journal of Sociology 76, 492–513.

Holland, P.W., Leinhardt, S., 1971. Transitivity in structural models of small groups. Comparative Group Studies 2, 107–124.

Hunter, A., 1980. Symbolic Communities. University of Chicago Press, Chicago.

Ibarra, J., 1992. Homophily and differential returns: sex differences in network structure and access in an advertising firm. Administrative Science Quarterly 37, 422–477.

Jacobson, D., 1985. Boundary maintenance in support networks. Social Networks 7, 341–351.

Leinhardt, S., 1972. Developmental change in the sentiment structure of children's groups. American Sociological Review 37, 202–212.

Marsden, P., 1987. Core discussion networks of Americans. American Sociological Review 52, 122–131.

Marsden, P., 1988. Homogeneity in confiding relations. Social Networks 9, 100–125.

McPherson, J.M., Smith-Lovin, L., 1987. Homophily in voluntary organizations: status distance and the composition of face-to-face groups. American Sociological Review 52, 370–379.

McPherson, J.M., Popielarz, P.A., Dbronic, S., 1992. Social networks and organizational dynamics. American Sociological Review 57, 153–170.

Milgram, S., 1967. The small world problem. Psychology Today 1, 61–67.

Newcomb, T.M., 1968. Interpersonal balance. In: Abelson, R., Aronson, E., Mcguire, W., Newcomb, T., Rosenthal, M., Tannebaum, P. (Eds.), Theories of Cognitive Consistency: A Sourcebook. Rand McNally, Chicago.

Parks, M.R., Stan, C.M., Eggert, L.L., 1983. Romantic involvement and social network involvement. Social Psychology Quarterly 46, 116–131.

Price, K.O., Harburg, E., Newcomb, T.M., 1966. Psychological balance in situations of negative personal attitudes. Journal of Personality and Social Psychology 3, 265–270.

Robinson, D.T., Balkwell, J.W., 1995. Density, transitivity, and diffuse status in task-oriented groups. Social Psychology Quarterly 58, 241–254.

Segal, M.W., 1979. Varieties of interpersonal attraction and their interrelationships in natural groups. Social Psychology Quarterly 42, 253–261.

Travers, J., Milgram, S., 1969. An experimental study of the small world problem. Sociometry 32, 425–443.

Tuma, N.B., Hallinan, M.T., 1979. The effects of sex, race, and achievement on schoolchildren's friendships. Social Forces 57, 1265–1280.

Zeggelink, E., 1994. Dynamics of structure: an individual oriented approach. Social Networks 16, 295–333.