Heterogeneity, Consolidation, and Intergroup Relations
Tests of Blau’s Hypotheses

ABSTRACT: The present study uses data from the first wave of an ongoing panel survey study to test some hypotheses derived from Peter M. Blau’s macrostructural theory of social integration. More specifically, the tests reported in this article focus on the effects of heterogeneity on the rate of intergroup relations. According to Blau, an increase in heterogeneity contributes to an increase in social integration as measured by the rate of intergroup relations, so the more a community is differentiated by gender, the more male–female friendships there are. Blau also predicted that the effect of heterogeneity is weakened when differences along multiple lines overlap, so when gender is correlated with occupation and friends are chosen with respect to gender, there are fewer friendships between members of different occupations than there would be if gender and occupation were independent. I test these predictions against the data from the Polish Panel Survey POLPAN; its first wave was carried out in 1988 on a random and nationwide sample of 5,817 people ages twenty-one to sixty-five. The tests reported in this study use a version of Blau’s theory developed by Thomas J. Fararo and John Skvoretz, who translated the original theory into a mathematical equation. The results of the tests raise doubts as to whether the theory is an accurate description of the relationship between social differentiation and integration. More specifically, the rates of marital and friendship ties with respect to several dimensions of differentiation—such as education or occupation—turn out not to be predicted adequately by the models developed by Fararo and Skvoretz.

The objective of this study is to empirically test theorems derived from Peter M. Blau’s macrostructural theory of intergroup relations (Blau 1977, 1994; Blau and Schwartz...
The theorems in question are concerned with the effect of heterogeneity of a population along a given dimension or characteristic on the population’s social integration, as measured by the rate of social ties spanning salient social divisions along that dimension. The theorems belong to the theory’s most fundamental theorems and have inspired much research, both empirical (Blau, Blum, and Schwartz 1982; Blum 1985; Fitzpatrick and Hwang 1992; Rytina 1982; Skvoretz 1990b) and theoretical (Fararo 1981; Fararo and Skvoretz 1989; Heckathorn and Rosenstein 2002; McPherson and Ranger-Moore 1991; Ridgeway 2006; Rytina and Morgan 1982; Skvoretz 1983).

The theory of intergroup relations has been recognized for its formal character and deductive fertility, features that are important criteria for the assessment of scientific theory (Fararo 2001a: ch. 1). Originally, the theory was presented by Blau in a natural language, but its postulates, or axioms, were explicitly stated and theorems were derived from them using deductive reasoning. Elaborations by Thomas J. Fararo and John Skvoretz (Fararo 1981, 1989; Fararo and Skvoretz 1984, 1989; Skvoretz 1983; Skvoretz and Fararo 1986) resulted in a restating of the original theory in a mathematical format that turned out to have a number of advantages. First, the formalization by Fararo and Skvoretz provided specific and unambiguous definitions of the key theoretical concepts. Second, the theoretical relationships became more precise and rigorous and thus more falsifiable. Third, the theory’s deductive fertility improved, enabling the authors to derive new theorems and propositions that could not have been deduced from the original formulation by Blau; instances of each of these developments are given below.

The studies by Fararo and Skvoretz also revealed a major scope restriction that was implicitly assumed by Blau in the theory of intergroup relations (Fararo 1981; Skvoretz 1983). The restriction pertains to the homophily bias, a fundamental concept in the theory. Homophily means a tendency for people to associate with members of their in-group, because of which social ties between members of the same group tend to occur more frequently than if they were formed by chance (McPherson, Smith-Lovin, and Cook 2001). That is, social ties, such as friendship or marriage, tend to be concentrated within groups and this makes social integration problematic. More specifically, the homophily bias increases within-group integration and decreases between-group societal integration.

However, the homophily bias is not given any particular explanatory role in the theory of intergroup relations. The bias is assumed to operate, but only because such an assumption adds significance to the problem of social integration through the between-group ties (Skvoretz 1991). Other than that, no special attention is given to this concept: The effect of variability in the strength of the homophily bias on the rate of intergroup ties or its possible interactions with heterogeneity are not investigated in the theory of intergroup relations. Instead, the strength of the bias is assumed to be the same in all the groups along the given dimension of differentiation. This assumption, however, is highly debatable in light of the research on marital homogamy (Domański and Przybysz 2007; Kalmijn 1998; Schwartz and Mare 2005; Smits 2003; Smits, Ultee, and Lammers 1998; Ultee and Luijkx 1990).
Fararo and Skvoretz added another simplifying assumption to the theory’s scope conditions (Foschi 1997; Walker and Cohen 1985), namely, the assumption that each person in the population has, on average, the same number of associates, regardless of the group or category that he or she belongs to (Fararo 1981; Skvoretz 1983, 1991). This latter assumption is also restrictive in that it cannot be true for all types of social relations, but in the case of some of them it comes close to being satisfied, with marriage as a primary example of such a relation.1 However, some studies ask subjects to name a specified number of associates, so this assumption is not as arbitrary as it appears to be at first.

In the present study, I rely on the formalization by Fararo and Skvoretz, rather than on the original formulation by Blau, to test the heterogeneity theorem. The theorem has two versions, simple and complex. The former refers to characteristics that are independent of other characteristics and it predicts that the rate of intergroup ties along a characteristic is positively related to the population’s heterogeneity with respect to that characteristic (Blau 1994). The complex version, in turn, applies to characteristics that are associated—or, as Blau puts it, “consolidated”—with other characteristics and it states that the effect of heterogeneity is limited by the consolidation, because the latter is predicted to be negatively related to the rate of intergroup ties (ibid.).

**Dependent and independent variables**

Underlying the theory of intergroup relations is a “distributional” conception of social structure, which is characteristic of macrosociology (see, e.g., Fararo 2001b; Fararo and Kosaka 2003). In that conception, social structure is defined in terms of the distribution of members of a population among social groups or categories, each group comprising individuals who are identical in regard to a given characteristic. The groups constitute equivalence classes. Most commonly, there are multiple characteristics or dimensions of social differentiation, each of which gives rise to a separate set of such equivalence classes. To speak of a society’s occupational or income structure is to speak, essentially, of a distribution of members of that society along dimensions of occupation and income, respectively. In the theory of intergroup relations, the focus is not so much on the distribution as such, but on its properties. If a characteristic gives rise to differentiation that is purely classificatory (or nominal), so that no “natural” ordering of the categories exists, the structural property of interest is heterogeneity. In turn, if the differentiation is of a graduated (or stratified) character, so that the categories are ordered in a nonarbitrary way, then the property of interest is inequality (Blau 1994; Fararo and Skvoretz 1989). My focus in this study is on the former quantity.

**Population Heterogeneity**

Suppose there is a population C of size N that is divided into m distinct categories by an equivalence relation in regard to a nominal variable A. Let $N_i$ be the number
of members of $C$ who belong to category $i$ and let $p_i = N_i/N$ be the proportion in that category. Heterogeneity of $C$ in regard to $A$ is defined as the probability that two individuals selected randomly from $C$ are in different categories (Blau 1994: 13–14). Formally, the heterogeneity is given by the following formula (Lieberson 1969):

$$H(A) = 1 - \sum_{i=1}^{m} p_i^2 = \sum_{i=1}^{m} p_i - \sum_{i=1}^{m} p_i^2 = \sum_{i=1}^{m} p_i (1 - p_i).$$  \hfill (1)

$H(A)$ varies from 0 to $1 - 1/m$. The lower bound corresponds to a situation in which all the members of $C$ belong to a single category, so that there is no differentiation with respect to $A$, while the upper bound corresponds to a situation in which the population is distributed evenly among all categories, so that the proportion in each category is equal to $1/m$.

**Relational Heterogeneity**

In all theorems to be presented in this section $H(A)$, the population’s heterogeneity with respect to $A$ is the basic independent variable, while the dependent variable is the population’s relational heterogeneity with respect to that characteristic, denoted by $H^r(A)$, or the probability that associates in a relation of a given type are in different categories. In other words, both $H(A)$ and $H^r(A)$ may be operationalized in terms of a proportion of intergroup pairs, but the former takes into account all possible pairs of members of $C$, while the latter only the pairs that satisfy a certain social relation, $R$. For instance, if the characteristic of interest is gender, and the relation of interest is friendship, then $H(A)$ indicates gender heterogeneity of the population, while $H^r(A)$ measures gender heterogeneity of friends, or the probability that friends are of different gender. If half the population under study is female and 90 percent of all friendships are between people of the same gender, then $H(A)$ in that population equals 0.5, while $H^r(A)$ equals 0.1.

**Consolidation of Characteristics**

Unlike the concepts of population and relational heterogeneity, that of consolidation is a complex concept. By consolidation, Blau (1994: 14–15) originally meant a situation in which differences with respect to one characteristic correspond to or overlap with differences along other characteristics to a greater or lesser extent. In the language of statistics, consolidation is simply association or correlation, so that “states” of one characteristic covary with states of other one(s). Although Blau defined consolidation broadly to span multiple dimensions of social differentiation (Blau 1994; Blau, Becker, and Fitzpatrick 1984), the present study discusses the effect of consolidation of only two such dimensions. This special case is reasonably simple to provide a starting point for working out more complex cases. But even the simplest case of consolidation of only two dimensions has a number of varieties to consider.
The varieties involve an associational bias that is activated by differentiation along a characteristic in the social relation of a given kind. Homophily, as discussed above, is one example of such a bias. Another example is heterophily, the opposite of homophily, which means a preference for associating with members of the out-group, or a tendency for people to choose dissimilar others as associates (Heckathorn and Rosenstein 2002; McPherson, Smith-Lovin, and Cook 2001; Skvoretz 1983). The extension of the theory of intergroup relations to cover the heterophily bias is due to Skvoretz (1983). Blau’s (1977) original formulation, as well as his later work on the theory of intergroup relations (Blau, Becker, and Fitzpatrick 1984; Blau and Schwartz 1984), was limited to the analysis of the homophily bias only.

With only two characteristics, there are three cases to consider:

- consolidation of two characteristics along which homophily in social relations of a given kind is activated;
- consolidation of two characteristics such that differentiation along one gives rise to homophily in social relations of a given kind, while differentiation with respect to the the other gives rise to heterophily; and
- consolidation of two characteristics along which heterophily in social relations of a given kind exists.

Further cases can be obtained by crossing the type of bias with the type of characteristic (nominal vs. graduated), but the present analysis is limited to nominal dimensions of social differentiation, so only the three cases listed above will be considered. They are covered in detail only in the extended version of the theory, developed by Skvoretz (1983), because, as mentioned before, Blau ignored the problem of heterophily; consequently, his version studies only the first of the three cases (Blau 1994).

The differences between the three situations concern:

1. specific measures of consolidation; and
2. the effect of consolidation of relational heterogeneity.

As for the former issue, it is interesting to note that the measures of consolidation used by Blau and his associates in their original research on the theory of intergroup relations were chosen arbitrarily. Skvoretz (1983) showed that there is no need for the measures to be chosen arbitrarily because a specific measure can be derived from the mathematical formulation of that theory. In the case of two homophilous characteristics, the measure in question is the familiar Goodman–Kruskal $\tau$. In the case of two heterophilous dimensions, the measure is Skvoretz’s $\sigma$, a statistic that has no precedent in the statistical literature. In the “mixed” case, when one characteristic is homo- and the other heterophilous, both $\tau$ and $\sigma$ are used (for a proof, see Skvoretz 1983).

Because Goodman–Kruskal $\tau$ is familiar to sociologists and is discussed in a typical introduction to statistics, I do not present its properties here. But because Skvoretz’s $\sigma$ is not as commonly known, I present its definition and interpretation.
Suppose that there are two nominal characteristics, $A$ and $B$, both of which are heterophilous in the sense that differentiation with respect to these characteristics activates the heterophily bias in a social relation of a given type. The equivalence relation in regard to $A$ divides the population into $m$ distinct categories, while the equivalence relation in regard to $B$ divides it into $n$ categories. Let $p_{ij}$ denote the proportion of the population in category $ij$ of the joint distribution of both characteristics, and let $p_i$ and $p_j$ be the proportions in categories $i$ of $A$ and $j$ of $B$, respectively. Then, Skvoretz’s $\sigma$ is given by the following equations (Skvoretz 1983: 374):

\[
\sigma_{A|B} = \frac{\left[1 - H(A)\right] - \sum_{i=1}^{m} \sum_{j=1}^{m} p_{ij} \frac{p_i - p_{ij}}{1 - p_{ij}}}{1 - H(A)},
\]

\[
\sigma_{B|A} = \frac{\left[1 - H(B)\right] - \sum_{i=1}^{m} \sum_{j=1}^{m} p_{ij} \frac{p_j - p_{ij}}{1 - p_i}}{1 - H(B)}.
\]

The complex expression in the numerator of equation (2) is equal to the probability that two people are equivalent in regard to $A$, given they are different in regard to $B$. In other words, this expression tells us how likely we are to pick an $A$-equivalent pair from among $B$-dissimilar pairs. The numerator in equation (2) subtracts that probability from the reciprocal of the population’s heterogeneity with respect to $A$, or the probability of picking at random an $A$-equivalent pair from among all pairs in the population under consideration; $\sigma_{A|B}$ compares this difference to a maximum possible difference between the two quantities. Equation (3) is interpreted similarly. Just as Goodman–Kruskal $\tau$, Skvoretz $\sigma$ is asymmetric, ranges from 0 to 1, and could be interpreted in terms of proportional reduction in error (Skvoretz 1983).

**Social Structural Theorems**

Having presented basic concepts of the theory of intergroup relations, I now turn to discussing relationships between them. As mentioned before, my interest here is in two general classes of these relationships, simple and complex. In the former case, relational heterogeneity depends only on population heterogeneity, but the latter case also includes the effect of consolidation. Within each class, specific variants of the relationships are proposed, depending on the type of bias that is activated along the given dimension of differentiation.

**Simple Heterogeneity Theorem**

In Blau’s original formulation, the simple heterogeneity theorem states that heterogeneity promotes intergroup relations by improving opportunities for contact between
members of different categories (Blau 1994: 31). In other words, an increase in heterogeneity increases the probability of an encounter between members of different categories. The mere fact that they are more likely to meet by chance when the heterogeneity increases makes it more likely that they will form a deep, long-lasting intimate social relationship. This proposition reflects a major assumption underlying the theory of intergroup relations: that the events that occur in the course of social life depend, at least in part, on the opportunities for these events to occur (Blau 1994; Skvoretz and Mayhew 1988). Fararo (1981: 146) and Skvoretz (1983: 361–63) translated this proposition into the following mathematical equations:

\[ H_R(A) = \begin{cases} (1 - \epsilon_A)H(A), & \text{when } A \text{ is homophilous}, \\ \delta A = (1 - \delta)H(A), & \text{when } A \text{ is heterophilous}. \end{cases} \] (4)

In equation (4) \( \epsilon_A \) denotes the strength of the homophily bias with respect to \( A \); it may be interpreted in terms of the probability that the homophily bias is active when a particular tie is being formed. Alternatively, \( \epsilon_A \) may be seen as the proportion of the population driven by the in-group preference in choosing associates. This is to say that there is some hypothetical latent trait that divides the population into two subpopulations, with one of them comprising members who stick to the in-group preference when selecting an associate and the other comprising those who are indifferent as to whether or not their associates are in the same category. Consequently, the former select their associates “purposively” from their in-group, while the latter select their associates “randomly” from among all the members of the population. Quite obviously, it is only the latter whose choices are affected by the population’s heterogeneity. Thus, the strength of the homophily bias is a parameter that mitigates the effect of heterogeneity on the rate of intergroup relations. The term \( \delta_A \) in equation (4) is interpreted accordingly as the proportion of the population that stays with the out-group preference in selecting an associate for a social relation.

If social ties with respect to \( A \) were formed on a purely random basis, relational heterogeneity would equal population heterogeneity. Homophily and heterophily biases make the former depart from randomness. The magnitude of that departure is called salience of a characteristic. More specifically, let \( H_R(A) \) be the amount of relational heterogeneity that would be observed if social ties were formed randomly with respect to \( A \). That ties form randomly with respect to a characteristic means there are no systematic tendencies affecting the likelihood of a tie between two people regardless of whether they belong to the same category. That is, if social relations of a given type form randomly with respect to \( A \), \( \epsilon_A = \delta_A = 0 \), so that \( H_R(A) \) equals \( H(A) \). Salience of \( A \), denoted by \( S(A) \), is simply the magnitude of the difference between \( H_R(A) \) and \( H(A) \), that is, \( S(A) = |H_R(A) - H(A)| \). After appropriate substitutions and manipulations, we arrive at:

\[ S(A) = \begin{cases} \epsilon_A H(A), & \text{when } A \text{ is homophilous}, \\ \delta A[1 - H(A)], & \text{when } A \text{ is heterophilous}. \end{cases} \] (5)
An interesting implication for the effect of heterogeneity on the rate of intergroup ties follows from equation (5) (Skvoretz 1983). In the case of a homophilous characteristic, the greater its salience, the less intergroup ties there are relative to chance expectations. Hence, when heterogeneity along \( A \) increases, relational heterogeneity along \( A \) increases in the absolute sense, but declines in the relative sense. In Figure 1, \( Q_1 \) is clearly above \( P_1 \), so there are absolutely more intergroup ties when \( H(A) \) is greater. At the same time, \( Q_0Q_1 \) is longer than \( P_0P_1 \), meaning that salience of \( A \) in the former case is greater, so that there are relatively less intergroup ties.

If \( A \) is heterophilous rather than homophilous, it is clear from equation (4) that \( H_r(A) \) is an increasing function of \( H(A) \) as well. Further, because \( A \) is heterophilous, intergroup ties are now statistically overrepresented, meaning that they occur with a greater frequency than they would if association was independent of category membership with respect to \( A \). Thus, salience of a heterophilous characteristic means there are more intergroup ties than would result from random association.

But equation (5) shows that salience of a heterophilous characteristic declines with heterogeneity, so that there are less intergroup ties relative to chance expectations when heterogeneity is greater. Again, this point is illustrated in Figure 1: We see that $Q_2$ is above $P_2$, which means that increasing population heterogeneity produces more integroup ties in the absolute sense. However, $Q_0Q_2$ is clearly shorter than $P_0P_2$, which means that salience is smaller in the former case, so there are fewer intergroup ties in the relative sense. These considerations show that the effect of heterogeneity is quite general and independent of the type of bias that is actually at play (Skvoretz 1983).

**Extension of the Simple Heterogeneity Theorem**

Let us now turn to a set of theorems concerning the effect of social structural differentiation on the rate of intergroup ties in a situation when:

- there are two dimensions of differentiation, $A$ and $B$, or characteristics that divide the population into two sets of mutually disjunctive categories, and
- these two characteristics are correlated, or “consolidated.”

In such a situation, ties can be intergroup with respect to $A$, $B$, or both. To illustrate, if the social relation of interest is friendship, and the characteristics dividing the population are gender and occupation, then one may be interested in the rate of friendships between people who

- are of different gender,
- belong to different occupations, or
- differ in terms of both gender and occupation.

This last case, however, is mostly of theoretical interest and is not explored in this study.

I start by presenting theorems that apply to the consolidation of two homophilous characteristics, a case to which Blau limited his analysis of the effects of consolidation in the original formulation of the theory of intergroup relations. Subsequently, I cover situations involving heterophilous characteristics. However, I do not go into details of derivation of the theorems discussed in this subsection. Readers interested in these details are encouraged to consult Skvoretz (1983) or Karpiński (2008).

**Two Homophilous characteristics**

When there are two characteristics, $A$ and $B$, both of which give rise to homophily in a social relation, then the following equations apply:

$$H_r(A) = (1 - \varepsilon_A)H(A) - \varepsilon_B(1 - \varepsilon_A)H(A)\tau_{AB}.$$  

(6)
where \( \tau_{AB} \) and \( \tau_{BA} \) denote Goodman–Kruskal \( \tau \), an asymmetric measure of association between two nominal characteristics that is interpreted in terms of proportional reduction in error, as mentioned earlier. The equations confirm Blau’s informal theorizing (see Blau 1994: chs. 2 and 3) that consolidation reduces the rate of intergroup ties because when two characteristics are strongly consolidated, associates who belong to the same category along one characteristic are very likely to be members of the same category along the other characteristic as well. However, equations (6) and (7) add certain qualifications to that proposition.

First, the strength of the homophily bias has to be nonzero along both \( A \) and \( B \) for consolidation to affect relational heterogeneity. This can easily be shown by setting one of the homophily-bias parameters to zero. For instance, if we let \( \varepsilon_B = 0 \), then equation (6) reduces to equation (4) for a homophilous characteristic, and \( \tau_{AB} \) will have no effect on relational heterogeneity even if the consolidation is at its maximum. Second, and more important, consolidation does not have an independent main effect on relational heterogeneity; instead, it affects the latter variable only in interaction with population heterogeneity (Skvoretz 1983; Skvoretz and Fararo 1986). This is an important result with serious implications for empirical tests of the theory of intergroup relations.

A curious implication concerning salience follows from equations (6) and (7). According to Blau’s original conception, salience derives from homophily. That is, a characteristic is salient to the extent that homophily with respect to it exists in a population under study and in a given social relation. This view is supported by equation (5), because it implies that there is no salience when there is no homophily bias. But in the case of consolidation of two homophilous characteristics, the link between homophily and salience is less obvious. Using the definition of salience proposed above, we obtain:

\[
S(A) = |H_{RA}(A) - H_{RA}(A)| = \varepsilon_A H(A) + \varepsilon_B (1 - \varepsilon_A) H(A) \tau_{AB},
\]

(8)

\[
S(B) = |H_{RB}(B) - H_{RB}(B)| = \varepsilon_B H(B) + \varepsilon_A (1 - \varepsilon_B) H(B) \tau_{BA}.
\]

(9)

What is so curious about this implication is that it shows that a characteristic may remain salient even if there is no homophily along it! That is, a characteristic may appear to be salient if it is consolidated with another characteristic that is homophilous in a given social relation. This can be seen by setting \( \varepsilon_A \) to zero. Equation (8) can then be reduced to:

\[
S(A) = \varepsilon_B H(A) \tau_{AB}.
\]

There would be fewer intergroup ties along \( A \), relative to chance expectations, even though there is no homophily with respect to \( A \)! But because there are fewer intergroup ties, \( A \) will appear to be homophilous. The departure from randomness
results, however, not from \( A \) being homophilous, but from its being consolidated with \( B \), which is genuinely homophilous.

**One Homophilous and One Heterophilous Characteristic**

The models discussed in the preceding paragraph can be applied to a variety of social relations and characteristics, such as friendship in a group differentiated along gender and occupation, marriage in a community divided by ethnic background and religion, or sociometric choices among schoolchildren whose tastes differ in regard to music and sport. In turn, in this subsection, I introduce models that are most readily applicable to marriage with respect to some homophilous characteristic (say, occupation) in a society that allows only for marriages between individuals of different sex.\(^3\) I start the introduction using more general terms, however.

Suppose there are two characteristics, \( A \) and \( B \), that divide the population, and that the former is homophilous, while the latter is heterophilous. The strength of the in-group bias along \( A \) is denoted by \( \varepsilon_A \) and the strength of the out-group bias along \( B \), by \( \delta_B \). The rates of intergroup ties along \( A \) and \( B \) were shown by Skvoretz (1983) to be equal to:

\[
H_r(A) = (1 - \varepsilon_A)H(A) + \delta_B(1 - \varepsilon_A)[1 - H(A)]\sigma_{a|B}, (10)
\]

\[
H_r(B) = \delta_B + (1 - \delta_B)H(B) - \varepsilon_A(1 - \delta_B)H(B)\tau_{a|B}. (11)
\]

All that was said about the effect of consolidation on relational heterogeneity in the preceding paragraph holds for the relationship in equation (11). But in equation (10), consolidation, as measured by Skvoretz \( \sigma \), defined by equations (2) and (3), has an opposite effect on the rate of intergroup ties.

Recall that earlier it was shown that the effect of population heterogeneity, \( H(A) \), was quite general and independent of whether differentiation with respect to \( A \) gave rise to homophily or heterophily. But equation (10) implies that the same cannot be said about the effect of consolidation: it is not as general as the effect of heterogeneity and clearly depends on the types of characteristics that are consolidated. To be more precise, if one is interested in deriving the amount of relational heterogeneity with respect to a given characteristic (let us call it “focal characteristic”) that is consolidated with another characteristic, then the effect of their consolidation depends on whether differentiation in regard to the other characteristic gives rise to homophily or heterophily.

In equation (10), the focal characteristic, \( A \), is homophilous, while the other, \( B \), is heterophilous. In this case, consolidation of the focal characteristic with the other one, as measured by \( \sigma_{a|B} \) has a positive effect on the rate of intergroup ties. Further, after rearranging terms in that equation we arrive at:

\[
H_r(A) = (1 - \varepsilon_A)H(A) + \delta_B(1 - \varepsilon_A)\sigma_{a|B} - \delta_B(1 - \varepsilon_A)H(A)\sigma_{a|B},
\]

which shows that consolidation affects the rate of intergroup ties, both directly
and indirectly, through its interaction with heterogeneity. In turn, in equation (11),
the focal characteristic, \( B \), is heterophilous, but the other \( A \), is homophilous and 
consolidation of the former with the latter has exactly the same effect on relational 
heterogeneity as described in the preceding subsection (see comments to equations [6] and [7]).

We now turn to the interpretation of equations (10) and (11) in terms of marriage 
with respect to occupation. Because the population of interest is one in which only 
heterosexual marriages are allowed, we need to consider the effect of distribution of 
the population along both occupation and gender, the former being a homophilous 
characteristic, and the latter—heterophilous. Also, the strength of the out-group 
bias along gender is assumed to be at its maximum: \( \delta_B = 1 \), so equations (10) and 
(11) reduce to:

\[
H_r(A) = (1 - \varepsilon_A)H(A) + (1 - \varepsilon_A)[1 - H(A)]\sigma_{A|B}, 
\]

\[
H_r(B) = 1. 
\]

Let us focus on equation (12). It shows that whenever \( \sigma_{A|B} \) is greater than zero, 
the consolidation of occupation with gender causes marriages between individu-
als from different occupational categories to occur more frequently than if the two 
characteristics were independent. That is, if women are “overrepresented” in some 
categories, and men in others, some people are “forced,” by the very structure of 
the population, to search for a spouse outside their own occupational category even 
if they preferred the spouse to be a member of the same occupation. Put in another 
way, a “raw” estimate of the rate of intermarriage in regard to occupation (or any 
other homophilous characteristic, for that matter) is likely to overestimate the actual 
rate. Thus, an adjustment is needed to obtain a correct estimate. Rearranging the 
terms in equation (12) gives:

\[
\frac{H_r(A)}{H(A)} = \frac{H(A)}{[1 - H(A)]\sigma_{A|B}} = (1 - \varepsilon_A)H(A). 
\]

For \( H(A) > 0 \) and \( \sigma_{A|B} > 0 \), the factor on the left-hand side of equation (14) is 
less than 1, and shows what fraction of the actual rate of intermarriage with respect 
to \( A \) results from the effect of heterogeneity along that characteristic and not from 
its consolidation with \( B \).

This problem of consolidation of a homophilous focal characteristic with gender 
in the case of marriage was recognized by Blau and his associates as well (Blau 
1994; Blau, Becker, and Fitzpatrick 1984; Blau, Blum, and Schwarts 1982; Blau 
and Schwartz 1984; Rytina et al. 1988). However, they did not approach it as a 
thoretical problem. Instead, they only proposed an ad hoc correction for the “ex-
cess” of the rate of intermarriage resulting from the consolidation, but Skvoretz’s 
(1983) extension of the theory of intergroup relations shows that such arbitrary 
decisions are not necessary, because a solution of this problem follows logically 
from the mathematical formulation of the theory’s postulates.
Two Heterophilous Characteristics

Finally, there is the case of consolidation of two heterophilous characteristics. However, this particular case is interesting only theoretically, because heterophilous characteristics are not as common as homophilous ones (McPherson, Smith-Lovin, and Cook 2001). Because of space limitations, I omit presentation of the mathematical equations appropriate for this case. Suffice it to say that the effect of consolidation on relational heterogeneity in this case is the same as in equation (10); that is, relational heterogeneity is an increasing function of consolidation, as measured by Skvoretz’s $\sigma$, and is affected by it both directly and indirectly, through consolidation’s interaction with heterogeneity. Readers interested in a more detailed discussion of this case are encouraged to consult Skvoretz (1983).

Data, Variables, and Analytical Procedures

Data

In order to evaluate the models presented in the preceding section, I use data from the first wave of the panel study Social Structure in Poland, 1988–2008 (hereafter, POLPAN), which was carried out in 1988 on a probability sample of men and women in the age range of twenty-one to sixty-five with 5,817 subjects. A detailed description of the panel study, its methodology and findings are presented in Slomczynski et al. (1989), Slomczynski (2000, 2002), and Slomczynski and Marquart-Pyatt (2007).

Variables

In 1988, one of the thematic modules in the POLPAN questionnaire contained items asking respondents about their social life: the number of close friends they have and how long they have known them; whether the close friends are their classmates, colleagues, neighbors, or relatives; and whether their friends know each other. In turn, some questions concerning the characteristics of the respondents’ closest friend followed, such as their gender, age, level of education, or occupation. Further, another section of the questionnaire asked subjects about their own age, education, and occupation and, if they were married, about these same characteristics of their spouse. I use responses to these items to estimate rates of intermarriage along age, education, and occupation, as well as rates of intergroup friendship ties along gender, age, education, and occupation. The unit of analysis is voivodship; for each voivodship, estimates of relational heterogeneity, population heterogeneity as well as consolidation are computed. Because there were forty-nine voivodships in 1988, there are forty-nine observations for each variable of interest.

In the 1988 wave of POLPAN, occupation was coded using the Polish Social Classification of Occupations, as developed by Pohoski and Slomczynski (1978).
Following suggestions in Domanski (2004: ch. 9), the original classification was transformed into broad social-occupational categories: (1) intelligentsia, (2) routine nonmanual workers, (3) small proprietors, (4) skilled manual workers, (5) unskilled manual workers, and (6) agricultural laborers. For the sake of the present analysis, the first three categories were merged, and so were categories 4 and 5, resulting in a three-category classification: (1) nonmanual workers, (2) manual workers, and (3) farm workers.

As for the degree of education, nine levels were originally distinguished in the POLPAN study, but for the sake of the present analysis the original classification was replaced by a three-category classification: (1) less than high school, (2) high school, and (3) more than high school.

Finally, as for age, it is recorded in years in the POLPAN data set, but for the purpose of this analysis it was transformed into a categorical variable with five states: (1) less than twenty-five years old, (2) twenty-five to thirty-four years old, (3) thirty-five to forty-four years old, (4) forty-five to fifty-four years old, and (5) fifty-five years old or more.

Thus, there are four characteristics for which estimates of the theoretical variables can be obtained using the POLPAN data: education, occupation, gender, and age.

**Analytical Procedures**

In order to estimate the models presented in the previous section, I use a method proposed by Skvoretz (Fararo and Skvoretz 1989; Skvoretz 1983). Let $H_j(a)$ denote the amount of heterogeneity with respect to $a$ in the $j$th voivodship. Let $O_j$ be the number of intergroup ties of a given type with respect to $a$ in the $j$th voivodship and let $I_j$ be the number of in-group ties of this type along that characteristic.

Now we can look at the process of formation of social ties as a series of Bernoulli trials, with an intergroup tie interpreted as a “success” and an in-group tie interpreted as a “failure.” Suppose we are interested in evaluating equation (4) for a homophilous characteristic. According to that equation, the probability of a success for voivodship $j$ is equal to $P(s) = (1 - \varepsilon_a)H_j(A)$, and the probability of a failure for that voivodship is given by $P(f) = 1 - P(s) = 1 - (1 - \varepsilon_a)H_j(A)$. The likelihood of exactly $O_j$ successes and $I_j$ failures is given by:

$$L_j(\varepsilon_a) = [(1 - \varepsilon_A)H_j(A)]^{O_j} [1 - (1 - \varepsilon_A)H_j(A)]^{I_j}.$$  

(15)

The product of the likelihoods for all voivodships gives a likelihood function that will be used to estimate $\varepsilon_A$:

$$L(\varepsilon_a) = \prod_{j=1}^{w} L_j(\varepsilon_a) = \prod_{j=1}^{w} [(1 - \varepsilon_A)H_j(A)]^{O_j} [1 - (1 - \varepsilon_A)H_j(A)]^{I_j}.$$  

(16)

The estimate of the strength of homophily along $A$ is the value of $\varepsilon_A$ that maximizes equation (16). However, the calculations are easier to perform when it is a
natural logarithm of equation (16) that is subject to maximization. Also, because $\varepsilon_A$ is interpreted as a proportion, it has to satisfy the requirement $0 \leq \varepsilon_A \leq 1$. In order to find the value of $\varepsilon_A$ that maximizes equation (16) under this constraint, I employed the function optim() in R (R Development Core Team 2009).

The fit of the model is assessed by comparing it with the baseline model. That is, let $\ln[l(0)]$ be the value of the natural logarithm of equation (16) under the assumption of the null model, or when $\varepsilon_A = 0$ holds, and let $\ln[l(\hat{\varepsilon}_A)]$ be the value of that function at the estimated strength of the homophily bias. Then, the statistic:

$$-2[\ln L(0) - \ln L(\hat{\varepsilon}_A)],$$

has a chi-square distribution with one degree of freedom. If the value of equation (17), which is referred to as deviance, is large enough, then the simple heterogeneity model fits the data significantly better than the null model. Evaluating the complex equations (6) and (7) follows the same procedure, with $P(S)$ and $P(F)$ specified accordingly. When the fit of the complex models is evaluated, it involves comparisons with the simple model rather than the baseline.

Results of the Analysis

Friendship

Table 1 contains estimates of the strength of the in-group preferences along the various dimensions under study as well as the corresponding log likelihoods or natural logarithms of the likelihood function (16) at these estimated strengths. The table is divided into four “panels,” each of which gives information as to the effects of heterogeneity and consolidation along a particular characteristic. To illustrate, the top panel pertains to differentiation and association with respect to degree of education. The baseline model, assuming that ties form irrespective of associates’ educational background, has $\ln[L(0)] = -3,563.6$, which is obtained by substituting $\varepsilon_A = 0$ in equation (16), along with the estimates of educational heterogeneity and the number of intergroup and in-group ties in each voivodship. In turn, the log likelihood for the simple model equals $\ln[L(\hat{\varepsilon}_A)] = -3,059.5$, and the estimate of the strength of homophily is $\hat{\varepsilon}_A = 0.42$, which is to say that 42 percent of the time closest friends are selected purposefully from among the members of one’s own educational category and 58 percent of the time they are selected randomly, regardless of their education.

The top panel of Table 1 also displays results regarding three further models for association along education. They are all complex models that assume education to be the focal characteristic and occupation, gender, and age, respectively, to be the other characteristic. In the first of the three complex models, the one assuming consolidation of education with gender, the strength of homophily bias along education is $\hat{\varepsilon}_A = 0.42$, while the estimate of the strength of the bias along gender is zero, suggesting lack of in-group preference with respect to gender. Similarly,
modeling social association with respect to education as dependent upon both educational heterogeneity and education’s consolidation with gender yields $\hat{\varepsilon}_a = 0.408$ and $\hat{\varepsilon}_b = 0.249$, so 40 percent of the time subjects “fall prey” to homophily along education, and a quarter of the time they succumb to in-group preference with respect to age. Clearly, educational homophily turns out to be stronger than age homophily. According to these estimates, then, the homophily bias along

<table>
<thead>
<tr>
<th>Model</th>
<th>Log likelihood</th>
<th>$\hat{\varepsilon}_{\text{focal}}$</th>
<th>$\hat{\varepsilon}_{\text{other}}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Focal characteristic: degree of education</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Null</td>
<td>$-3,563.6$</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Degree of education</td>
<td>$-3,059.5$</td>
<td>$0.420$</td>
<td>—</td>
</tr>
<tr>
<td>Complex models</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Education by gender</td>
<td>$-3,059.5$</td>
<td>$0.420$</td>
<td>$0.000$</td>
</tr>
<tr>
<td>Education by age</td>
<td>$-3,009.2$</td>
<td>$0.408$</td>
<td>$0.249$</td>
</tr>
<tr>
<td>Education by occupation</td>
<td>$-3,058.2$</td>
<td>$0.353$</td>
<td>$0.354$</td>
</tr>
<tr>
<td>Focal characteristic: occupation</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Null</td>
<td>$-3,069.0$</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Occupation</td>
<td>$-2,133.6$</td>
<td>$0.569$</td>
<td>—</td>
</tr>
<tr>
<td>Complex models</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Occupation by education</td>
<td>$-2,133.1$</td>
<td>$0.507$</td>
<td>$0.429$</td>
</tr>
<tr>
<td>Occupation by gender</td>
<td>$-2,133.6$</td>
<td>$0.569$</td>
<td>$0.000$</td>
</tr>
<tr>
<td>Occupation by age</td>
<td>$-2,097.5$</td>
<td>$0.543$</td>
<td>$0.617$</td>
</tr>
<tr>
<td>Focal characteristic: gender</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Null</td>
<td>$-3,502.5$</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Gender</td>
<td>$-1,612.8$</td>
<td>$0.807$</td>
<td>—</td>
</tr>
<tr>
<td>Complex models</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gender by education</td>
<td>$-1,612.8$</td>
<td>$0.806$</td>
<td>$0.248$</td>
</tr>
<tr>
<td>Gender by occupation</td>
<td>$-1,612.3$</td>
<td>$0.792$</td>
<td>$0.794$</td>
</tr>
<tr>
<td>Gender by age</td>
<td>$-1,592.7$</td>
<td>$0.802$</td>
<td>$0.437$</td>
</tr>
<tr>
<td>Focal characteristic: age category</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Null</td>
<td>$-5,121.9$</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Age</td>
<td>$-3,424.8$</td>
<td>$0.502$</td>
<td>—</td>
</tr>
<tr>
<td>Complex models</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age by education</td>
<td>$-3,424.8$</td>
<td>$0.502$</td>
<td>$0.000$</td>
</tr>
<tr>
<td>Age by occupation</td>
<td>$-3,424.8$</td>
<td>$0.502$</td>
<td>$0.000$</td>
</tr>
<tr>
<td>Age by gender</td>
<td>$-3,424.8$</td>
<td>$0.500$</td>
<td>$0.503$</td>
</tr>
</tbody>
</table>

Table 1

Results of Fitting Heterogeneity Models to Data on Friendship Patterns
education is a bit weaker than the simple model suggests, and the interpretation is that some part of this bias is induced by education’s consolidation with age. Put in another way, these estimates suggest that because friends tend to be of similar ages and that education is consolidated with age, there appears to be more educational homophily than there actually is. The estimates for the remaining models are interpreted similarly.

Let us now look at the figures presented in Table 2. These figures report on results of checking the fit of the models listed in Table 1. For each model, except the null models, the fit statistic is computed, using equation (17), comparing that model to a simpler one: a simple model is compared with a null one, and a complex model with a simple one. As mentioned, in each case the fit statistic, or deviance, has a chi-square distribution with a single degree of freedom. Table 2 lists the deviances and their respective \( p \)-values. Whenever the fit statistic has a \( p \)-value less than 0.05, it is concluded that the model being tested fits the data better than the model tested against. To illustrate, the simple model of educational heterogeneity can be said to fit the data better than the baseline model, because it has deviance equal to 1,008.3, which is not likely to have happened by chance. At the same time, the complex model of education’s consolidation with gender fails to fit data any better than the simple model of educational heterogeneity: The deviance for the former model is zero, so there is no improvement in fit.

More generally, the figures in Table 2 lead to the following conclusions:

• the simple heterogeneity model fares much better than the baseline in the case of all four characteristics; but
• the complex model fits the data better than the simple model only in three cases out of twelve, and all three cases involve age as the other characteristic.

Thus, these results seem to give little support to the formal models developed by Fararo and Skvoretz (Fararo 1981, 1989; Fararo and Skvoretz 1984; Skvoretz 1983; Skvoretz and Fararo 1986), and more generally to the theory of intergroup relations as such (Blau 1994), as these models constitute a “translation” of Blau’s informal theorizing into mathematical equations (see, e.g., Skvoretz 1990b). However, the theory could still be defended on the grounds that not all results in Tables 1 and 2 are negative, or inconsistent with the theory’s predictions, because in four cases the estimates of the strength of homophily along the other characteristic are zero. As seen in the discussion above, consolidation has no effect on relational heterogeneity along the focal characteristic if there is no homophily with respect to the other characteristic (see comments to equations [6] and [7]). Notice, however, that defending the theory on these grounds requires “symmetry” in the estimates of the parameters of the complex models. But there is no such symmetry: The strengths of the homophily biases along age and education are estimated at 0.502 and 0, respectively, if age is the focal characteristic, and 0.249 and 0.408 if education is the focal characteristic, while
defending the theory using the above argument would be correct if the estimates were the same in both situations.

This latter remark draws attention to another problem with the estimates in Table 1, namely, their instability. For instance, the strength of the bias along education is estimated at 0.420 in the simple model of educational heterogeneity and 0, 0.248, and again 0 in the three complex models that have education as the other characteristic. The estimates for gender vary even more widely, from 0 to 0.807. So, either the estimation procedure proposed by Skvoretz (Fararo and Skvoretz 1989; Skvoretz 1983) is unreliable (in the sense of measurement reliability) or the complex models are too simple to be able to capture the strength of the biases adequately. That is, if some part of the homophily bias with respect to a characteristic can be “induced” by its correlation with another characteristic, then it is equally likely that the biases along two consolidated characteristics may not be estimated adequately because of their both being consolidated with still a third characteristic that also affects rates of intergroup ties along the focal one. In fact, it is such multidimensional consolidation that Blau had in mind in his original formulation of the theory of intergroup

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>Test</th>
<th>Deviance</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Education</td>
<td>Simple vs. baseline</td>
<td>1,008.3</td>
<td>0.00</td>
</tr>
<tr>
<td>Education by gender</td>
<td>Complex vs. simple</td>
<td>0.0</td>
<td>1.00</td>
</tr>
<tr>
<td>Education by age</td>
<td>Complex vs. simple</td>
<td>100.5</td>
<td>0.00</td>
</tr>
<tr>
<td>Education by occupation</td>
<td>Complex vs. simple</td>
<td>2.5</td>
<td>0.11</td>
</tr>
<tr>
<td>Occupation</td>
<td>Simple vs. baseline</td>
<td>1,870.7</td>
<td>0.00</td>
</tr>
<tr>
<td>Occupation by education</td>
<td>Complex vs. simple</td>
<td>1.0</td>
<td>0.32</td>
</tr>
<tr>
<td>Occupation by gender</td>
<td>Complex vs. simple</td>
<td>0.0</td>
<td>1.00</td>
</tr>
<tr>
<td>Occupation by age</td>
<td>Complex vs. simple</td>
<td>72.3</td>
<td>0.00</td>
</tr>
<tr>
<td>Gender</td>
<td>Simple vs. baseline</td>
<td>3,779.4</td>
<td>0.00</td>
</tr>
<tr>
<td>Gender by education</td>
<td>Complex vs. simple</td>
<td>0.0</td>
<td>0.89</td>
</tr>
<tr>
<td>Gender by occupation</td>
<td>Complex vs. simple</td>
<td>1.0</td>
<td>0.31</td>
</tr>
<tr>
<td>Gender by age</td>
<td>Complex vs. simple</td>
<td>40.1</td>
<td>0.00</td>
</tr>
<tr>
<td>Age</td>
<td>Simple vs. baseline</td>
<td>3,394.4</td>
<td>0.00</td>
</tr>
<tr>
<td>Age by education</td>
<td>Complex vs. simple</td>
<td>0.0</td>
<td>1.00</td>
</tr>
<tr>
<td>Age by occupation</td>
<td>Complex vs. simple</td>
<td>0.0</td>
<td>1.00</td>
</tr>
<tr>
<td>Age by gender</td>
<td>Complex vs. simple</td>
<td>-0.0</td>
<td>1.00</td>
</tr>
</tbody>
</table>
relations (Blau 1994; Blau, Becker, and Fitzpatrick 1984). Extension of the formal models developed by Fararo and Skvoretz to cover consolidation of three, or possibly more, characteristics, is possible, and it may be worthwhile to explore this path in the future.

**Interrmarriage**

Before concluding, I briefly discuss the results of fitting the simple and complex models to data on rates of intermarriage along education and occupation in voivodships. Because the age of the subject’s spouse was not recorded in the first wave of POLPAN, it is not possible to estimate rates of age intermarriage from POLPAN data. Also, because marriage in Poland is recognized as a bond between a man and a woman, relational heterogeneity of marriage with respect to gender is always 1, whatever the amount of gender heterogeneity, so the case of intermarriage with respect to gender is not really interesting. This leaves us with only two characteristics for which appropriate measures of heterogeneity and consolidation can be obtained from the POLPAN data.

As for the tests results discussed in the previous section, recall that the measures of relational heterogeneity of friendship along the characteristics of interest were estimated directly from the data, using the proportion of friendship ties between individuals differing on the characteristic in a given voivodship. However, in the context of marriage such an estimate of relational heterogeneity is not valid if the characteristic of interest is consolidated with gender, because, as equations (10) and (11) made clear, such consolidation increases relational heterogeneity. Thus, the “raw” estimate of the rate of intermarriage overestimates the “true” rate. It is, therefore, necessary to “adjust” the raw rate. Equation (14) offers an adjustment that has the advantage of being theory driven, rather than ad hoc. With the estimates of the adjusted rates of intermarriage, tests of the simple and complex model follow the procedures described earlier.

Table 3 gives estimates of the strength of marital homogamy or homophily in the marriage relationship, with respect to education and occupation. While it is impossible to obtain estimates of rates of intermarriage along age from the POLPAN data, consolidation of each of the two characteristics with age can be estimated, which allows the performance of tests of complex models with age as the other characteristic. Interpretation of the figures in Table 3 is the same as those in Table 1. For instance, the estimate of the amount of educational homogamy is 0.411, roughly the same as that of educational homophily in the friendship relation. In turn, the estimate of the strength of occupational homogamy is 0.143, much less than the corresponding estimate for friendship.

However, the general pattern of findings in the case of intermarriage is the same as in the case of friendship: As Table 4 shows, the simple model fits the data better than the baseline, but the complex model does better than the simple model only in two cases out of four, and they both have age as the other characteristic.
The fact that the complex models with age as the other characteristic consistently dominate corresponding simple models may suggest that age is somehow “fundamental” when it comes to selecting an associate, whether friend or spouse, but this conclusion, however tempting, may well be unfounded if age is consolidated with still a third characteristic. Clearly, further research is necessary to cover such consolidation of more than two characteristics.

### Conclusion and Discussion

The research presented in this study gives only limited empirical support to the formal variant of the theory of intergroup relations; the original theory was developed by Blau (1977; 1994), but the formalization is due to the work of Fararo and Skvoretz (see Fararo 1981; Fararo and Skvoretz 1984, 1989; Skvoretz 1983; Skvoretz and Fararo 1986). As a matter of fact, because predictions derived from the Fararo–Skvoretz models disagree with data on patterns of friendship and marriage in most cases, the present research suggests that the formalization be rejected or at least reformulated.

The models developed by Fararo and Skvoretz have the advantage of being simple and straightforward, but these features come at a cost, as they require two idealizing assumptions to be made. The two assumptions are:

- the strength of the homophily bias is the same in all categories of a given characteristic; and

<table>
<thead>
<tr>
<th>Model</th>
<th>Log likelihood</th>
<th>$\hat{\varepsilon}_{\text{local}}$</th>
<th>$\hat{\varepsilon}_{\text{other}}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Focal characteristic: degree of education</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Null</td>
<td>$-3,322.7$</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Degree of education</td>
<td>$-2,875.6$</td>
<td>0.411</td>
<td>—</td>
</tr>
<tr>
<td>Complex models</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Education by age</td>
<td>$-2,825.8$</td>
<td>0.398</td>
<td>0.286</td>
</tr>
<tr>
<td>Education by occupation</td>
<td>$-2,875.6$</td>
<td>0.406</td>
<td>0.028</td>
</tr>
<tr>
<td>Focal characteristic: occupation</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Null</td>
<td>$-2,039.6$</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Occupation</td>
<td>$-1,993.0$</td>
<td>0.143</td>
<td>—</td>
</tr>
<tr>
<td>Complex models</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Occupation by education</td>
<td>$-1,992.9$</td>
<td>0.118</td>
<td>0.096</td>
</tr>
<tr>
<td>Occupation by age</td>
<td>$-1,954.4$</td>
<td>0.088</td>
<td>0.624</td>
</tr>
</tbody>
</table>
each person in the population initiates and receives the same number of social ties, regardless of the category that he or she belongs to.

Clearly, these assumptions are debatable in light of the research on marital homogamy (Domański and Przybysz 2007; Kalmijn 1998; Schwartz and Mare 2005; Smits 2003; Smits, Ultee, and Lammers 1998; Ultee and Luijkx 1990). This research shows that social categories differ from one another in regard to the strength of the in-group bias. For instance, in the study by Domański and Przybysz (2007), homogamy with respect to education turns out to be the strongest at the extremes of the hierarchy of education. Also, frequencies of marital pairs in particular educational groups fail to satisfy the second of the above simplifying assumptions. In some categories of education, the average number of marital ties exceeds the average for the whole population while in others it falls below the population average.

It is possible to relax the two simplifications, as shown by Skvoretz (1991; see also Karpiński 2008), which results in a number of models that are more accurate descriptively, but lack the beauty and simplicity of the models presented earlier in this chapter. In fact, general social-structural theorems—such as those concerning the effects of heterogeneity and consolidation on the rate of intergroup ties, discussed in this paper—cannot be derived from the more descriptively accurate models proposed by Skvoretz. This illustrates a point made by Douglas D. Heckathorn (1984) that there is often a trade-off between theory’s analytical power, descriptive accuracy, and explanatory scope.

Also, Fararo and Skvoretz’s formalization of the theory of intergroup relations implies that the relationship between the rate of the intergroup ties and heterogeneity is linear. In other words, it implies that the magnitude of the effect of heterogeneity is the same, regardless of whether the heterogeneity is low or high. That is, the formal variants of Blau’s original theorems predict that the increase in the rate of intergroup ties will be the same when population heterogeneity with respect to A

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>Test</th>
<th>Deviance</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Education</td>
<td>Simple vs. baseline</td>
<td>894.3</td>
<td>0.00</td>
</tr>
<tr>
<td>Education by age</td>
<td>Complex vs. simple</td>
<td>99.6</td>
<td>0.00</td>
</tr>
<tr>
<td>Education by occupation</td>
<td>Complex vs. simple</td>
<td>0.0</td>
<td>0.92</td>
</tr>
<tr>
<td>Occupation</td>
<td>Simple vs. baseline</td>
<td>93.2</td>
<td>0.00</td>
</tr>
<tr>
<td>Occupation by education</td>
<td>Complex vs. simple</td>
<td>0.1</td>
<td>0.75</td>
</tr>
<tr>
<td>Occupation by age</td>
<td>Complex vs. simple</td>
<td>77.2</td>
<td>0.00</td>
</tr>
</tbody>
</table>
increases from 0.01 to 0.1 and when it goes up from 0.91 to 1. But this prediction seems problematic as well.

When there is little differentiation with respect to \( A \), increasing population heterogeneity along that characteristic may have hardly any effect on the rate of intergroup relations, because changing the heterogeneity from “almost nil” to “very little” does not necessarily produce a qualitative change in the amount of heterogeneity that is perceived or experienced by members of the community. In other words, the community will still be seen as homogeneous in regard to the characteristic \( A \). But when the differentiation is already very large, increasing it still further may contribute nothing to the frequency of intergroup ties, for basically the same reason: If there is very much differentiation, further increases are not likely to produce a qualitative change in the amount of heterogeneity that members of the given community experience.

Put in another way, there are likely to be two “threshold” levels of population heterogeneity. If the actual level of heterogeneity is below the lower threshold or above the upper one, it will have no bearing on the rate of intergroup relations, because it will be either “too little,” or “too large” to affect the rate at all. Only in-between these two thresholds is population heterogeneity likely to affect relational heterogeneity (and the effect may well be linear). If this conjecture is correct, then future work on the theory of intergroup relations should focus on a method of establishing these thresholds.

As this article shows, the formal version of the theory of intergroup relations fails to adequately predict the patterns in data concerning rates of marriages and friendships between members of different educational and occupational groups. But it would be unwise to completely dismiss the original theory: first, because the theory asks important questions about structural sources of social integration; second, because what fails is not the theory as such, but its particular mathematical representation. So, it is possible to develop the original theory further, but without relying on the mathematical apparatus employed by Fararo and Skvoretz.

Notes

1. However, for marriage to satisfy the assumption men and women have to be represented in equal proportions in each group.

2. Fararo and Skvoretz (Fararo 1981, 1983; Skvoretz 1985, 1990a; Skvoretz, Fararo, and Agneessens 2004) distinguish between two general classes of the associational biases: relational and compositional. The former type pertains to relations among nodes in a network. The latter type, in turn, pertains to attributes of the nodes. Homophily and heterophily are both examples of compositional biases.

3. Certainly, it is not the only possible application or interpretation of the models to be presented. One can easily apply these models to cooperative ties among scientists pursuing an international research project within a scientific discipline. If the project is required to be international, then choices of associates are likely to be heterophilous with regard to nationality, but homophilous with respect to discipline. The marriage interpretation is more commonsensical, however.

5. For an updated version of this classification, see Domański, Slomczynski, and Sawniński (2009).

References

Rytina, Stephen; Peter M. Blau; Terry C. Blum; and Joseph Schwartz. 1988. “Inequality and Intermarriage: A Paradox of Motive and Constraint.” *Social Forces* 66, no. 3: 645–75.


To order reprints, call 1-800-352-2210; outside the United States, call 717-632-3535.